

**UNIVERSITAT POLITÈCNICA DE VALÈNCIA**

**INSTITUTO DE RECONOCIMIENTO MOLECULAR  
Y DESARROLLO TECNOLÓGICO**



**Study of strategies for genetic variant  
discrimination and detection by optosensing**

**Ph.D. THESIS**

Submitted by

**Ana Lázaro Zaragoza**

Ph.D. Supervisors:

Prof. Dr. Ángel Maquieira Catalá

Prof. Dr. Luis Antonio Tortajada Genaro

**Valencia, July 2022**





UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



ÁNGEL MAQUIEIRA CATALÁ and LUIS ANTONIO TORTAJADA GENARO, PhD. in Chemistry and professors at the *Universitat Politècnica de València*,

CERTIFY:

That the work “*Study of strategies for genetic variant discrimination and detection by optosensing*” has been developed by Ana Lázaro Zaragoza under their supervision in the Instituto de Reconocimiento Molecular y Desarrollo Tecnológico at the *Universitat Politècnica de València*, as a thesis Project in order to obtain the degree of PhD. in Chemistry at the *Universitat Politècnica de València*.

Valencia, May 9<sup>th</sup> 2022

Prof. Dr. Ángel Maquieira Catalá

Prof. Dr. Luis Antonio Tortajada Genaro



***“En medio del odio descubrí que había dentro de mí un amor invencible. En medio de las lágrimas me pareció que había dentro de mí una sonrisa invencible. En medio del caos descubrí que había en mi interior una calma invencible. Me di cuenta, a pesar de todo, que en medio del invierno había dentro de mí un verano invencible. Y eso me hace feliz. Porque no importa lo duro que el mundo empuje en mi contra; en mi interior hay algo más fuerte, algo mejor, empujando de vuelta”.***

Adaptación El Verano, Albert Camus



## AGRADECIMIENTOS

Aquello que en su día parecía tan lejano, hoy estoy a punto de conseguirlo. Tras estos cuatros años inolvidables, siento una mezcla de sentimientos, alegría, satisfacción, emoción, e incluso orgullo. La frustración, agobios y ansiedad que me han acompañado durante este largo proceso, hoy han desaparecido y solo queda la recompensa de sentir que el esfuerzo y el sacrificio han valido la pena. Pero soy muy consciente que no hubiera podido llegar hasta aquí sola. Por eso me gustaría agradecer a todas las personas que me han apoyado durante estos años y han hecho posible que de manera directa o indirecta pueda cumplir mi sueño de ser doctora.

En primer lugar, a mis tutores de tesis, por hacer que mi paso por el doctorado haya sido tan provechoso y constructivo. Gracias Ángel, por haberme dado la oportunidad de formar parte de este grupo de investigación y haberme dedicado tu tiempo y esfuerzo. Has sido para mí, un referente por tu pasión y perseverancia en el trabajo. Y a ti, Luis, gracias por guiarme y orientarme a lo largo de todos estos años. Gracias por haberme apretado cuando tocaba, pero también por la comprensión en los momentos más difíciles. Por tus valiosas sugerencias, tu paciencia infinita y, las innumerables horas que has invertido conmigo con el fin de formarme como una científica crítica y profesional. Espero que te sientas orgulloso del trabajo que hemos logrado juntos.

Así mismo deseo expresar mi gratitud a la Generalidad Valenciana por la beca predoctoral ACIF/2017 que me ha brindado todos los recursos y herramientas necesarios para realizar esta tesis. Además, estoy muy agradecida a la beca BEFPI/2020 que me ha permitido disfrutar de una maravillosa estancia en Eindhoven University of Technology. Gracias especialmente a Maarten Merkx por darme la oportunidad de formar parte de su grupo *Protein Engineering*, por su trato tan amable y cercano. Ha sido una experiencia magnífica en la que he aprendido mucho tanto en lo profesional como en lo personal, me ha servido para abrir la mente, y sin duda alguna ha supuesto un antes y después en mi vida.

A los profesores del departamento con los que he tenido la suerte de coincidir. Sergi, M<sup>a</sup> José, Javi, Miguel Ángel, Miguel Peris, David, Nuria, Pilar, Rosa, Eva, Patricia, Marichu, y Júlia. Gracias por vuestra amabilidad, sonrisas y simpatía. Por poner orden en el laboratorio y solucionar dudas y problemas siempre con cariño. Por todos los buenos consejos y por las palabras de ánimo. Os admiro y espero haber podido absorber todas las enseñanzas que me habéis brindado.

A todos los compañeros de laboratorio, que son responsables de que recuerde esta etapa con una sonrisa en mi cara por todos los buenos ratos que hemos compartido

juntos. Gracias por la complicidad construida a lo largo del tiempo. Me habéis hecho disfrutar sobre todo en las comidas en el césped al sol con el campamento tipo camping que montábamos. Nunca olvidaré lo bien que nos lo pasamos en el viaje a la boda de Dani, el zapato perdido de Salva en el contenedor de cristal, y como Miquel sabe bien que no hay nada mejor para los domingos por la mañana que una Coca-Cola “normal”. Lo divertido que eran los congresos a vuestro lado, lo ansiosos que esperábamos que llegara el día del jamón, las bromas constantes e interminables. Gracias por saber entender cómo me costaba arrancar la semana, por aguantar mis recitales musicales, que empeoraron cuando me compré el micrófono. También por tranquilizarme y acompañarme en los momentos no tan buenos. Son recuerdos que voy a llevar para siempre en mi corazón.

Algunos ya habéis terminado, a otros os queda poco para acabar y otros empezasteis cuando yo estaba en la recta final. Gracias a Miquel, Salva, Dani, Noelle, Eric, Zeneida, Estrella, Cintia, Gabi, Edurne, Pilar, William, Vicky, Julieth, Paola, Yeray, Amadeo, Aitor y Maribel. A Pedro, mi sevillano favorito, por compartir gustos musicales y darlo todo cuando escuchamos *reggaeton*. A Mjota que con su arte singular siempre estaba dispuesta a ayudarme con todas mis dudas sin cansarse. Mención especial a Augusto por conseguir levantarme cada vez que tropezaba (tanto en sentido literal como metafórico). Y, por último, a mi Saris por ser mi compañera de batallas. Ver tus ganas de crecer como persona e investigadora y como has luchado contra las adversidades consiguiendo lograr tus objetivos ha sido todo un ejemplo para mí. Te convertiste en mi apoyo, en las orejas que escuchaban atentamente mis inquietudes, en el brazo que me arropaba cuando lloraba y en la mano que me sostenía cuando todo me venía grande. Me siento extremadamente afortunada por haberme encontrado con un grupo de profesionales maravillosos y trabajar en un lugar donde me he podido sentir como en casa desde el primer día.

También quiero agradecer a mis amigos. Por poder contar con ellos para todo sin importar la distancia. En particular a Sandra, Carol, Mireya, Imma, Blanca y Silvia. Vuestra amistad es un tesoro para mí. Me ilumináis en mis días grises y me motiváis para que alcance todo aquello que deseo. Soy muy feliz de haber crecido con vosotras y mantener un vínculo tan bonito. Cuando estamos juntas parece que no pase el tiempo y solo nos basta una mirada para comprendernos. Por seguir entendiéndonos como lo hacemos, disfrutando tanto juntas y continuar siendo cómplices en muchas más aventuras. Sois imprescindibles en mi vida y os quiero millones.

Por supuesto también quiero dedicar unas palabras a mi familia. Sois el pilar más valioso de mi vida y mi mayor suerte. Gracias por enseñarme lo que significa el amor



incondicional. Os quiero con todo mi corazón a todos, y os voy a nombrar porque tengo la necesidad de demostrar lo importante que sois para mí.

A mi familia de Silla, mis abuelos Eduardo y Paqui, mi tía Laura, y tío Paco, y mis tetes Itziar e Isaac. He crecido con vosotros, me habéis enseñado el valor de la familia y me habéis hecho ser quien soy. Siempre estáis ahí en primera fila para celebrar conmigo las cosas buenas y apoyarme en las malas. Gracias por valorarme y confiar en que lo lograría.

A mi familia de Aranjuez, mis abuelos Eduardo y Visi, mis tías Mamen, Mar, María, Jaje, Kelly y Laura. A mis tíos José, Fede, Fermín y Oliver. A mis primas María, Ana, Sofia, y Olivia, y primos Fabian, Alex, Mateo y Máximo. No nos vemos todo lo que nos gustaría, pero cuando estamos todos juntos es una fiesta que guardo en mi corazón. Saber que siempre puedo contar con vosotros no tiene precio. Gracias por animarme a seguir adelante y esperar lo mejor de mí. En especial a mis primas María, Itziar y Ana. Nuestra relación es muy fuerte y única, traspasa la amistad. Sabéis que os considero mis hermanas y que os quiero muchísimo.

Ahora en inglés para mi familia holandesa, Albert, Janet, Rick, Thomas y Robby. Thank you for making me feel at home even when there are more than 1,950 Km away. Robby, my love, you are my soulmate and the best partner. I am so lucky to be able to go through life by your side. I love you to the stars and beyond. Forever and always.

Y como no, gracias a mis padres por cuidarme y quererme tanto. Por estar a mi lado de manera incondicional. Saber que pase lo que pase puedo contar con vosotros me da una paz que tranquiliza a todos mis demonios. Gracias por todo lo que habéis hecho por mí que me ha permitido cumplir hoy un gran sueño. Sin vosotros nada tendría sentido. Os debo todo lo que soy. Sois mi hogar, el lugar al que siempre quiero volver sea cual sea mi viaje. En especial gracias a mi madre, que es la persona más bondadosa e inteligente que conozco. Es increíble como eres capaz de sacrificarte por los tuyos y como te desvives por mí. Gracias por creer en mí, apoyarme y darme fuerzas cuando más lo necesitaba. Me has enseñado a no hundirme frente a las adversidades. Gracias por estar siempre al pie del cañón, preparada para descifrar mis miradas. Gracias por dejarme volar sola, aunque ese vuelo te quitara el sueño. Por ser mi guía, soporte y refugio; y todo al mismo tiempo. Sinceramente pienso que nuestra conexión es de otra dimensión. Esta tesis, resultado del esfuerzo, dedicación y sacrificio, te la dedico a ti, mami, porque también es tuya.

**¡Muchas gracias a todos por formar parte de mi vida!**



## ABSTRACT

Current medicine is moving towards a more personalized approach based on the patients' molecular diagnosis through the study of specific biomarkers. Diagnosis, prognosis and therapy selection, applying this molecular principle, rely on identifying specific variations in the human genome, such as single nucleotide variations (SNV). A wide range of technologies is available to detect these biomarkers. However, many of the employed methods have limitations such as high cost, complexity, long analysis times, or requiring specialized personnel and equipment, making their massive incorporation in most healthcare systems impossible. Therefore, there is a need to research and develop analytical solutions that provide information on genetic variants that can be implemented in different health scenarios with competitive and economically feasible performances.

The main objective of this thesis has been to develop innovative strategies to solve the challenge of multiple detection of genetic variants that are found in a minority amount in patient samples, covering the demands associated with the clinical setting. Research tasks focused on the combination of allelic discrimination reactions with selective DNA amplification and the development of versatile optical detection systems.

In order to meet the wide range of needs, in the first chapter, the analytical performances of the polymerase chain reaction (PCR) were improved by incorporating a thermocycling step and a blocking agent to amplify selectively minority variants that were monitored by real-time fluorescence. In the second chapter, allelic discrimination was achieved by combining oligonucleotide ligation with recombinase polymerase amplification (RPA), which operates at a constant temperature, allowing point-of-care (POC) detection. SNV identification was carried out by hybridization in microarray format, using Blu-Ray technology as the assay platform and detector. RPA was integrated with allele-specific hybridization chain reaction (AS-HCR), in an array format to genotype SNV from genomic DNA on a chip in the third chapter. The reading of the results was performed using a smartphone. In the last chapter, a new bioluminescent reagent was synthesized. It was applied to real-time and endpoint DNA biomarker monitoring based on bioluminescence resonance energy transfer (BRET), eliminating the need for an excitation source.

All the strategies allowed specific recognition of the target variant, even in samples containing as few as 20 copies of target genomic DNA. Sensitive (limit of detection 0.5% variant/total), reproducible (relative standard deviation < 19%), simple (3 steps or less), fast (short times of 30-200 min) results were achieved, allowing simultaneous analysis of several genes. As proof of concept, these strategies were

applied to detect and identify biomarkers associated with colorectal cancer and cardiological diseases in clinical samples. The results were validated by comparison with reference methods such as next-generation sequencing (NGS) and PCR, proving that the technical requirements and cost-effectiveness were improved. In conclusion, the developed research made it possible to develop genotyping tools with competitive analytical properties and versatile, applicable to different healthcare scenarios, from hospitals to limited-resource environments. These results are promising since they respond to the demand for alternative technologies for molecular diagnostics allowing progress in disease monitoring and personalized medical treatment.

## RESUMEN

La medicina actual se dirige hacia un enfoque más personalizado basándose en el diagnóstico molecular del paciente a través del estudio de biomarcadores específicos. Aplicando este principio molecular, el diagnóstico, pronóstico y selección de la terapia se apoyan en la identificación de variaciones específicas del genoma humano, como variaciones de un único nucleótido (SNV). Para detectar estos biomarcadores se dispone de una amplia oferta de tecnologías. Sin embargo, muchos de los métodos en uso presentan limitaciones como un elevado coste, complejidad, tiempos de análisis largos o requieren de personal y equipamiento especializado, lo que imposibilita su incorporación masiva en la mayoría de los sistemas sanitarios. Por tanto, existe la necesidad de investigar y desarrollar soluciones analíticas que aporten información sobre las variantes genéticas y que se puedan implementar en diferentes escenarios del ámbito de la salud con prestaciones competitivas y económicamente viables.

El objetivo principal de esta tesis ha sido desarrollar estrategias innovadoras para resolver el reto de la detección múltiple de variantes genéticas que se encuentran en forma minoritaria en muestras biológicas de pacientes, cubriendo las demandas asociadas al entorno clínico. Las tareas de investigación se centraron en la combinación de reacciones de discriminación alélica con amplificación selectiva de DNA y el desarrollo de sistemas ópticos de detección versátiles.

Con el fin de atender el amplio abanico de necesidades, en el primer capítulo, se presentan resultados que mejoran las prestaciones analíticas de la reacción en cadena de la polimerasa (PCR) mediante la incorporación de una etapa al termociclado y de un agente bloqueante amplificando selectivamente las variantes minoritarias que fueron monitorizadas mediante fluorescencia a tiempo real. En el segundo capítulo, se logró la discriminación alélica combinando la ligación de oligonucleótidos con la amplificación de la recombinasa polimerasa (RPA), que al operar a temperatura constante permitió una detección tipo *point-of-care* (POC). La identificación de SNV se llevó a cabo mediante hibridación en formato micromatriz, utilizando la tecnología Blu-Ray como plataforma de ensayo y detección. En el tercer capítulo, se integró la RPA con la reacción de hibridación alelo específica en cadena (AS-HCR), en formato array para genotipar SNV a partir de DNA genómico en un chip. La lectura de los resultados se realizó mediante un smartphone. En el último capítulo, se presenta la síntesis de un nuevo reactivo bioluminiscente que se aplicó a la monitorización de biomarcadores de DNA a tiempo real y final de la RPA basada en la transferencia de energía de resonancia de bioluminiscencia (BRET), eliminando la necesidad de una fuente de excitación.

Todas las estrategias permitieron un reconocimiento específico de la variante de interés, incluso en muestras que contenían tan solo 20 copias de DNA genómico diana. Se consiguieron resultados sensibles (límite de detección 0.5% variante/total), reproducibles (desviación estándar relativa < 19%), de manera sencilla (3 etapas o menos), rápida (tiempos cortos de 30-200 min) y permitiendo el análisis simultáneo de varios genes. Como prueba de concepto, estas estrategias se aplicaron a la detección e identificación en muestras clínicas de biomarcadores asociados a cáncer colorrectal y enfermedades cardiológicas. Los resultados se validaron por comparación con los métodos de referencia secuenciación de nueva generación (NGS) y PCR, comprobándose que se mejoraban los requerimientos técnicos y la relación coste-eficacia. En conclusión, las investigaciones llevadas a cabo posibilitaron desarrollar herramientas de genotipado con propiedades analíticas competitivas y versátiles, aplicables a diferentes escenarios sanitarios, desde hospitales a entornos con pocos recursos. Estos resultados son prometedores al dar respuesta a la demanda de tecnologías alternativas para el diagnóstico molecular y permitir avanzar en el seguimiento de enfermedades y tratamiento médico personalizado.

## RESUM

La medicina actual es dirigeix cap a un enfocament més personalitzat basant-se en el diagnòstic molecular del pacient a través de l'estudi de biomarcadors específics. Aplicant aquest principi molecular, el diagnòstic, pronòstic i selecció de la teràpia es recolzen en la identificació de variacions específiques del genoma humà, com variacions d'un únic nucleòtid (SNV). Per a detectar aquests biomarcadors, es disposa d'una àmplia oferta de tecnologies. No obstant això, molts dels mètodes en ús presenten limitacions com un elevat cost, complexitat, temps d'anàlisi llargues o requereixen de personal i equipament especialitzat, la qual cosa impossibilita la seua incorporació massiva en la majoria dels sistemes sanitaris. Per tant, existeix la necessitat d'investigar i desenvolupar solucions analítiques que aporten informació sobre les variants genètiques i que es puguin implementar en diferents escenaris de l'àmbit de la salut amb prestacions competitives i econòmicament viables.

L'objectiu principal d'aquesta tesi ha sigut desenvolupar estratègies innovadores per a resoldre el repte de la detecció múltiple de variants genètiques que es troben en forma minoritària en mostres biològiques de pacients, cobrint les demandes associades a l'entorn clínic. Les tasques d'investigació es van centrar en la combinació de reaccions de discriminació al·lèlica amb amplificació selectiva de DNA i al desenvolupament de sistemes òptics de detecció versàtils.

Amb la finalitat d'atendre l'ampli ventall de necessitats, en el primer capítol, es presenten resultats que milloren les prestacions analítiques de la reacció en cadena de la polimerasa (PCR) mitjançant la incorporació d'una etapa al termociclat i d'un agent bloquejant amplificant selectivament les variants minoritàries que van ser monitoritzades mitjançant fluorescència a temps real. En el segon capítol, es va aconseguir la discriminació al·lèlica combinant el lligament d'oligonucleòtids amb l'amplificació de la recombinasa polimerasa (RPA), que en operar a temperatura constant va permetre una detecció tipus *point-of-care* (POC). La identificació de SNV es va dur a terme mitjançant hibridació en format micromatriu, utilitzant la tecnologia Blu-Ray com a plataforma d'assaig i detecció. En el tercer capítol, es va integrar la RPA amb la reacció d'hibridació al·lèlica específica en cadena (AS-HCR), en format matriu per a genotipar SNV a partir de DNA genòmic en un xip. La lectura dels resultats es va realitzar mitjançant un telèfon intel·ligent. En l'últim capítol, es presenta la síntesi d'un nou reactiu bioluminescent que es va aplicar al monitoratge de biomarcadors de DNA a temps real i final de la RPA basada en la transferència d'energia de ressonància de bioluminescència (BRET), eliminant la necessitat d'una font d'excitació.

Totes les estratègies van permetre un reconeixement específic de la variant d'interès, fins i tot en mostres que només contenen 20 còpies de DNA genòmic diana. Es van aconseguir resultats sensibles (límit de detecció 0.5% variant/total), reproduïbles (desviació estàndard relativa < 19%), de manera senzilla (3 etapes o menys), ràpida (temps curts de 30-200 min) i permetent l'anàlisi simultània de diversos gens. Com a prova de concepte, aquestes estratègies es van aplicar a la detecció i identificació en mostres clíniques de biomarcadors associats a càncer colorectal i a malalties cardiològiques. Els resultats es van validar per comparació amb els mètodes de referència seqüenciació de nova generació (NGS) i PCR, comprovant-se que es milloraven els requeriments tècnics i la relació cost-eficàcia. En conclusió, les investigacions dutes a terme van possibilitar desenvolupar eines de genotipat amb propietats analítiques competitives i versàtils, aplicables a diferents escenaris sanitaris, des d'hospitals a entorns amb pocs recursos. Aquests resultats són prometedors en donar resposta a la demanda de tecnologies alternatives per al diagnòstic molecular i permetre avançar en el seguiment de malalties i tractament mèdic personalitzat.



## LIST OF PUBLICATIONS

- 1- Lázaro, A., Yamanaka, E. S., Maquieira, A., & Tortajada-Genaro, L. A. (2019). Allele-specific ligation and recombinase polymerase amplification for the detection of single nucleotide polymorphisms. *Sensors and Actuators B: Chemical*, 298, 126877. Impact factor: 7.34; Q1 (SJR 2020:1.6); CiteScore: 14.0
- 2- Lázaro, A., Tortajada-Genaro, L. A., & Maquieira, Á. (2021). Enhanced asymmetric blocked qPCR method for affordable detection of point mutations in *KRAS* oncogene. *Analytical and Bioanalytical Chemistry*, 413(11), 2961-2969. Impact factor: 3.74; Q1 (SJR 2020:0.86); CiteScore: 6.2
- 3- Lázaro, A., Maquieira, A., & Tortajada-Genaro, L. A. (2022). Discrimination of Single-Nucleotide Variants Based on an Allele-Specific Hybridization Chain Reaction and Smartphone Detection. *ACS sensors*, 7(3), 758-765. Impact factor: 7.71; Q1 (SJR 2020: 2.06); CiteScore: 10.3
- 4- Lázaro, A., Maquieira, Á., Merk M., & Tortajada-Genaro, L. A. Development of a BRET-based intercalating dye for DNA biomarker detection (in writing phase).

### **Communications in conferences:**

1. Lázaro A., Tortajada-Genaro L. A., Maquieira A. Reconocimiento selectivo de alelos minoritarios mediante la técnica COLD-PCR combinada con ensayos de hibridación in XI International Workshop on Sensors and Molecular Recognition (5<sup>th</sup> and 6<sup>th</sup> of July, 2017), at the Universitat Politècnica de València. ISBN: 978-84-697-5069-8 (pp 138-141)
2. Jiménez I., Lázaro A., Martorell S., Tortajada-Genaro L. A., Maquieira A. Mutational analysis of *BRAF* oncogene in colorectal cancer patients in XI International Workshop on Sensors and Molecular Recognition (5<sup>th</sup> and 6<sup>th</sup> of July, 2017), at the Universitat Politècnica de València. ISBN: 978-84-697-5069-8 (pp 121-125). Best Poster Award.
3. Tortajada-Genaro L. A., Martorell S., Lázaro A., Palanca S., Maquieira A. Detección de mutaciones en oncogenes en el tratamiento dirigido al cáncer colorrectal. 2017. *Actualidad Analítica*. 60, ISSN 2340-8006 (pp. 21-24). Best Poster Award.
4. Tortajada-Genaro L. A., Yamanaka, E. S., Martorell S., Lázaro A., Puchades R., Maquieira A. Estabilidad de ácidos nucleicos: análisis termodinámico para el bio-reconocimiento de mutaciones y polimorfismos en la XXXVI Reunión Bial de la Real Sociedad Española de Química (RSEQ 2017) (29<sup>th</sup> of June, 2017) en Sitges.

5. Lázaro A., Tortajada-Genaro L. A., Maquieira A. Advanced blocked DNA amplification of mutant variants related to cancer treatment in XII International Workshop on Sensors and Molecular Recognition (5<sup>th</sup> and 6<sup>th</sup> of July, 2018), at the Universitat de València. ISBN: 978-84-09-05881-5 (pp 261-264).
6. Lázaro, A., Yamanaka, E. S., Maquieira, A., & Tortajada-Genaro, L. A. Blu-ray technology for classifying patients with cardiovascular diseases in XII International Workshop on Sensors and Molecular Recognition (5<sup>th</sup> and 6<sup>th</sup> of July, 2018), at the Universitat de València. ISBN: 978-84-09-05881-5 (pp 257-260).
7. Lázaro A., Tortajada-Genaro L. A., Maquieira A. Análisis de mutaciones minoritarias mediante biosensado óptico in Encuentro de Doctorandos (27<sup>th</sup> of June, 2019), at the Universidad Politécnica de Valencia. Poster selected among the 8 best.
8. Lázaro A., Tortajada-Genaro L. A., Maquieira A. Enrichment of clinically significant low-level mutations in *KRAS* gene based on blocked PCR in XIII International Workshop on Sensors and Molecular Recognition (4<sup>th</sup> and 5<sup>th</sup> of July, 2019), at the Universitat Politècnica de València. ISBN 978-84-09-15385-5 (pp 131-134).
9. Villamayor M., Lázaro A., Tortajada-Genaro L. A., Maquieira A. Sistema de detección de gluten en alimentos mediante biosensado genético in XIII International Workshop on Sensors and Molecular Recognition (4<sup>th</sup> and 5<sup>th</sup> of July, 2019), at the Universitat Politècnica de València. ISBN 978-84-09-15385-5 (pp 23-26).
10. Tortajada-Genaro L. A., Lázaro A., Morais S., Maquieira A. Fast DNA biosensing based on gold nanoparticles and consumer electronic devices in online Rapid Methods Europe 2021 (19<sup>th</sup> of February, 2021).
11. Tortajada-Genaro L. A., Lázaro A., Maquieira A. Hybridization chain reaction for the optical detection of oncogene variants inl XIV International Workshop on Sensors and Molecular Recognition (5<sup>th</sup> of July, 2021) at the Universitat de València.

## LIST OF ABBREVIATIONS

ARMS	Amplification refractory mutation system
ASA	Allele-specific amplification
ASO	Allele-specific oligonucleotide
ATP	Adenosine triphosphate
BD	Blu-ray disc
bp	Base pair
BRET	Bioluminescence resonance energy transfer
CD	Compact disc
Cq	Quantification cycle
ctDNA	Circulating tumor DNA
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside triphosphates
dUTP	Deoxyuridine triphosphate
dsDNA	Double stranded DNA
DVD	Digital versatile discs
FDA	Food and Drug Administration
FFPE	Formalin-fixed, paraffin-embedded tissue
FRET	Fluorescence resonance energy transfer
cDNA	Complementary DNA strand
gDNA	Genomic DNA
HCR	Hybridization chain reaction
HDA	Helicase-dependent amplification
LAMP	Loop-mediated isothermal amplification
LCR	Ligation chain reaction
LNA	Locked nucleic acid
LOD	Limit of detection
MALDI-TOF	Matrix-assisted laser induced-time of flight
Mb	Megabase
MDA	Multiple displacement amplification
NASBA	Nucleic acid sequence-based amplification
NGS	Next-generation sequencing

NTC	Non-template control
PC	Polycarbonate
PCR	Polymerase chain reaction
PNA	Peptide nucleic acids
POC	Point of care
qPCR	Quantitative polymerase chain reaction
RCA	Rolling circle amplification
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RPA	Recombinase polimerase amplification
SDA	Strand displacement amplification
ssDNA	Single stranded DNA
SNV	Single nucleotide variants
T <sub>m</sub>	Melting temperature
WT	Wild-type

# CONTENTS

<b>1. INTRODUCTION .....</b>	<b>1</b>
<b>1.1 BIOMARKERS.....</b>	<b>1</b>
1.1.1 Concept .....	1
1.1.2 Types of biomarkers .....	2
1.1.3 Fundamentals of DNA biomarkers .....	4
1.1.3.1 Types of mutation.....	5
1.1.3.2 Single nucleotide variants.....	5
1.1.4 Relationship with human health .....	7
1.1.4.1 Colorectal cancer .....	8
1.1.4.2 Pharmacogenetics of cardiovascular diseases .....	9
1.1.5. Detection of biomarkers .....	11
<b>1.2 CONVENTIONAL TECHNOLOGIES .....</b>	<b>13</b>
1.2.1 Sequencing.....	13
1.2.2 Microarrays.....	15
1.2.3 Enzyme-based methods .....	19
1.2.3.1 Polymerases-based methods .....	19
1.2.3.2 Ligation-based methods .....	24
1.2.3.3 Endonucleases.....	26
<b>1.3 NEW BIORECOGNITION STRATEGIES.....</b>	<b>28</b>
1.3.1 Improvement of PCR methods .....	28
1.3.2 Alternative methods to PCR.....	34
1.3.2.1 CRISPR/Cas methods.....	34
1.3.2.2 Enzymatic isothermal amplification methods .....	37
1.3.2.3 Non-enzymatic isothermal amplification methods.....	48
<b>1.4 NOVEL DETECTION STRATEGIES.....</b>	<b>55</b>
1.4.1 Transduction modes .....	55
1.4.2 Consumer electronic devices .....	58
1.4.2.1 Optical digital disc technology .....	59
1.4.2.2 Smartphone technology.....	62
<b>1.5. METHODS COMPARISON.....</b>	<b>65</b>
1.5.1 Recapitulation .....	65
1.5.2 Unsolved challenges.....	67
<b>1.6 REFERENCES.....</b>	<b>68</b>
<b>2. OBJECTIVES .....</b>	<b>81</b>
<b>3. EXPERIMENTAL SECTION AND RESULTS .....</b>	<b>85</b>

<b>3.1 CHAPTER 1: ENHANCED ASYMMETRIC BLOCKED qPCR METHOD FOR AFFORDABLE DETECTION OF POINT MUTATIONS IN <i>KRAS</i> ONCOGENE .....</b>	<b>87</b>
3.1.1 Abstract .....	88
3.1.2 Introduction .....	88
3.1.3 Materials and Methods.....	90
3.1.4 Results.....	91
3.1.5 Discussion .....	97
3.1.6 Conclusions .....	99
3.1.7 References .....	100
3.1.8 Supplementary information .....	102
<b>3.2 CHAPTER 2: ALLELE-SPECIFIC LIGATION AND RECOMBINASE POLYMERASE AMPLIFICATION FOR THE DETECTION OF SINGLE NUCLEOTIDE POLYMORPHISMS .....</b>	<b>111</b>
3.2.1. Abstract .....	112
3.2.2 Introduction .....	112
3.2.3 Materials and methods.....	114
3.2.4 Results and discussion .....	117
3.2.5 Conclusions .....	123
3.2.6 References .....	123
3.2.7 Supplementary information .....	125
<b>3.3 CHAPTER 3: DISCRIMINATION OF SINGLE NUCLEOTIDE VARIANTS BASED ON ALLELE-SPECIFIC HYBRIDIZATION CHAIN REACTION AND SMARTPHONE DETECTION .....</b>	<b>139</b>
3.3.1 Abstract .....	140
3.3.2 Introduction .....	140
3.3.3 Materials and methods.....	142
3.3.4 Results and discussion .....	143
3.3.5 Conclusions .....	153
3.3.6 References .....	154
3.3.7 Supplementary information .....	156
<b>3.4 CHAPTER 4: DEVELOPMENT OF A BRET-BASED INTERCALATING DYE FOR DNA BIOMARKER DETECTION.....</b>	<b>179</b>
3.4.1 Abstract .....	180
3.4.2 Introduction .....	181
3.4.3 Materials and methods.....	183
3.4.4 Results and discussion .....	185
3.4.5 Conclusions .....	196
3.4.6 References .....	197
3.4.7 Supplementary Information .....	200

<b>4. RESULTS DISCUSSION .....</b>	<b>211</b>
<b>5. GENERAL CONCLUSIONS .....</b>	<b>219</b>

## LIST OF FIGURES

### **Figures of introduction:**

Figure 1. Outline of the types of mutation clinically relevant. ....	6
Figure 2. Scheme of the relationship of biomarkers to personalized medicine and human health.....	7
Figure 3. Effect of anti-EGFR monoclonal therapy on K-ras signaling.....	9
Figure 4. Mechanism of warfarin anticoagulant action.....	10
Figure 5. Microarray technology workflow. ....	15
Figure 6. Formats of PCR monitoring.....	23
Figure 7. MLPA reaction phases.....	26
Figure 8. Application of CRISPR/Cas systems for sensitive SNV detection.. ....	36
Figure 9. Mechanism of NASBA amplification. ....	38
Figure 10. Mechanism of SDA amplification.....	39
Figure 11. Mechanism of RCA amplification. ....	40
Figure 12. Mechanism of MDA amplification. ....	41
Figure 13. Mechanism of LAMP amplification.. ....	42
Figure 14. Mechanism of HDA amplification. ....	43
Figure 15. Mechanism of RPA amplification.....	44
Figure 16. RPA metric studies. ....	45
Figure 17. Different detection approaches using RPA.....	46
Figure 18. Mechanism of HCR amplification. ....	48
Figure 19. HCR metric studies. ....	49
Figure 20. Different detection approaches using HCR. ....	50
Figure 21. Different transduction modes of the biorecognition process. ....	56
Figure 22. Comparison between FRET and BRET approaches. ....	57
Figure 23. Outline of the most common compact discs. ....	61
Figure 24. Examples of applications that use smartphones as readers of the biorecognition process. ....	63

### **Figures of chapter 1:**

Summary figure. Overview chapter 1 .....	87
Figure 1. Scheme of the mechanism of the EAB-qPCR method. ....	92
Figure 2. Discrimination effect depending on the EAB-qPCR conditions.....	93
Figure 3. Amplification curves with different formats. ....	94
Figure 4. Assay sensitivity of EAB-qPCR.....	95
Figure 5. Mutational analysis of cancer patients.....	97
Figure SI.1. Melting curve plotted.....	103
Figure SI.2. Real time curves depending on the thermal cycling and the presence of blocker.....	105



Figure SI.3. Conventional qPCR method. ....	106
Figure SI.4. (Top) Thermal cycling for conventional qPCR (left) and EAB-qPCR (right). (Bottom) Estimation of detection limits from patient samples (biopsied tissues). ....	108

**Figures of chapter 2:**

Summary figure. Overview chapter 2 .....	111
Figure 1. Scheme of the multiplexed discrimination assay for single-base changes..	115
Figure 2. Integration of ligation and universal RPA. ....	118
Figure 3. Optical detection of RPA products (array format).. ....	120
Figure 4. Discrimination map plotted from the response signals of patients (wild-type and mutant reactions) for each polymorphism. ....	122
Figure SI.1. Structure of ligation probes. ....	125
Figure SI.2. Sensing of hybridization assay in array format based on Blu-ray technology .....	126
Figure SI.3. Set-up of array detection based on signal-to-noise ratio (SNR). ....	128
Figure SI.4. Array layout. ....	135
Figure SI.5. Examples of array images from patient samples.....	136
Figure SI.6. Electropherogram obtained in Sanger sequencing for the VKORC1 gene, indicating the SNP position rs9923231. ....	137

**Figures of chapter 3:**

Summary figure. Overview chapter 3 .....	139
Figure 1. Scheme of genotyping by AS-HCR with colorimetric detection.....	144
Figure 2. HCR method setup.. ....	145
Figure 3. Smartphone detection of RPA-HCR products for the <i>KRAS</i> and <i>NRAS</i> targets. .....	147
Figure 4. Multiplexing capability.. ....	149
Figure 5. Comparison of the assay sensitivity of AS-HCR and PCR-based method. .	150
Figure 6. Identification of specific mutations from cancer patients using the AS-HCR method.....	152
Figure SI.1. Structure of the H1 and H2 oligonucleotides.....	156
Figure SI.2. Design of the oligonucleotides for the AS-HCR method:.....	157
Figure SI.3. Evaluation of conventional HCR reaction conditions.....	161
Figure SI.4. Linear vs. HCR approaches.....	163
Figure SI.5. Assembly for the smartphone-based detection of HCR products. ....	164
Figure SI.6. Optimization of chip immunostaining for smartphone-based detection in the colorimetric mode. ....	165
Figure SI.7. Assessment of the array quality obtained. ....	166
Figure SI.8. HCR method from RPA products. ....	167
Figure SI.9. Images obtained from patient samples.. ....	170

**Figures of chapter 4:**

Summary figure. Overview chapter 4.....	179
Figure 1. Scheme of DNA biomarker detection using bioluminescence sensing. ....	186
Figure 2. Control of NanoLuc synthesis. ....	188
Figure 3. NanoLuc-thiazole orange conjugates using lysine residues. ....	190
Figure 4. NanoLuc-thiazole orange conjugates using cysteine residue. ....	191
Figure 5. Comparison of NanoLuc-thiazole orange conjugates obtained following different reaction pathways. ....	193
Figure 6. Effect of DNA concentration in the luminescence signal using the conjugate complex. ....	194
Figure 7. Different transducers modes of biomarker DNA amplification.....	196
Figure SI.1. Synthetized NanoLuc luciferases.....	201
Figure SI.2. Thiazole orange succinyl ester as an intercalating agent.. ....	203
Figure SI.3. Reactivity of the reagents. ....	204
Figure SI.4. Control with the unconjugated NanoLuc and thiazol orange deactivated. ....	206
Figure SI.5. Luminescence results using variable concentration of thiazol orange (TO) in the solution. ....	206
Figure SI.6. Functionalization of thiazole orange (TO) with crosslinker (CL). Chromatogram (left) and mass spectrum (right).....	208
Figure SI.7. Chromatograms of reaction products from the functionalization of thiazole orange (TO) with crosslinker (CL). ....	209
Figure SI.8. BRET emission calculated from the difference between ssDNA (1 $\mu$ M) and dsDNA signal (1 $\mu$ M).....	210

## LIST OF TABLES

### **Tables of introduction:**

Table 1. The ten types of biomarkers by FDA-NIH Biomarker Working Group (2021) ...	3
Table 2. Types of mutations and their effect .....	6
Table 3. Comparison of the leading next-generation sequencing technologies. ....	14
Table 4. Comparison between probe immobilization strategies.....	16
Table 5. Factors affecting $T_m$ and hybridization.....	19
Table 6. SNV detection by HRM. ....	21
Table 7. Different strategies of ligation-based methods.....	25
Table 8. Characteristics of commonly used oligonucleotides. ....	29
Table 9. Characteristics of the main SNV genotyping techniques from modifications or combination of the described strategies. ....	31
Table 10. Comparison of COLD-PCR variants.....	33
Table 11. Comparison between characteristics of enzymatic isothermal amplification techniques. ....	47
Table 12. Examples of different HCR sensing approaches classified according to the transduction principle. ....	51
Table 13. Techniques currently available for SNV genotyping. ....	66

### **Tables of chapter 1:**

Table SI.1. Tested blockers for assay optimization of EAB-qPCR method.....	103
Table SI.2. Sequences of oligonucleotides used for the EAB-qPCR method. ....	104
Table SI.3. Threshold cycle ( $C_t$ ), percentage of blocking (% B) and discrimination factor ( $\Delta C_t$ ) for conventional blocked qPCR and EAB-qPCR.....	105
Table SI.4. <i>KRAS</i> mutation analysis of 20 clinical samples by studied methods .....	109

### **Tables of chapter 2:**

Table 1. Spot mean SNRs depending on the ligation mixture (single, duplex or triplex) and the forward primer used in the amplification reaction (wild-type or mutant). ....	121
Table SI.1. Variants associated to cardiovascular diseases.....	130
Table SI.2. List of tested oligonucleotides.....	131
Table SI.3. Cross-hybridization calculations of ligation probes, expressed as variation of free energy for template-probe complexes ( $\Delta G$ , kcal/mol).....	132
Table SI.4. List of oligonucleotides for RPA controls.....	134
Table SI.5. Results for example assays .....	138
Table SI.6. Assignment to genetic populations.....	138

### **Tables of chapter 3:**

Table SI.1. Thermodynamic analysis of the AS-HCR oligonucleotides.....	158
Table SI.2. Sequences of the oligonucleotides used for the AS-HCR method.....	159

Table SI.3. Sequences of the oligonucleotides used as biosensing process controls. ....	162
Table SI.4. Comparison of the PCR-chip and RPA-HCR methods.....	168
Table SI.5. The mutation analysis of the clinical samples by PCR-based method and NGS technique for the target SNVs. ....	171
Table SI.6. Examples of different HCR sensing approaches classified according to the sensing principle.....	172
Table SI.7. Currently available platforms and closely related SNV strategies for <i>RAS</i> genotyping.....	173
Table SI.8. Potential applications of the AS-HCR and examples showing the clinical relevance of correct genotyping. ....	175
<b>Tables of chapter 4:</b>	
Table SI.1. Overview intercalating agents characteristics .....	202
<b>Tables of results discussion:</b>	
Summary table. Comparison between the strategies developed in the thesis. ....	218

# **1. INTRODUCTION**



## **1.1 BIOMARKERS**

### **1.1.1 Concept**

In recent decades, technological and economic progress, public health measures, as well as the achievements of modern medicine have made it possible for the world's average life expectancy to increase significantly from 52.5 years in 1960 (the first global statistic) to an average of 72.4 today (1).

**Molecular medicine** must be highlighted among the areas that have led to this benefit in human health. It is based on the description at the molecular level of the current state of the cell. It allows identifying the presence of a disease, determining the probability of its development, and monitoring the course of pathology or the effects of treatment.

Molecular medicine is crucial both in developed countries, where chronic illnesses such as cancer, cardiovascular diseases, diabetes, or dementia are the leading cause of death per year; and in developing countries, where infectious diseases are one of the main threats to public health. This, coupled with other global health problems, such as the current hazard of pandemics or the emergence of antibiotic-resistant bacteria due to globalization, has increased its relevance (2).

Molecular medicine is based on the use of molecular diagnostics. The term molecular diagnostics encompasses a range of techniques used to identify and analyze biological markers or **biomarkers**. The most widely accepted definition of a biomarker is the one proposed by Biomarkers Definitions Working Group 2001 as "An objectively measurable characteristic indicative of a physiological condition, pathological process, or pharmacological response." In other words, its presence, absence, or concentration level will make it possible to detect a disease even before symptoms appear and also the effectiveness of treatment from the early stages of its administration. Therefore, a biomarker provides information that broadens the knowledge of pathophysiological mechanisms in a simple, non-invasive way that does not jeopardize the patient (3).

The ideal biomarker should reduce the uncertainty provided by clinical data, and thus, should meet the following chemical-analytical requirements:

- Be sensitive and specific.
- Be univocally associated with a specific stage, process, or disease.
- Be present in standard and easy to obtain samples (blood, urine, or saliva).
- Be detected quickly and inexpensively.
- Be ethically acceptable.
- Be used for preventive actions.

Currently, scientific advances are helping to identify new biomarkers. Based on the study of the parameters altered by a pathology, they are identified, then those that provide the most information is selected, and finally, they must be validated. For a biomarker to be valid, it must be measurable by a well-established procedure with high sensitivity and specificity, and its clinical significance must be well determined (4).

### **1.1.2 Types of biomarkers**

Based on the classification of biomarkers into ten types updated in 2021 according to the Food and Drug Administration (FDA) and the National Institutes of Health (NIH) Biomarker Working Group as shows Table 1 (5). In this thesis, they have been grouped into three categories according to the type of information and application they provide:

- **Disease biomarkers.** They provide information about the pathological mechanism of a disease, allowing early detection and tracking its progression over time.
- **Response therapeutic biomarkers.** They characterize the processes between the administration and the effect of a drug.
- **Surrogate endpoints biomarkers.** They offer valuation criteria and substitute measures of clinical outcome.

Some databases collect the different biomarkers, such as the GOBIOM (6), which currently contains more than 130,000, but biomarkers can be divided into three main types according to their nature:

- **Nucleic acids** such as DNA, are the genetic information carrier inside the cells. Moreover, RNA orchestrates and regulates gene expression.
- **Proteins** are responsible for carrying out various cellular functions (structural, transport, regulatory, or defense) translated in the ribosomes.
- **Metabolites** are molecules whose production provides valuable information about variations determined by genetic information, regulation and protein expression, and environmental influence.

The main staple of biomarkers is constituted by point a: DNA, the blueprint of life and the central focus of this thesis.



**Table 1.** The ten types of biomarkers by FDA-NIH Biomarker Working Group (2021) (5) (own adaptation).

Name	Definition	Example
<b>Disease biomarkers</b>		
Diagnostic	Correlates with the disease or condition of interest, allowing early identification of patients and identifying individuals with a disease subset	HbA1c to discriminate individuals with type 2 diabetes mellitus
Disease	Measures the progression or follow-up of diseases and the effect of exposure or toxicity to medical products or environmental agents recurrently	RNA Hepatitis C virus to evaluate therapy's impact in chronic hepatitis C patients
Prognostic	Offers knowledge on the natural course of the disease. They are associated with the individual probability of survival or repair of the disease	C-reactive protein to detect acute myocardial infarction patients with a more significant probability of recurrent coronary artery disease
Susceptibility/Risk	Shows the predisposition to have a disease or medical condition in a person who apparently does not develop the disease or its symptoms at this moment	Factor V Leiden distinguishes individuals with a high probability of suffering from deep vein thrombosis
<b>Response therapeutic biomarker</b>		
Pharmacodynamic	Indicates the effect after exposure to medicinal products or environmental factors	Serum LDL cholesterol to evaluate the repercussion of dietary changes in patients with hypercholesterolemia
Safety	Measures the degree of toxicity previous or posterior exposures to medical interventions or environmental agents	Hepatic aminotransferases and bilirubin to analyze the risk of hepatotoxicity
Predictive	Provides information about the patient's response to the treatment received. Their analysis allows the classification of patients as susceptible or not to receive specific treatments	<i>BRCA1</i> and <i>BRCA2</i> mutations to identify women with platinum-sensitive ovarian cancer who would benefit from Poly (ADP-ribose) polymerase inhibitors
<b>Surrogate endpoints biomarkers</b>		
Validated Surrogate Endpoint	Is measured in an analytical technique and a scientific framework that can meaningfully elucidate the physiology, pharmacology, toxicology, or clinical significance of the results	HIV-RNA for monitoring the clinical progress of patients with immunodeficiency virus
Reasonably Likely Surrogate Endpoint	Replaces a clinical evaluation and expects to predict clinical benefits based on epidemiology, therapeutics, pathophysiology, or other scientific evidence, but it requires more clinical data to be validated	Serotype-specific antibodies against <i>Streptococcus pneumoniae</i> for forecasting the risk of developing pneumococcal pneumonia
Candidate Surrogate Endpoint	Based on the published evidence, currently, there is insufficient data to demonstrate whether it will produce a beneficial clinical effect, as it is under investigation	The durable response rate for determining overall survival in immune-Checkpoint Inhibitors (ICI)-treated patients

### **1.1.3 Fundamentals of DNA biomarkers**

The study of DNA biomarkers is crucial since it contains all the genetic information, which is the key to vital phenomena in living organisms. For this reason, these genetic biomarkers have become an ideal and versatile target for molecular diagnostics in the genomic era (7).

The concept of genomics originated in 1986 when Thomas Roderick used it to refer to the molecular study of the: content, organization, function, and evolution of the genetic information contained in a complete genome. Since the mid-1970s, there had a great deal of interest in unveiling the secrets of life contained in the DNA sequence of *Homo sapiens*. The entire human genome was sequenced in the most ambitious international biology study known as the Human Genome Project (HGP), which revealed that the nuclear genome contains 3,200 Mb of DNA and 21,000 genes. Furthermore, that the differences between humans on the genetic level are only 0.1%. Isolation of genes whose mutations are responsible for diseases, development of genetic diagnostic tests and new treatments, diagnosis of biological paternity and biological kinship, identification of biological remains and post-mortem individuals were the most immediate benefits of the HGP (8,9).

On one hand, new studies were carried out, such as the **ENCyclopedia Of DNA Elements Project** (ENCODE). It collected all primary and mature transcripts and elucidated that a high percentage of the transcripts detected do not encode proteins. On the other hand, new related research areas, such as metagenomics, transcriptomics, proteomics, pharmacogenomics, or nutrigenomics, started to develop. They have involved relevant improvement of our understanding of human biology, allowing us to re-define the following key concepts:

- **Gen** as the storage unit of genetic information that is passed on to subsequent generations. It consists of a linear chain of nucleotides that includes the regions containing the information and the sequences responsible for its correct expression. Related terms are locus (loci, plural), which means the location on a chromosome where a gene is physically situated; and allele, one of the distinct forms of a gene that occur at a particular locus. At the same time, the set of alleles from different loci inherited together is known as a haplotype (10).

- **Genome** as the total genetic material endowment of an organism. A genome is much more than a set of genes because it also contains repeated sequences, intergenic regions, and regulatory sequences. The use of this information depends on complex biochemical machinery called gene expression. The concept of genotype refers to the

genetic constitution of an organism, whereas phenotype is the observable manifestation of the genotype (11).

- **Mutation** as any change in the phenotype that is due to a modification of the hereditary material of an individual. The mutations are the primary source of genetic variability and the basis of evolution. The result of mutation is the appearance of new alleles. This leads to the distinction of wild-type alleles, fully functional form of a gene, and mutant alleles, a form of a gene that has been derived from the wild-type form by mutation. In this context, it is possible to distinguish homozygote individuals, presenting two identical alleles of a gene, and heterozygous, who have two different alleles. In heterozygosity, the relationships between alleles of a gene can be studied, being dominant or recessive, depending on whether or not the allele is expressed in the phenotype of an heterozygous individual, respectively (12).

#### *1.1.3.1 Types of mutation*

Several types of mutations can be classified according to different criteria. Hereunder, the most interesting from the clinical point of view are discussed (Figure 1).

Concerning the cell type in which the mutations appear, **germline** mutations occur in the cells that produce gametes, so these mutations will not be expressed in the individual that contains them. Still, they will be transmitted to the next generation, which will present them from birth and, as they will be stable throughout life, they can be measured at any time. However, **somatic** mutations happen in somatic cells, i.e., in all except for germ cells. In this case, the mutations affect the individual carrier and can occur throughout the individual's life but are not transmitted to the offspring.

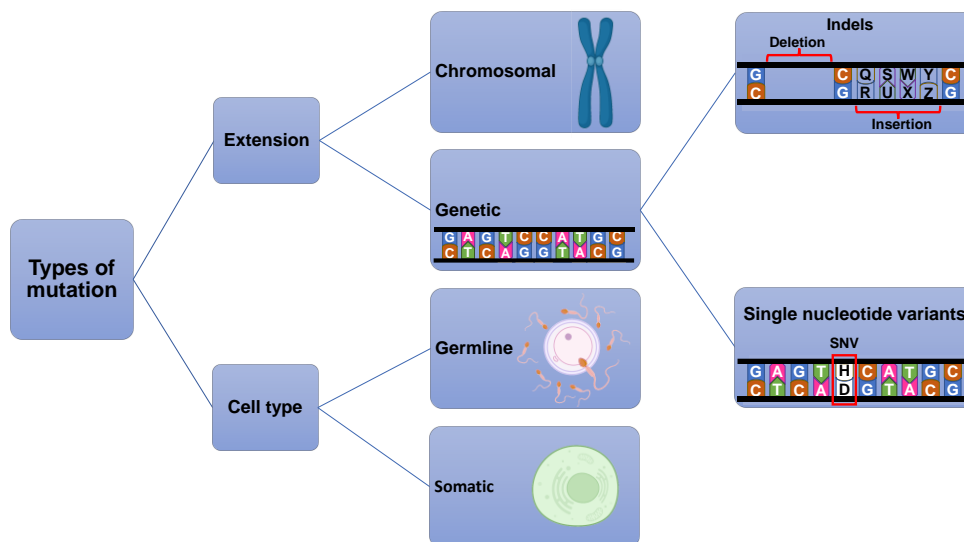
According to the extent of the affected genetic material, it can be classified as **chromosomal** or **genetic**. Chromosomal mutations are changes in the hereditary material resulting from the rearrangement of parts of the chromosome (structural) or the presence of an abnormal number of chromosomes (numerical). These will mainly result in gene dosage defects. In contrast, genetic mutations are variations in the DNA sequence that occur in individual genes. There are regions of a gene in which mutations appear more frequently called hot spots (10).

#### *1.1.3.2 Single nucleotide variants*

Among the genetic mutations are distinguished:

- **Insertion, deletions or indels.** When one or more nucleotides are incorporated or missed. One of the most known types is copy-number variations, implying large DNA fragments (more than one kilobases).

▪ **Single nucleotide variants (SNV) or point mutation.** When one nucleotide is replaced by another at a determined position. SNV are known as a transition if they change a nucleotide from a purine to another purine or from a pyrimidine to another pyrimidine. The latter is the most common type. Alternatively, as transversion, when a purine is replaced by a pyrimidine and vice versa (11)



**Figure 1.** Outline of the types of mutation clinically relevant.

The genetic mutations can alter the mRNA sequence, which in turn may affect the conformation, structure, function, activity, or expression level of the protein. As a result of these mutations, different situations can appear, as the Table 2 shows.

**Table 2.** Types of mutations and their effect.

Mutation	Aminoacid change	Change in the protein
Silent	No	No
Missense	Yes	Yes
Neutral	Yes, other with similar properties	No
Non-sense	Yes, introduce premature stop codon	Yes, shorter
Frameshift	Yes, alter the reading frame changing the polypeptide sequence or deleting a stop codon	Yes, different or longer protein

Mutations are responsible for the existence of **polymorphisms**, that indicate the diversity of species. The term polymorphism refers to the existence of two or more variants for a given locus. The least common variant must be found in at least 1% of the study population. The HapMap Project collected a catalog of polymorphisms present in the human genome. The study described that those polymorphisms could range from a single nucleotide to large regions of the chromosome. The most common variants are

**Single Base Polymorphisms (SNP).** In fact, SNPs make up 90% of the human genome since they occur with a frequency of 1 per 100-300 base pairs (13).

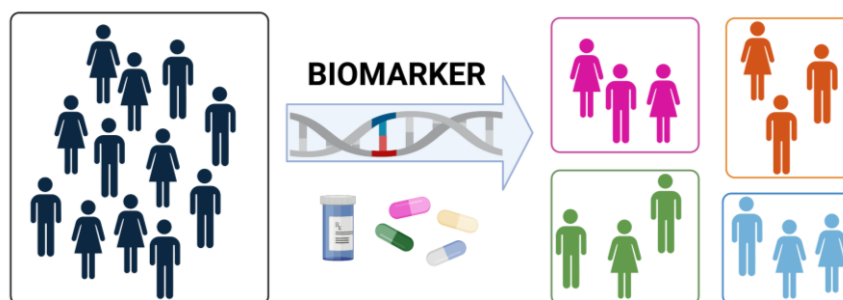
Thus, SNP are considered a form of point mutation with sufficient evolutionary success to become fixed in a significant part of the gene pool. Another finding is that some SNP can significantly affect the phenotype, depending on where they occur and the subsequent consequence on protein function and expression. Indeed, **Genome-Wide Association Studies (GWAS)** have mediated the fast discovery of genetic variants influencing illness susceptibility and therapy response (14).

Thanks to this and other projects, there are now publicly accessible databases containing descriptions of more than 15 million human genetic variants that can serve as biomarkers due to their correlation with diseases and provide essential information for medical care (15).

#### 1.1.4 Relationship with human health

For many years, medical care placed patients according to their medical signs and physiological status on the same level, stipulating standardized treatments for everyone. The development of biomarkers has led to a new approach, namely, personalized medicine. The pursuit of which has been a goal already raised by Hippocrates, the father of medicine, when he said, "it is much more important to know what kind of person has the disease than what kind of disease the person has." Personalized medicine aims to tailor treatments to the individual molecular characteristics of each patient and has enormous potential to improve medical care, making it safer and more efficient (7).

Biomarkers provide the tools required by personalized medicine to make an early diagnosis of a disease, recognize the efficacy of treatment from the beginning of its administration, and facilitate drug discovery. In addition, this approach has the capacity and potential to be used in numerous types of diseases, classifying the patient in the appropriate population group and the choice of a precision treatment (Figure 2).



**Figure 2.** Scheme of the relationship of biomarkers to personalized medicine and human health.

In particular, this thesis will discuss in detail the application of SNV biomarkers, and the technologies used for their detection in the context of colon cancer and in pharmacogenetics targeting cardiovascular diseases.

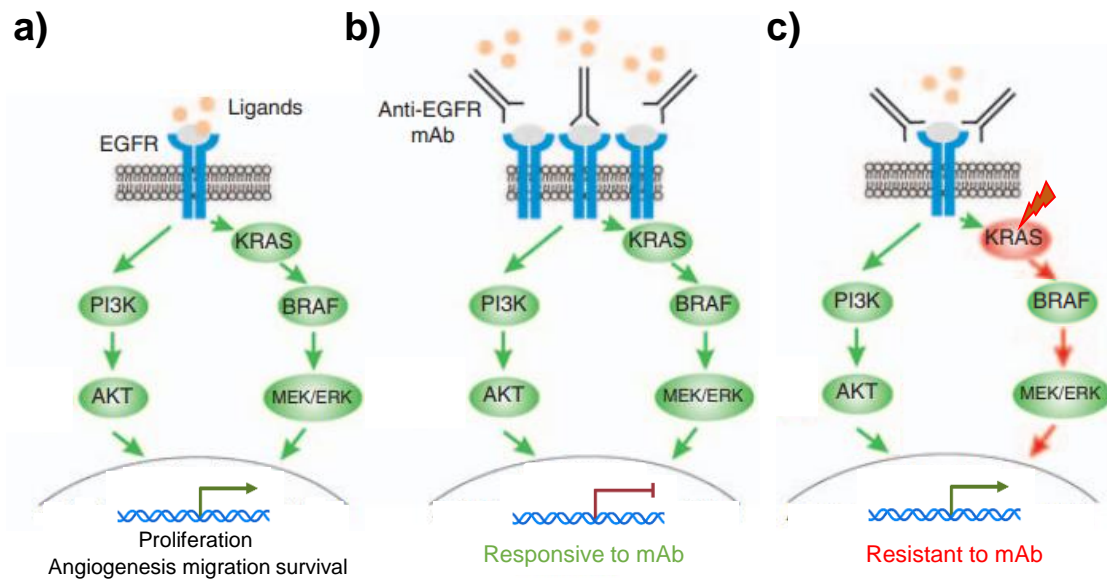
#### *1.1.4.1 Colorectal cancer*

Colorectal cancer (CRC) is currently one of the most prominent types of cancer due to its high incidence and mortality in developed countries (16). CRC occurs after a long process of malignant transformation of glandular epithelial cells located in the large intestine. Molecularly, CRC is a heterogeneous illness defined by different mutations in various receptor tyrosine kinases or downstream components of intracellular pathways activated by these receptors.

The *KRAS* pro-oncogene encodes K-Ras protein, a GTPase that participates in the RAS-MAPK signaling cascade and is involved in the transcription of critical processes in the cell, such as proliferation, maturation, differentiation, or apoptosis. The *KRAS* gene is mutated in approximately 40% of CRC. These are SNV mutations, usually restricted to codons 12, 13, and 61, all of which result in constitutive activation of the K-Ras protein (17).

In particular, codons 12 and 13 contain the small non-polar amino acid glycine. The SNV mutation causes the substitution to other amino acids that project their bulkier side chains into the GDP/GTP binding pocket. In this way, they interfere by the steric hindrance with GTP hydrolysis, reducing the intrinsic activity of Ras GTPase and its affinity for GAP, making inactivation by GAP impossible. Consequently, K-Ras remains constitutively GTP-bound and active (18). In the case of codon 61, changes of the bulky and polar glutamine reduce the rate of intrinsic GTP hydrolysis that causes K-Ras to become constitutively activated (19).

In recent years for patients with metastatic CRC (mCRC), different therapeutic strategies have been implemented to improve the antitumor effect of chemotherapy treatment and minimize the side effects of treatments. Cetuximab and panitumumab are two FDA-approved monoclonal antibodies that act as competitive antagonists to epidermal growth factor receptor (EGFR) ligands. These antibodies block the EGFR and prevent the activation of downstream pathways. Nevertheless, this elegant immunotherapy only shows favorable response rates in around 10% and 30% of cases, respectively (Figure 3) (20).



**Figure 3.** Effect of anti-EGFR monoclonal therapy on K-ras signaling. a) Overview of the EGFR pathway and its effectors. a) Treatment sensitivity for KRAS wild-type patients. b) Treatment resistance for patients with KRAS mutations. Image modified from Molecular determinants of anti-EGFR sensitivity and resistance in metastatic colorectal cancer. Figure adapted from (21).

Several clinical studies have revealed that patients without mutations in the *KRAS*, *NRAS*, and *BRAF* genes exhibit better clinical effects, while mCRC patients with mutations in these genes are not good candidates for monoclonal antibodies therapy. This is because these genes act downstream of EGFR, and thus if they are mutated, the administration of anti-EGFRs does not prevent the sustained and deregulated activation of the RAS-MAPK signaling pathway (Figure 3) (21). In addition, some studies in this field have shown that carriers of these mutations have shorter survival than colorectal cancer patients without these mutations (22).

In fact, it is standard clinical practice, recommended by the American Society of Clinical Oncology (ASCO), the National Comprehensive Cancer Network (NCCN), the European Medicines Agency (EMA), and the European Society for Medical Oncology (ESMO), to perform mutational analysis of *KRAS*, *NRAS*, and *BRAF* genes in patients with mCRC before initiating anti-EGFR therapy for its value as a predictive and prognostic biomarker (23).

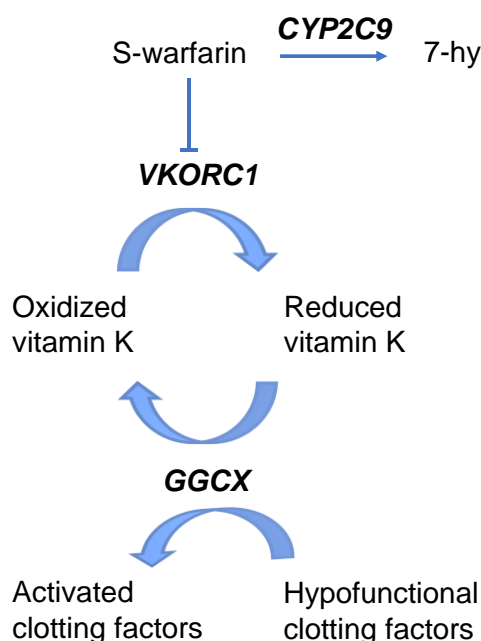
#### 1.1.4.2 Pharmacogenetics of cardiovascular diseases

Pharmacogenetics focuses on the study of the influence that genetic variations in individuals have on drug response. The pharmacological action involves both pharmacokinetic (release, absorption, distribution, metabolism, and excretion of drugs) and pharmacodynamic (interaction of the drug with its target and effect) processes. Pharmacogenetic studies have demonstrated that DNA biomarkers are responsible for

the 20 to 95% variability in drug response. In particular, the main biomarkers within pharmacogenetics are SNV.

One of the most prominent examples is cardiovascular diseases, the top cause of mortality in the western world. One million Spanish must take anticoagulant drugs every day to prevent a cardiovascular accident. Pharmacogenetics allows determining the correct dosage of certain drugs related to cardiovascular disease and thrombosis, such as warfarin, among others (24).

Warfarin is an oral vitamin K antagonist anticoagulant derived from 4-hydroxycoumarin. It inhibits the enzyme encoded by the vitamin K epoxide reductase complex 1 (*VKORC1*). Thus, it decreases the reduced form of vitamin K available for the synthesis of active coagulation factors II, VII, IX, and X and of protein C and S. Since the reduced form of vitamin K is a cofactor for the gamma-glutamyl carboxylase (*GGCX*) in the gamma-carboxylation of the glutamic acid residues of these factors. Without the carboxylation of these residues, activation of these factors will not occur, decreasing thrombin formation and fibrin binding (25). This leads to avoid coagulation. Noteworthy, the warfarin metabolization is through the enzyme encoding by the *CYP2C9* gene, which belongs to the cytochrome P450 family 2, subfamily C, polypeptide 9 (Figure 4).



**Figure 4.** Mechanism of warfarin anticoagulant action.

Pharmacogenetic studies have determined that genetic variants in *VKORC1* and *CYP2C9* explain 10-45% of the interindividual variation in warfarin dosage, depending on the studied population. The main challenge is to determine the presence of these biomarkers accurately. In addition, the dose may vary throughout the treatment course,



and an incorrect dose can cause a wide gradation of side effects associated, such as thromboembolism and bleeding (26).

The rs9923231 variant of the *VKORC1* gene (-1639G > A) is an SNV located in the promoter region. The patients with the A allele have less enzyme function; therefore, less anticoagulant is needed to inhibit its function, while those with the G allele have a higher function of the enzyme and, accordingly, will require higher warfarin doses (27). On the other hand, SNV in the *CYP2C9* gene affects the cytochrome activity level. In the European population, the most common variant is the wild type *CYP2C9\*1*, but two variants are also found in heterozygosity with one amino acid substitution for another, *CYP2C9\*2* (rs1799853), which is an R144C substitution and *CYP2C9\*3* (rs1057910) with the I359L change. Both cause a decrease in enzyme activity, and therefore these patients require lower doses of anticoagulants (28). In fact, the Clinical Pharmacogenetics Implementation Consortium (CPIC) consider that the best option before starting treatment would be to provide genotyping services for the most critical DNA biomarkers involved (29).

#### **1.1.5. Detection of biomarkers**

As seen in the previous section, medical societies consider that early SNV detection is crucial in the daily routine diagnosis and prognosis of diseases. The interest in these variants grows parallel with the great need for personalized treatment. Therefore, there is a need for tools to detect and quantify SNV. Although remarkable progress has been made in the application of SNV biomarker detection in clinical practice, there are still analytical challenges to overcome:

- **Resemblance.** Target SNV discrimination among the 6400 Mb contained in the diploid human genome must be achieved, despite their molecular, structural, and chemical similarity, as the only difference is a change in the nitrogenous base type (A, C, T, or G).

- **Amount.** Depending on the biomarker, in some cases, SNV are present from the germline in homozygosity or heterozygosity. For example, SNV biomarkers are found in a 1:1 ratio in the samples analyzed when dealing with heterozygous individuals in pharmacogenetics. However, SNV is due to somatic mutations in many situations and is found in low concentrations relative to the native variant. For instance, although it is usually also present in homozygosity or heterozygosity in the tumor cell, the proportion of mutated cancer cells is in the minority in samples. Therefore, it must be detected in the presence of a significant excess of unmutated DNA from wild-type cells. Similarly, in diagnosing some prenatal genetic diseases, the identification of certain SNV from free

fetal DNA is required, but these are found in smaller quantities than native maternal DNA (30).

- **Sampling.** The sample type can also influence the assay quality. Sampling has been evolving toward a less invasive process in all healthcare settings. The aim is to be less risky, cheaper, cause minor discomfort to the individual, not require additional processing, and be easier to handle, which implies another technological challenge to SNV detection. Thus, in early prenatal diagnosis, non-invasive tests based on the study of free fetal DNA in maternal peripheral blood are being used; saliva for some pharmacogenetic tests or nasopharyngeal samples to detect infectious diseases. Therefore, the sampling process affects the integrity of the extracted DNA, an added obstacle.

- **Preservation.** In the case of cancer, solid biopsies of tumor samples from paraffin-embedded and formalin-fixed are the most commonly used in clinical laboratories. Thus, in addition to the lower SNV concentration, the preservation process causes degradation and fragmentation of the nucleic acids, reducing the number of copies available and making the detection of the mutant even more problematic (31).

Consequently, technologies in routine medical care capable of overcoming these difficulties in SNV detection are needed to facilitate the advancement of personalized medicine.

## 1.2 CONVENTIONAL TECHNOLOGIES

In order to determine the genotype of a particular SNV, different discrimination strategies have been developed. In this section, the conventional technologies most used will be discussed.

### 1.2.1 Sequencing

**DNA sequencing** determines the arrangement of the four nucleotides A, T, C, and G in the genome sequence. The first sequencing method was proposed in 1977 by Maxam and Gilbert as a **chemical method** (32). It modified the DNA molecule with a specific nucleotide causing its breakage and subsequently separating the fragments by electrophoresis. Ten months later, Sanger et al. published the method that would become the most popular and is still used today, an **enzymatic sequencing** based on DNA replication using terminators dideoxynucleotides (ddNTPs) that inhibit DNA synthesis at specific sites. Both methods were difficult to handle and risky due to the radioactivity emitted by the gels. Afterward, it was succeeded in **automating the Sanger method** with slight improvements, producing cyclic sequencing that exponentially increased the number of products generated. Furthermore, the fragments were separated by capillary electrophoresis, accelerating the speed of the process (33).

These advances led to the signing of the Human Genome Project in 1990, and the Whole-genome Shotgun sequencing, which facilitated the assembly of the sequences. The final version of the human genome was published in 2004 at the cost of approximately 3 billion dollars (8). These very high costs stimulated the search for massive sequencing techniques known as **next-generation sequencing (NGS)**.

Thus, the **second-generation techniques** emerged, all of them with two phases that are specific to each method. The first, *in vitro* phase, in which the genomic libraries are constructed to avoid the laborious task of cloning the DNA. The second, *in silico* phase, in which the bioinformatic analysis of the results is performed.

One of the significant limitations of second-generation technologies is the requirement of advanced bioinformatics techniques and computational infrastructure to handle the amount of data generated, which is beyond the reach of most research groups and clinical care. **Third-generation sequencers** were developed to avoid these problems and reduce costs, since they do not require library construction (34). The main technologies of each generation are compared in the Table 3.

NGS techniques have offered excellent opportunities in screening novel SNV on the genome-scale, for both known and unknown, as they do not require prior knowledge of the sequence and allow their specific identification (35). However, on one hand, most

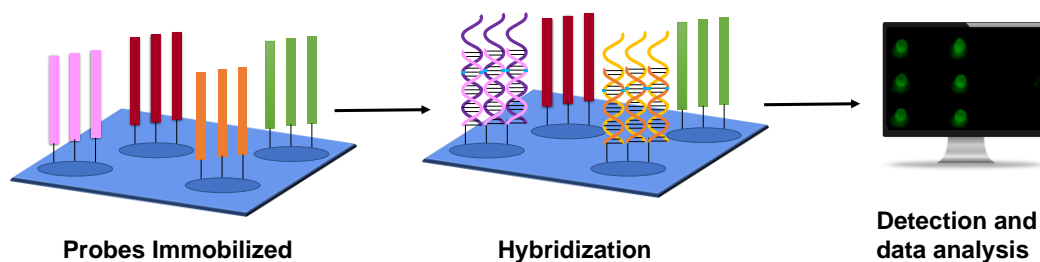
of these sequencing data are assigned to uninformative wild-type sequences, increasing the analysis complexity. On the other hand, their instrumentation (up to 10<sup>5</sup> €) and operating costs are high (up to 10<sup>3</sup> €), and analysis times are generally long. They have poor repeatability and low sensitivity (10<sup>-2</sup>-10<sup>-5</sup> mutant to wild-type DNA), producing more errors than other conventional methods. These errors tend to be false positives in base assignment and correlate with the position in the read, accumulating more frequently towards the end of the read or around homopolymeric sequences (36). These are some of the challenges facing SNV genotyping by NGS techniques to be compatible with routine clinical use.

**Table 3.** Comparison of the leading next-generation sequencing technologies.

Technology	Comments	Advantages	Disadvantages
<b>Second-generation techniques</b>			
<b>Pyrosequencing</b> (37)	Use of polymerase and emulsion PCR Chemiluminescence of pyrophosphate released during synthesis	Long reads of 400-600 bp Suitable for <i>novo</i> genome sequencing Short time (<1 day)	Confusing homopolymeric sequences High cost
<b>SOLID</b> (38)	Use of ligase and emulsion PCR Fluorescence of labeled dNTP	Lower error rate Low price	Small lengths of 35-75 bp Long time (7 days)
<b>Illumina</b> (39)	Use of polymerase and bridge PCR with reversible terminators Fluorescence of labeled dNTP	Reads of 32-250 bp High throughput	Long times (10 days) High error rate Expensive equipment
<b>Ion Torrent</b> (40)	Use of polymerase and emulsion PCR pH changes due to released H <sup>+</sup> during synthesis	Faster in reading (2h) Long reads of 100-400 bp Cheaper equipment	Elevated reagent cost High error rate in homopolymer sequences
<b>Third-generation techniques</b>			
<b>True Single Molecule Sequencing</b> (34)	Real-time fluorescence of labeled individual DNA molecules without amplification step	Need less DNA amount Simple sample preparation	Suitable for genome resequencing Small lengths of 25-45 bp Complex Instrumentation
<b>Single-Molecule Real-Time</b> (41)	Use of polymerase but not template amplification Real-time fluorescence of labeled dNTP using the "zero modes" waveguide	Short time (2-4 h) Long reads of 10 <sup>3</sup> -10 <sup>4</sup> bp	Highest error rates due to fluctuations Expensive equipment
<b>Electron microscopy</b> (42)	Read the DNA sequence directly on an electronic image	Reduces the reagents costs	Need to replicate a DNA template strand to mark the nucleotides with bases modified
<b>Nanopore</b> (43)	Specific electric currents generated by passing each DNA molecule through a nanopore	Suitable for the rapid in situ detection Portable genome sequencer	Higher error in sequence reads Need confirmation by other technologies

### 1.2.2 Microarrays

DNA microarray assays use solid support on which DNA sequences (called probes, responsible for recognition) are immobilized forming an ordered array and on which the DNA to be detected (called target) is placed. Previously the target usually undergoes an amplification process that increases its copy number and is labeled to be detected. The microarray format is based on the hybridization (recognition and subsequent binding) of complementary bases between two DNA strands. Since this binding is highly stable and selective, it allows identifying and quantifying the target present in the sample, separating it from the other components (Figure 5) (44).



**Figure 5.** Microarray technology workflow.

Its advantages include multiplexing since several samples can be analyzed simultaneously in parallel in a short time and at a reduced cost. Indeed, its small size and simplicity facilitate the miniaturization of assay devices with high selectivity using small sample volumes and different detection modes. Although prior knowledge of the sequence to be detected is mandatory. Microarrays consolidated as an excellent technique for applications such as screening, gene expression analysis, genotyping, and multiple detections of infectious agents and SNV, allowing their identification and reaching sensitivities comparable to those obtained by other reference genotyping methodologies, such as sequencing and PCR methods (45).

Hereunder it will be discussed the different key factors to achieve a correct SNV discrimination using DNA microarrays.

#### ■ Platforms

The nature of the material where the probe-target interaction occurs is essential, as the platforms' intrinsic chemical and physical properties will determine the microarray assay's performance. Typically, the surfaces used are flat and uniform.

Among the materials that can serve as supports, the most commonly used are glass, followed by silicon oxide, metals, and carbon (graphite, fullerenes, graphene, or diamond). **Synthetic organic polymers** commonly referred to as plastics, such as polycarbonate (PC), polystyrene (PS), polymethylmethacrylate (PMMA), and

polypropylene (PP) are considered an attractive alternative. In particular, PC has good mechanical properties, easy probe anchoring, low background signals, wide availability, and is very affordable. It is also hydrophobic, transparent to UV/V radiation, and does not present autofluorescence, making it a very suitable surface for microarray technology (46).

### ■ Immobilization strategies

Achieving microarrays that are reproducible, reliable, robust, and cost-effective is a demanding task. First, the **probe patterning** is performed. In this process is necessary to precisely control the temperature and humidity to avoid the coffee ring effect. This accumulates the probes in the border of the drops during their drying. The ex-situ patterning is an interesting strategy since it uses automated robotic microarrayers that deposit small volumes of probes in precise positions, reducing their consumption to picolitre-nanoliter. In addition, the contactless printing of the probes directly to the surface reduces contamination problems and damage to the deposition device or substrate and provides more excellent spot uniformity (47).

Second, the probe **immobilization** occurs i.e. the biomolecules bind to a surface losing mobility by different strategies, as compared in Table 4.

Table 4. Comparison between probe immobilization strategies.

Strategies	Type of bond	Advantages	Disadvantages
<b>Adsorption</b>	Van der Waals forces Hydrophobic interactions Hydrogen bonds Electrostatic interactions	Simple Fast Low cost	Random orientation Desorption
<b>Entrapment</b>	Retention forces	Stability High density Similar to in vivo	Random orientation Limited by mass transfer
<b>Affinity</b>	Specific interactions by bioaffinity	Simple Orderly orientation High specificity	Laborious Slow
<b>Covalent</b>	Chemical bonding between the functional probe groups and the surface	Stability Orderly orientation High density	Surface treatment Need of linker molecules

In **adsorption**, the probes are placed in contact with the support for a specific time and conditions, adsorbing them to the surface. This technique does not involve any functionalization of the support or the ligand, so it generally does not affect the activity of the receptor. The resulting layer is usually heterogeneous and randomly oriented, which can influence the hybridization performance. However, because the bond between bioreceptor and support is not very strong, changes in temperature, pH, and ionic strength can lead to desorption of the probes (48).

In **entrapment**, large quantities of probes are irreversibly retained in a polymeric material with a 3-D structure that allows flow between the reagents and sample solutions. It is helpful for cells and enzymes, although on a practical level, it presents diffusion problems producing many non-specific signals.

In **affinity** immobilization takes advantage of the specific interaction between a pair of complementary biomolecules that form a resistant complex due to their high-affinity constants, resulting in high selectivity (49). Among the most employed pair are biotin-avidin/streptavidin, Protein A and lectin-sugar. Avidin/streptavidin show excellent immobilization efficiency since both have four binding sites with a high affinity for biotin ( $10^{-14}$ ). Although, streptavidin is usually chosen because it has a low (5) isoelectric value than avidin (10.5), allowing a more specific binding (50).

In **covalent**, the probes are chemically attached to the surface through modifications at their 5' end allowing irreversible binding of biomolecules. In this approach. For this, the surface has to be activated and functionalized before probe attachment. The main drawback is the need for a linker with functional groups that serve as an anchor between the probe and the surface, such as carbodiimide, organosilanes and glutaraldehyde can be highlighted (51).

In general, the choice of an immobilization strategy is determined by the physicochemical properties of both platform and functional groups of the DNA probes. In the thesis, functionalization of biotinylated probes was selected, which will bound to streptavidin of the printing buffer, and this complex, in turn, is easily anchored to the surface by adsorption. In this way, a robust bond is achieved without functionalizing the surface, avoiding a previous and tedious step, and making it possible to control the immobilization process.

### ■ Design of probes

Probe design is a critical step in the SNV recognition process. Allele-specific oligonucleotides (ASO) probes is an attractive strategy to achieve an efficient and specific hybridization process since it allows a highly selective recognition for each SNV. This phenomenon is due to the thermal stability and base complementarity between a perfectly matched ASO probe and the target analyte, making it possible to distinguish point mutations (52). For these reasons, this type of probes has been used in this thesis.

The parameters that must be considered when designing ASO probes:

- **The length.** Short oligonucleotides (18-30 bp) are preferred as they increase selectivity since they do not result in complex conformational changes, steric hindrance, or formation of secondary structures that may interfere with the biorecognition process. On the other hand, their use decreases sensitivity. To overcome this drawback, including

a spacer between the probe and the surface is a standard procedure that improves the hybridization signal (53).

- **The sequence of oligonucleotides.** In turn, it is influenced by:
  - The percentage of G-C base content. It should be between 40% and 60%.
  - The melting temperature ( $T_m$ ) i.e., the temperature when half of the strands are in dsDNA duplex form and the other half in ssDNA.  $T_m$  is critical in complementary base pairing.
  - The cross-reaction between the probe itself and the other probes. For this, neither should be any complementary regions within the probe nor more than four repetitions of the same base.
  - The specificity with the target. When comparing the probe sequence with that of the genome, homologies of more than 70% or more than eight complementary bases in a row are found, the probe should be discarded (54).
  - The perfect-match target interaction. It is desirable to place the SNV in a central position on the probe, so that the probe-target interaction is sufficiently stable throughout the hybridization process. The stability of the double-strand can be quantified by defining the change in the hybridization reaction free energy,  $\Delta G_{hib}$ , calculated from the hydrogen bonds between each base pair and the interactions between the adjacent bases (55).

## ■ Hybridization

Hydrogen bonds are established between complementary base pairs of two DNA strands by a phenomenon called hybridization, resulting in a double helix. The use of microarrays for feasible SNV identification requires a great effort to optimize the hybridization conditions. The hybridization buffer must be studied in detail since the hydrogen bonds between the nitrogenous bases and the intramolecular forces stabilizing the double helix can be easily broken due to pH, temperature, or ionic strength changes. It is common to add reagents to achieve a more selective hybridization, such as formamide, which stabilizes ssDNA promoting the hybridization of only entirely complementary sequences and Denhardt's that saturates non-specific bonds by competition with macromolecules (56).

Factors affecting the speed, efficiency, or specificity of  $T_m$  and hybridization are summarized in Table 5. Equilibrium in astringency conditions must be found experimentally, even in the wash phase. Low stringency conditions (higher salt and low temperature) increase sensitivity but give non-specific and high background hybridization signals. In contrast, high astringency wash conditions (low salt and higher



temperature, closer to the hybridization temperature) can reduce background and only maintain the specific signal (57).

**Table 5.** Factors affecting  $T_m$  and hybridization.

Factor	Effects on $T_m$	Effects on hybridization
Temperature	Not applicable↓	Optimal < $T_m$ *
Ion strength	Increases with ↑ [Na <sup>+</sup> ]	Optimal [Na <sup>+</sup> ] ≈1M *
Destabilizing agents	Decreases with ↑ [formamide]	50% [formamide] restricts hybridization, < 50% [formamide] reduces hybridization speed, but favors specificity *
Percentage of mismatches	Decreases with ↑ % of mismatches	Every 10% decreases the hybridization rate by a factor of 2
Duplex length	Increases with ↑ length (without effect >500 bp)	Increases directly proportional with ↑ the length
Bases composition	Increases with ↑ %GC	Without effect

\* It is necessary to determine experimentally.

### 1.2.3 Enzyme-based methods

#### 1.2.3.1 Polymerases-based methods

The **Polymerase Chain Reaction (PCR)** was a technique discovered by Mullis in 1983 (58). PCR is considered one of the most significant advances in molecular biology in the 20th century due to its ability to exponentially increase the DNA sequence of interest, even if the starting amount is meager, in a simple, fast, and efficient manner. It is based on implementing in the laboratory the use of polymerases. These natural enzymes are responsible for cell replication, adding complementary nucleotides to a strand that serves as a template in the 5'-3' sense. Most polymerases cannot start synthesizing a chain of nucleotides and only can add nucleotides to a pre-existing chain, so it is necessary to use initiators or primers. The primers are short (18-20 bp) ssDNA that provide the specificity of the method since they define the region to be amplified within the genome. For each region, a pair of primers (forward and reverse) is used, and each one is complementary to one of the extremes of the DNA strands targets.

PCR consists of three-stage thermocycling process that repeats itself. First, denaturation, in which the hydrogen bonds that join the two chains are broken when the temperature is raised above 95 °C. Second, primer annealing, in which stable hydrogen bonds are formed between the primers and the complementary target sequence. For that, the temperature of the stage is adjusted to the  $T_m$  of the primers. Lately, with polymerase extension, the enzyme adds nucleotides to the new chain when the temperature is modified to its working optimum. Furthermore, the entire PCR process is

automated by a thermocycler, which controls the temperature required during the stages of each cycle (59).

Over the years, the PCR technique had improved and perfected. On one hand, *E. coli* polymerase, which denatured at high temperatures, was replaced by thermostable polymerases from thermophilic microorganisms resistant to high temperatures, being currently the most used Taq polymerase from *Thermus aquaticus* (60). Indeed, another of the first modifications was **PCR multiplex**, first described in 1988 by Chamberlanin JS et al., that use two or more sets of primers to simultaneously amplify different DNA targets in the same reaction, reducing assay time and reagent consumption. Nevertheless, it requires correct primer design to ensure method specificity (61). Concerning SNV detection, different variants have been developed that achieve excellent analytical performances, which will be discussed in the third section of this thesis.

On the other hand, post-PCR techniques had been established. **High-Resolution Melting (HRM)** was developed in 1997 by Ririe et al. and is based on increasing gradually the temperature and the use of a fluorophore, which is intercalated nonspecifically in the dsDNA amplicons. Thus, the amplicons are denatured, and the fluorescence emitted is reduced, obtaining a specific melting curve for each DNA sequence and thus differentiating between SNV, both known and unknown. Nowadays, HRM is widely used for SNV discrimination because it is fast, has a closed-tube format, has an outstanding cost/effectiveness ratio, and has high sensitivity. Its limit of detection of mutations in the presence of native DNA is generally between 3 and 10% (62). Different mechanisms have been adapted to obtain melting curves to improve SNV recognition from continuous monitorization of fluorescence during temperature changes, as shown in Table 6.

**Table 6.** SNV detection by HRM.

Mechanism	Target	Referencia
Inherent quenching of deoxyguanosine nucleotides of a single-labeled probe	Factor V Leiden <sup>1691G&gt;A</sup> MTHFR <sup>1298A&gt;C</sup> Hemoglobin S <sup>17A&gt;T</sup>	(63)
Fluorescence difference of LCGreen binding to dsDNA amplicon to identify heterozygotes and homozygotes	Six common $\beta$ -globin genotypes (AA, AS, AC, SS, CC, and SC)	(64)
Fluorescence difference of LCGreen I binding to small dsDNA amplicon increasing the Tm between homozygotes	Factor V Leiden <sup>1691G&gt;A</sup> prothrombin <sup>20210G&gt;A</sup> MTHFR <sup>1298A&gt;C</sup> HFE <sup>187C&gt;G</sup> Hemoglobin S <sup>17A&gt;T</sup>	(65)
3'-blocked, unlabeled probe and LCGreen	<i>BRAF</i> <sup>V600E</sup>	(66)
Unlabeled probe modified with LNA improving mismatch discrimination	Factor V Leiden <sup>1691G&gt;A</sup>	(67)
Snapback primer genotyping with saturating dyes	CFTR <sup>G542X</sup> CFTR <sup>I506V</sup> CFTR <sup>F508C</sup>	(68)
Microfluidic platform tailored to rapid serial PCR and high-speed melting	F2 <sup>c.97G&gt;A</sup> F5 <sup>c.1601G&gt;A</sup> , MTHFR <sup>c.665C&gt;T</sup> MTHFR <sup>c.1286A&gt;C</sup>	(69)

## ■ Formats of PCR

In terms of reaction kinetics, four phases can be distinguished (Figure 6a). First, an initial screening phase, where there is still no detectable amplification because tracking and binding of the primers to their target occur. A second phase, the exponential amplification, during which two double helices are synthesized at the end of a cycle, each consisting of a new daughter strand and the original strand. After repeating the cycles, a quantity equal to  $2^n$  is obtained, where  $n$  is the number of cycles. The third phase is where linear amplification is produced, where the growth of the reaction slows down because reagents are being consumed. Finally, in the late or plateau phase, saturation is reached since no more product is generated even though the number of cycles increases. It is due to the depletion of components and competition of the product found in enormous amounts for the binding sites of the primers.

Concerning the monitoring of the PCR reaction, it can distinguish:

- **End-point or conventional PCR** in which the products generated with the SNV of interest are detected after the reaction, usually by gel electrophoresis or fluorescence. As measurements are in the plateau phase, they are less accurate due to variations in reaction kinetics (Figure 6a). This approach only requires simple thermal cyclers costing around 2,000-3,000 €.

▪ **Real-time or quantitative PCR (qPCR)** allows the entire amplification process to be monitored, quantifying the products generated in the exponential phase (Figure 6a). qPCR requires the inclusion of references or monitoring standards in each determination and the construction of a curve; therefore, it is a relative quantification. Products are detected by incorporating a reporter into the conventional PCR reaction that emits fluorescence, such that the signal increase is proportional to the amount of DNA amplicon (70). qPCR needs a thermal cycler equipped with fluorescence sensors, which costs approximately 8,000-10,000 €, and the reporters can be:

➤ Non-specific fluorochromes, which are compounds that emit fluorescence when they are bound to dsDNA. The most used is SYBR Green I, which intercalates into the minor groove of dsDNA. Among its advantages are that it is cheap and that the same reagents can be used to analyze any fragment. In contrast, its main disadvantage is its lack of specificity because it can bind to primer-dimers or unspecific products, although this can be solved by performing a melting curve after PCR (71).

➤ Specific probes, which are short-functionalized oligonucleotides that hybridizes with the target SNV. Three different types can be distinguished, hydrolysis, hairpin, and hybridization probes, being the most representative for each case TaqMan, molecular beacons and fluorescein-LC Red probes. All are based on the use of fluorescence resonance energy transfer (FRET). The primers and the probes provide specificity in all cases, so they do not require melting curves. Still, this strategy needs the proper design of specific probes for each target, increasing the cost and assay complexity (72).

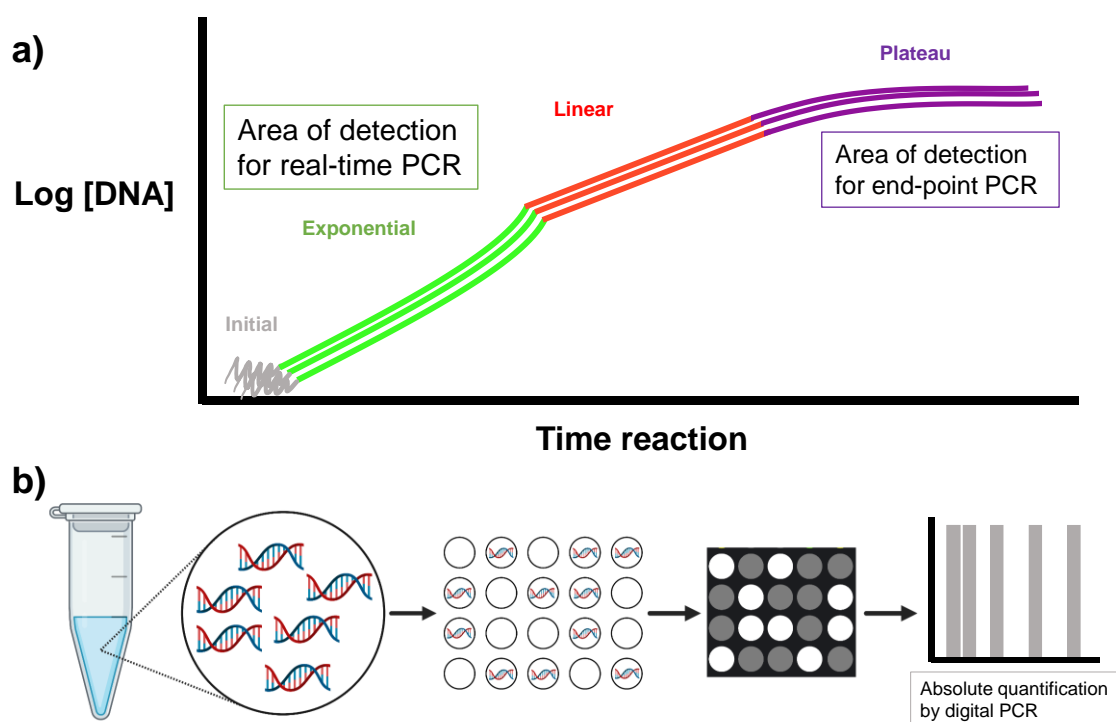
▪ **Digital PCR (dPCR)** allows absolute quantification. It is based on the partitioning of a sample into individual molecules in wells. Following PCR reactions, some amplicons that are generated contain the SNV of interest, while others do not, and will be used to count the exact number of target molecules accurately, sensitively, and without the need for calibration curves (Figure 6b). Though, the sensitivity of this technique is variable, as LOD is defined by the number of wells that can analyze (73). The idea that the PCR mixture can be compartmentalized into microdroplets was emerged to overcome this drawback. For that, single-molecule PCR is performed in the same microenvironment. It is the basis of the approaches:

➤ **Bead, Emulsion, Amplification, and Magnetisation (BEAMing)**, designed by Dressman et al. in 2003, involves magnetic particles attached to primers and template DNA compartmentalized in microemulsions. Each microemulsion contains a DNA template and a particle that allows single-molecule PCR. Upon completion, the emulsions are broken and magnetically collected. Then, they are incubated with the designed oligonucleotides achieving specific hybridization in solid phase. The particles

are counted by measuring the fluorescence they emit using flow cytometry allowing SNV detection (74).

➤ **Droplet Digital PCR (ddPCR)**, in 2011, Hindson et al. employed a droplet system from an oil-water emulsion so that a specific microfluidic system produces a massive partition of the sample into 20,000 droplets. These droplets work in the same way as individual tubes in a conventional PCR, but with the advantage that the format is much smaller (nanolitre scale), and amplification occurs massively. After PCR, each droplet is detected by flow cytometry and classified as positive or negative for SNV amplification. In turn, the DNA generated is quantified by fluorescence. Unlike the previous one, this is a homogeneous assay since all phases are performed in the emulsion state (75).

Both assays allow higher precision, absolute quantification, and high selectivity (from  $10^{-3}$  to  $10^{-8}$ ) compared to other PCR systems. However, dPCR is still time-consuming and requires specifically designed high-cost reagents (up to 20 € per sample) and expensive specialized instruments (up to  $10^5$  €), which are challenging to implement in routine diagnostic laboratories (76).



**Figure 6.** Formats of PCR monitoring. a) PCR reaction phases and their detection depending on using real-time PCR or end-point PCR. b) Absolute quantification mode by digital PCR.

### 1.2.3.2 Ligation-based methods

PCR has facilitated the development of a variety of detection systems to determine a nucleotide base mutation. The **ligase chain reaction (LCR)** has been developed to complement or as an alternative to PCR. Like PCR, it can amplify a specific DNA sequence. Since its discovery in 1991 by Barany, LCR has become an up-and-coming diagnostic technique for detecting these variants (77).

The fundamentals of LCR are the same as those applied in the **oligonucleotide ligation assay (OLA)**. It employs a DNA ligase enzyme, two allele-specific oligos, and a common probe aligned in adjacent positions, leaving a nick that the enzyme will seal through a phosphodiester bond. As a requirement, primers must be phosphorylated at their 5' end since a 3' hydroxyl-free group and a 5' phosphoryl end are required for ligase to act. The ligation only occurs when there is a perfect match between the last nucleotide of the primer and the DNA strand since DNA ligases have high specificity and fidelity. For this reason, ligation has been applied for SNV discrimination. However, this discrimination assay requires an amplification step to achieve adequate sensitivity (78). This drawback was solved with LCR, which uses a thermostable DNA ligase (usually *Thermus aquaticus* ligase), two primer pairs (4 in total), and ATP or NAD<sup>+</sup> as cofactor depending on the type of ligase used. Each primers pair hybridizes to one strand of the dsDNA. The amplification takes place in cycles, thus requiring a thermal cycler. It consists of a denaturation step (95 °C), primer annealing (the temperature varies according to the T<sub>m</sub> of the primer), and a ligation step (at the optimal temperature of the ligase). As a result, a fragment is generated exponentially for both strands in each cycle, which acts as a template in subsequent cycles (79).

LCR has higher specificity (100% vs. 90%) and sensitivity (97.6% vs. 99.5%) than PCR, but its main disadvantages are target-independent ligation and that it requires a large number of modified probes. On the other hand, with PCR, target-independent amplification results in primer-dimers of lower molecular weight and, therefore, easily distinguishable. However, with LCR, a product of the same size is produced, thus increasing background noise and leading to false positives.

Modifications of LCR have been reported to overcome these disadvantages (80). Table 7 shows those applied to SNV detection. The **ligase detection reaction (LDR)**, produces linear amplification, has multiplex capacity, but low sensitivity (81). The **Gap-Ligase Chain Reaction (GLCR)**, improves sensitivity, but requires a DNA polymerase and does not have multiplex capacity (82). Other alternatives were the **Real-Time Gap-LCR (QGLCR)**, which uses the FRET phenomenon to quantify the amount of starting

DNA (83) and the **Asymmetric Gap-Ligase chain reaction (AGLCR)**, which serves to detect SNV in RNA by pre-step with reverse transcriptase (84).

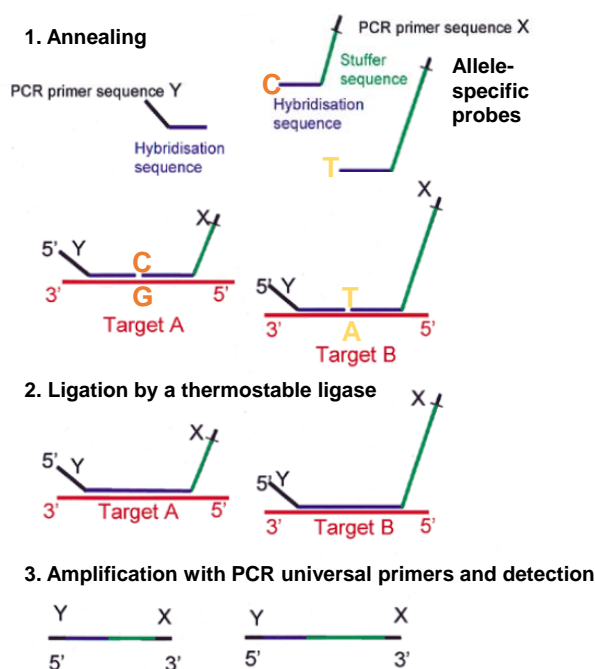
**Table 7.** Different strategies of ligation-based methods.

Strategies	Detection	LOD*	SNV detected
LDR	Microarray SERS	0.4-1 fmol	Codons 12-13 in <i>KRAS</i> SNV in <i>TP53</i>
	Capillary electrophoresis Colorimetric by magnetic beads and gold nanoparticles		Species of malaria
LCR	Fluorescence Colorimetric	10 copies	<i>HBsAg<sup>G154R</sup></i> Codons 12-13 in <i>KRAS</i>
	Electrochemiluminescence and magnetic beads Quantum dot		
QGLCR	FRET using real-time PCR equipment	0.001% mutant	<i>PRNP<sup>A136V</sup></i> SNV in <i>TP53</i> Codons 12-13 in <i>KRAS</i> CD17 (A>T) in <i>HBB</i>
AGLCR	Microparticle capture enzyme immunoassay	>20 copies	HCV RNA 45
SNIPlex genotyping system	Capillary electrophoresis Genetic analyzer (3730xl DNA Analyzer)	>1 ng	SNV in <i>Vitis vinifera</i> L. 521 SNV in the human genome
MLPA	Capillary electrophoresis Genetic analyzer	25 pmol	7 circulating markers of breast cancer SNV in <i>HBB</i>
	Microarray Electrochemical		<i>Yersinia enterocolitica</i>

\* **LOD:** expressed in percentage (%) or quantity (fmol, amol, copies, pmol), according to the results of each work.

Ligation has also been combined with amplification methods such as PCR. In this context, the **SNIPlex Genotyping System** stands out. With a tiny initial amount of target DNA and four different probes, several hundred of SNV can be genotyped with high sensitivity and specificity (85). The main problem is that it takes more than two days to perform the whole assay and the equipment required is a unique and expensive detector manufactured by Applied Biosystems (3730xl DNA Analyzer) (86).

Currently, the most widely used LCR-based SNV genotyping platform is **Multiplexed Ligation Dependent Probe Amplification (MLPA)**. MLPA is a robust, straightforward technique that can amplify several targets in a single reaction. Only the hybridized probes can be ligated and then amplified by PCR using primers with universal sequences (Figure 7) (87). Since 2002, MLPA has been commercialized as a trademark of MRC-Holland, allowing multiplex detection of at least 50 SNV at a time. Notwithstanding, it has the disadvantage of the high cost of the detection equipment (genetic analyzer). In addition, probe design and optimization of ligation conditions are difficult and complex, resulting in differences in amplification yield (88).



**Figure 7.** MLPA reaction phases. Adapted from (87) for SNV genotyping.

### 1.2.3.3 Endonucleases

Endonucleases are restriction enzymes that recognize a specific sequence in DNA and catalyze the cleavage of alpha-type phosphodiester bonds (5'-P, 3'-OH). Most of them cut within or very close to the recognized sequence and usually recognize palindromic sequences of 4-8 bp. Due to the high specificity of restriction enzymes, it is a widely used method of SNV discrimination in combination with different strategies (89).

Initially, **Southern blotting analysis** was used in conjunction with restriction enzymes to detect SNV by generation or ablation of restriction sites resulting from nucleotide substitution. However, this method required large amount of genomic DNA (gDNA) and was tedious. **Restriction Fragment Length Polymorphism PCR (RFLP-PCR)** emerged as an alternative. RFLP-PCR involves PCR amplification followed by digestion with an enzyme that selectively cuts mutant DNA, and the resulting fragments are run on an electrophoresis gel. This approach requires a small amount of DNA and has a sensitivity around  $10^{-3}$ , but cannot detect rare mutations (90). For example, it has recently been used to detect D614G mutation in SARS-CoV-2 (91).

However, the application of these assays requires that SNV must be in the region recognizable by the endonucleases, limiting their scope of application. In addition, there is the disadvantage that restriction enzymes lack high selectivity, for example: EcoRI cuts G/AATTC, but under suboptimal conditions, it can cut N/AATTN (92).

**Invader** assays using a restriction enzyme called **cleavage** which recognize a specific structure were developed to overcome these difficulties. They utilize an allele-



specific probe and an invasive oligonucleotide that hybridizes to the target and overlaps at the SNV site. The cleavage recognizes the overlapping structure formed by the target and the two probes and cleaves the allele-specific probe, releasing its 5' end, known as the flap, which serves to report the SNV (93). Results can be visualized by electrophoresis, mass spectrometry, or FRET, among other approaches. These assays can be multiplexed by designing multiple probes and using a specific flap for each SNV. A work was recently published using this strategy coupled with a universal microarray for genotyping 14 SNV in the *BRCA* gene (94).

However, experiments with Invaders require high amounts of target DNA or a previous amplification step to improve sensitivity; for this reason, raising the temperature in the reactions was performed. In this way, once the specific allele probe was cut, it could be denatured, initiating successive cleavage cycles. Hence, multiple flaps were generated per target, amplifying the signal. An example of this strategy was detected by two successive invasive reactions and FRET detection of 0.02% of mutant EGFR in the background wild-type (95). Its main drawback is the high cost of using multiple modified probes, as well as the complexity in the design and finding the working conditions of the reaction (96).

### **1.3 NEW BIORECOGNITION STRATEGIES**

As previously discussed, SNV are difficult to genotype because many are present at trace levels and only differ from the wild type by a single nucleotide. SNV genotyping methods require high sensitivity and specificity of target sequences. In addition, factors such as high robustness, speed, reproducibility, and low cost, must also be considered. Many of the current techniques do not achieve these performances and require centralized and well-equipped laboratories (92).

Combining molecular recognition techniques that are compatible with the daily routine with transduction systems are emerging solutions to facilitate the diagnosis and prognosis of the patients. The most innovative approaches described in the literature in recent years are discussed below.

#### **1.3.1 Improvement of PCR methods**

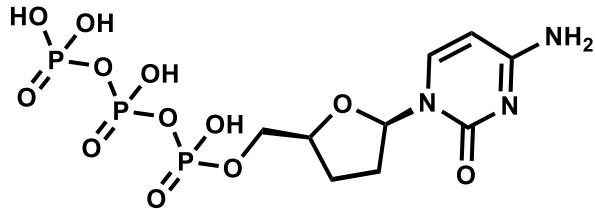
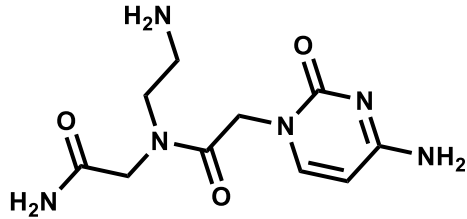
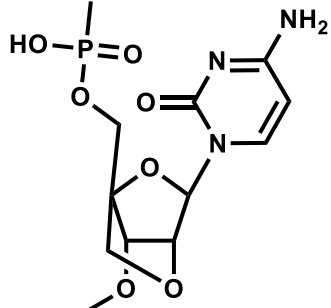
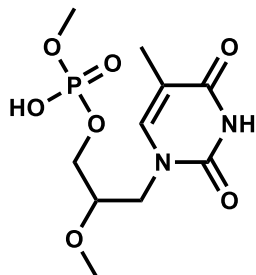
Low-cost PCR-based amplification systems accessible to minimally equipped laboratories have been developed to improve SNV discrimination performance and detect them with greater sensitivity and selectivity. Most of these methods use real-time PCR equipment, which is very common in clinical laboratories, so there is no need for post-processing that would increase complexity, cost, and duration of the assay (97). In this thesis, the most significant ones have been grouped according to their strategy.

- **Allele-specific primers.** These strategies are based on the selective amplification of SNV through the specific design of primers. They are simple, fast, and cost-effective strategies that rely on the ability of the DNA polymerase enzyme to synthesize in the 5-3' direction only when there is an entirely complementary primer binding to the template DNA, on the destabilization generated by a mismatch at the 3' position in the primer-template duplex, and on the lack of 3'-5' exonuclease activity of some polymerases (98).

They typically have low to moderate selectivity ( $10^{-1}$  to  $10^{-3}$ ) and generate false positives when the mutant content is less than 1-5% due to inadvertent amplification of wild-type DNA, resulting in a false mutant template that propagates in future PCR cycles. These results are improved by introducing near the 3' end additional mismatches. They are only suitable for known SNV, and both the specific amplification and the yield of the reaction will depend on the position of the mismatch and type of nucleotide incorporated (99). These techniques are known as Allele-Specific PCR (AS-PCR) (100), Amplification Refractory Mutation System PCR (ARMS-PCR) (101), or PCR amplification of specific alleles (PASA) (102).

▪ **Blocking agent.** These strategies incorporate a blocker agent in the amplification reaction, a oligonucleotide complementary to the wild-type allele. It hybridizes preferentially interfering in the primer annealing or strand elongation that does not contain the SNV of interest (103). In this way, mutant sequences are preferentially amplified, blocking native amplification. It is required that the blockers do not function as primers for DNA polymerases. Therefore, the oligonucleotides used are modified with different functionalization. The blocking agents commonly employed in the literature are presented in Table 8.

**Table 8.** Characteristics of commonly used oligonucleotides.

Name and chemical structure	Features
<p data-bbox="240 725 616 757">Dideoxynucleotides (ddNTP)</p> 	<p data-bbox="868 815 1300 882">Nucleotides lacking a 3'-OH group on the deoxyribose.</p>
<p data-bbox="240 1010 608 1041">Peptide nucleic acids (PNA)</p> 	<p data-bbox="868 1032 1300 1160">Structural analogs of DNA consisting of repeating N-(2-aminoethyl)-glycine units linked by peptide bonds.</p> <p data-bbox="868 1171 1300 1272">They contain neither the DNA pentose nor the phosphate groups. They are not negatively charged.</p>
<p data-bbox="240 1301 600 1332">Locked nucleic acids (LNA)</p> 	<p data-bbox="868 1420 1300 1554">Analogous to RNA, but the ribose ring is blocked by a methylene bridge connecting the 2'-O atom and 4'-C on the ribose.</p>
<p data-bbox="240 1682 576 1713">Xeno nucleic acids (XNA)</p> 	<p data-bbox="868 1805 1300 1906">Synthetic skeletal analogs of nucleic acids that have different sugars than natural nucleic acids.</p>

An appealing property of modified nucleic acids (PNA, LNA, XNA) is that hybridization is more thermodynamically favorable with perfectly matched DNA and any mismatch in the duplex destabilizes the binding (104,105).

The most commonly used blocker design strategies for binding to the native sequence are:

- Overlap with the primer (total or partial).
- Adjacent to the last nucleotide of the primer.
- Between the two target sites of the primer.

The selectivity of this technology is in the range of  $10^{-2}$  to  $10^{-5}$  mutant to wild-type DNA, so it detects low abundance mutations, but it requires previous steps to optimize and design the blockers properly (106,107).

The following Table 9 compares the characteristics of the main SNV genotyping techniques that add modifications or combine the two described strategies to achieve improvements aimed at solving the current problem of SNV detection in clinical sample

**Table 9.** Characteristics of the main SNV genotyping techniques from modifications or combination of the described strategies.

Name	Criteria	Advantages	Drawbacks	SNV	LOD (%mutant)	Reference
<b>AS- real time</b>	AS amplification combined with TaqMan in real-time.	Quantitative Simplicity Multiplexing	Not appropriate for the detection of rare alleles.	<i>BRAF</i> <sup>V600E</sup>	10	(108)
<b>Tetra-primer ARMS</b>	Two external primers located at different distances from the SNV Two internal ones with AS mismatches. An ARMS primer is covalently linked to a blocking sequence that, in turn, is coupled to a molecular beacon probe.	Reduces false positives. Single reaction tube.	Requires a final electrophoresis step. The inner/outer primer concentration ratio should be optimized. Not suitable for GC-rich SNV.	<i>JAK2</i> <sup>V617F</sup>	1	(109)
<b>Scorpions ARMS</b>	An ARMS primer is covalently linked to a blocking sequence that, in turn, is coupled to a molecular beacon probe.	Real-time Unimolecular interaction Accuracy Shorter reaction times	Scorpion probe design is very complex and tedious. Difficult to adapt to different assays.	<i>EGFR</i> <sup>L858R</sup>	1	(110)
<b>MAMA</b>	Add glycerol. Primers with a single mismatch for the mutant allele and double mismatches for wt.	Short reaction time (2-step PCR) Lower reagent concentration Inexpensive	Requires a final electrophoresis step. The extension efficiency depends on the type and nature of the mismatches.	<i>KRAS</i> <sup>G12D</sup>	0.1	(111)
<b>TaqMAMA</b>	Combines Taq probes with MAMA-PCR.	Can be directly and rapidly detected and quantified by fluorescence in a single PCR tube.	Requires an additional probe in addition to the primers. Can cause false positives depending on the position and type of mismatch.	<i>HPRT</i> <sup>c.215A&gt;G</sup> <i>HPRT</i> <sup>c.538G&gt;T</sup> <i>HPRT</i> <sup>c.580G&gt;C</sup>	0.1	(112)
<b>PAP-ASA</b>	Combines AS-amplification with pyrophosphate-activated polymerization (PAP).	Increased specificity because two independent processes must occur for an erroneous result: pyrophosphorolysis of mismatches and DNApol failure.	Theoretical results do not agree with practice because PAP efficiency will vary depending on terminal ddNMP and oligonucleotide size.	<i>DRD1</i> <sup>c.229A&gt;G</sup>	0.1	(113)

(continue)

(continued)

<b>Bi-PAP-ASA</b>	Uses two opposing 3-terminal blocked phosphorylated oligos with one nucleotide overlap at 3' end specific for the SNV	High selectivity because PAP is simultaneously performed in two directions	Requires a final electrophoresis step and a genetically engineered DNA polymerase with high pyrophosphorylation activity. Without the addition of PPi, false positives can be generated because oligonucleotide dimers cause non-specific pyrophosphorolysis.	<i>TP53</i> <sup>g.14060G4C</sup> <i>TP53</i> <sup>g.14060G4A</sup>	0.1	(114)
<b>ARMS-PNA/ AS-PNA</b>	An PNA replaces the nucleotide at the 3' end of the primers.	The high affinity of PNA for its target reduced false-positive results.	Requires a significant optimization in the adjustment of the Tm of the primers according to their concentration.	<i>EGFR</i> <sup>L858R</sup>	0.1-1	(115)
<b>LNA-AS-qPCR</b>	LNA incorporation into real-time probes.	Wild-type blocking allows genotyping of mutants	No resistance to exonucleases. Requires experimental design processes and optimization of the procedure.	<i>HFE</i> <sup>H63D</sup> <i>HFE</i> <sup>C282Y</sup>	0.1-1	(116)
<b>Chimers LNA-PNA</b>	PNA-containing oligonucleotide to block wild-type sequences and a fluorescent probe including LNA to detect the mutant amplicon.	PNA/LNA duplexes are more stable than PNA/ DNA or LNA/DNA duplexes	Expensive oligonucleotides and complex design.	<i>EGFR</i> <sup>L858R</sup>	0.1-1	(117)
<b>LATE -CLAMP- PCR</b>	Asymmetric PCR combined with clamp blocker and changes in the thermocycling process.	Allows the introduction of a detection step distinct from the hybridization step. Avoids the plateau phase of PCR, where there is no correlation between the total amount of PCR product and the initial target amount.	Hairpins blockers are very stable in solution. So, it is difficult to find the exact temperature that makes them thermodynamically favorable to bind to the target.	<i>KRAS</i> <sup>codon 12</sup>	0.1-1	(118)

dNMP; deoxyribonucleoside monophosphate

▪ **Modification of thermocycling.** They are based on including steps during thermocycling process with specific temperatures to amplify the SNV preferentially.

Within this group, it is worth highlighting a new technique had been developed for Li et al. in 2008, called **Co-amplification at a Lower Denaturation Temperature (COLD-PCR)**. Between the denaturation and primer banding stage, the temperature is lowered so that duplexes are re-formed between the alleles present, giving rise to both homoduplexes and heteroduplexes with a mismatch between wt and mutant. Subsequently, a critical temperature ( $T_c$ ) is used in thermocycling below  $T_m$ , which causes heteroduplexes to denature preferentially to homoduplexes, thus amplifying the mutated alleles and improving their detection by up to 100-fold (119). This technique has high sensitivity; it can detect  $10^{-1}$  to  $10^{-4}$  mutant alleles, both known and unknown. Moreover, it can be used in place of conventional PCR and can be combined with most existing assays due to its simplicity without requiring additional labor costs, as it does not use allele-specific primers, probes, or enzymes (106).

According to the bibliography, the different variants of this technique are presented in the Table 10. However, all require a significant amount of work to determine the correct critical denaturation temperature, and slight experimental variations ( $\pm 0.2$  °C) can altogether abolish mutation enrichment, so thermal cyclers with precise temperature definition are required. Furthermore, mutation enrichment depends on the sequence context, and therefore specific mutations in a DNA sequence may be more challenging to detect than others, so there is no guarantee that all low-level mutations are preferentially enriched (120,121).

**Table 10.** Comparison of COLD-PCR variants.

Variant	Full-COLD-PCR	Fast-COLD-PCR	Ice-COLD-PCR	Enhanced-Ice-COLD-PCR
<b>Features</b>	Needs an intermediate hybridization step	Eliminates intermediate denaturation and annealing steps	Adds a blocking oligonucleotide	It employs a modified blocking oligo with LNA.
<b>Advantages</b>	Enriches all possible mutations	Lower analysis time.	Improved wt-mutant selectivity.	Significant improvement of wt-mutant selectivity.
<b>Disadvantages</b>	The reaction is slower. The wt of the heteroduplex is also amplified.	Only homoduplexes with lower $T_m$ than wt are amplified.	Requires precise control of working temperatures.	Significant increase in the cost and complexity
<b>References</b>	(119)	(122)	(123)	(124)

### 1.3.2 Alternative methods to PCR

Although the development of analytical detection methods performed in a conventional laboratory setting remains an open field of research and innovation, in recent years there has been a focus on decentralization of molecular diagnostics. The development of portable platforms that can be applied to tediagnosis or in resource-poor regions where medical care is limited has become a primary goal. The main challenges are to achieve devices with a simple manufacturing procedure, using affordable and/or reusable materials; as well as to develop a user-friendly technology that exempts the supervision of a specialized professional, so that any user can run the test autonomously at home, in doctors' offices, in outpatient clinics, etc. This strategy is making it possible to expand the number of clinical scenarios where genetic biomarker-based methods support diagnosis. It is achieving a breakthrough for the rapid detection and addressing of emerging public health threats, as is being demonstrated with the SARS-Covid-19 crisis and facilitates the treatment of chronic diseases that require constant monitoring (125).

In this regard, other attractive alternatives have emerged to determine the levels of these biomarkers by expanding their use through point-of-care (POC) tests that bring together all the capabilities to provide information close to the patient. POC tools are portable, simple, affordable, use small sample and reagent volumes, and have miniaturization capabilities. At the same time they maintain sensitivity, specificity and robustness comparable to conventional methods used in large hospital laboratories.

In fact, the World Health Organization established a set of criteria for the ideal characteristics that all POC testing devices should meet, which are summarized under the acronym ASSURED: low cost, sensitive, specific, easy to use, fast, robust, not requiring bulky equipment and available to those who need it (126).

#### 1.3.2.1 CRISPR/Cas methods

It was a significant finding to discover that prokaryotes had evolved an adaptive immunity that allowed them to distinguish "foreign" DNA from "self" DNA and eliminate the foreign one. This process is carried out by DNA sequences grouped under the name **Regularly Clustered Interleaved Short Interspaced Palindromic Repeats**, abbreviated **CRISPR** (first characterized in 1993 by Francisco Mojica). It involves genes encoding a large and heterogeneous family of proteins called **Cas** (meaning CRISPR-associated) (127).

Since 2012 CRISPR has been used for genome editing applications surpassing the previously used nucleases (zinc finger nucleases and TALENs) because they are



much less expensive, time-consuming, more precise, and scalable. When this system is used for biotechnological purposes, it is required:

- **Endonuclease.** The most used is Cas9 from the bacterium *Streptococcus pyogenes*.
- **Guide RNA.** It consists of spacer RNA, including the 20-nucleotide target sequence and the scaffold RNA, the sequence necessary for Cas9 binding to the gRNA.
- **Presence of a Protospacer Adjacent Motif (PAM).** It consists in few nucleotides in the target DNA that serve as the binding signal of Cas9.

The Cas9-gRNA complex will bind to any genomic sequence with a PAM if the gRNA spacer is complementary to the target. In this way, cleavage of the three nucleotides upstream of the PAM sequence will occur. The target can be any DNA sequence of approximately 20 nucleotides, as long as it meets two conditions: it is unique in the genome, and it is located just upstream of a PAM. Thus, potential target sites are those that are 20 nucleotides upstream of the PAM (128).

Although PAM sequences for Cas9 are abundant throughout the human genome, they are not always correctly positioned to target particular genes. In addition, a sequence may have sufficient homology with the target one to recruit Cas9 to other parts of the genome. These sequences can be cut, referred to as "off-target effects," so mutations can occur unintentionally during attempts to use CRISPR-mediated genetic engineering.

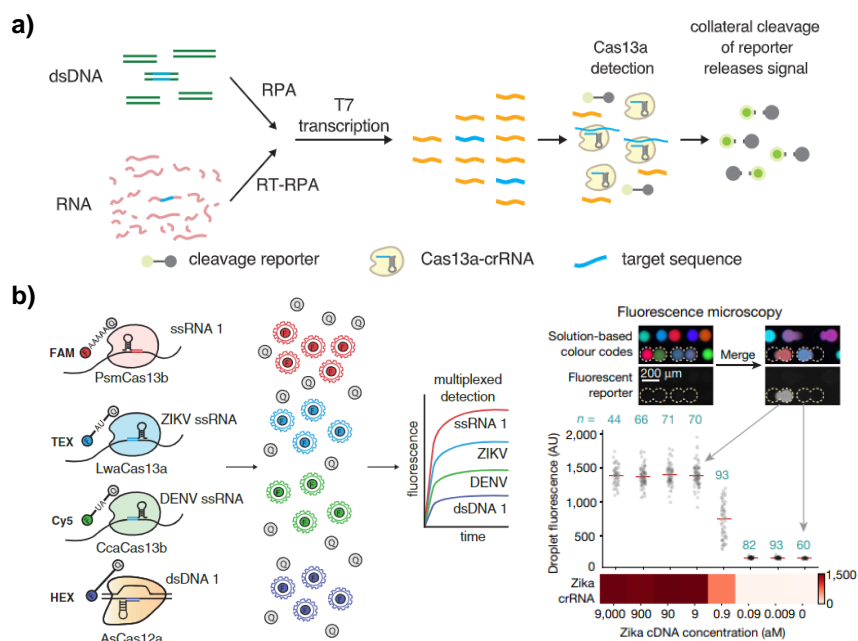
On the other hand, excision efficiency may increase or decrease depending on the specific nucleotides within the selected target sequence. Therefore, carefully examining the predicted activity within and outside the target sequence is necessary. Several gRNA designs programs have been developed to locate potential PAMs and target sequences. Also, modifications of synthetic nucleases, such as eSpCas9 and SpCas9-HF1, have been made to alter Cas9 activity and decrease off-target effects. On the other hand, as the requirement for a single acceptable PAM sequence remains a technical hurdle in utilizing CRISPR throughout the human genome, Cas9 homologs from different organisms have also been characterized (129).

In gene editing, once the cut is produced in the dsDNA, it will be repaired by one of the recombination mechanisms, which has been the application of CRISPR until very recently. Notwithstanding, since 2017, this strategy is also exploited for nucleic acid detection thanks to discovering new nucleases. These approaches using CRISPR as a discrimination strategy would stay in this step, avoiding the recombination process, and the cut obtained would be combined with different detection strategies.

Gootenberg et al. were the first to use this strategy as an analytical tool. They called it **Specific High Sensitivity Enzymatic Reporter UnLOCKing (SHERLOCK)**

and detected SNV with high sensitivity (attomoles). They used a new ribonuclease called Cas13a that can be programmed with gRNA to recognize a specific RNA target. So that, if starting from DNA, its transcription to RNA is previously required. In addition, Cas13a has collateral cleavage activity, which means that once it cuts its target, its nuclease activity continues and indiscriminately cuts nearby fluorescent RNA reporter molecules, allowing SNV to be detected. As the reporters are cut collaterally by Cas13a, the fluorescence quenching is avoided, and a fluorescent signal is generated that can be correlated with the initial target amount. In this way, they detected 0.1% of *BRAF*<sup>V600E</sup> and *EGFR*<sup>L858R</sup> mutants against the wild-type background. However, they needed to start from a high copy number, so for real samples they had to incorporate an initial amplification step by RPA (130,131) (Figure 8a).

To use CRISPR in this field, since the appearance of SHERLOCK, new CRISPR-associated endonucleases were discovered, such as Cas12a, which also has collateral activity, but its target is DNA (132). Furthermore, improvements were made allowing multiplexing, real-time and quantitative detection of SNV (133). In fact, SHERLOCK recently was combined with nanodroplets to detect the Zika virus and SARS-CoV-2 coronavirus simultaneously in multiple samples (134) (Figure 8b). However, it is tough to obtain these nucleases, as they are not commercially available. Therefore, they must be expressed and purified, which is not available to all laboratories. In addition, the need for a preliminary amplification step to achieve the required sensitivities making this technique not yet the most suitable for POC scenarios (135).



**Figure 8.** Application of CRISPR/Cas systems for sensitive SNV detection. a) Singleplex strategy using Cas13a adapted from (130). b) Multiplex detection using an enzymatic cocktail adapted from (133) and (134).

### 1.3.2.2 Enzymatic isothermal amplification methods

Although PCR is considered the reference method for SNV detection, current research focuses on finding alternative methods that overcome its limitations in portability, speed, and cost to reach those regions with limited resources. In this case, isothermal amplification-based techniques have been a successful tool for POC applications by avoiding the need for a thermal cycler (136).

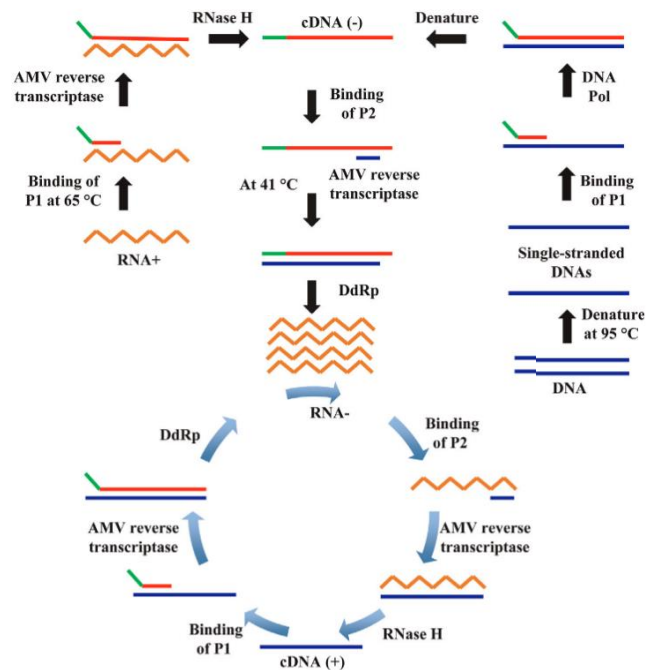
Numerous isothermal amplification techniques have been described to date. The mechanism used in these reactions mimics the amplification processes used *in vivo*. The need for high temperatures is circumvented by proteins that induce denaturation of the double helix, and selective amplification is achieved at constant temperature (37-65 °C), simplifying the design and operation of the system (137). In this thesis, the characteristics of the most used for SNV detection are described (Table 11).

#### ■ Nucleic acid sequence-based amplification (NASBA)

In 1991 Compton et al. described NASBA which copies the retroviral replication strategy of RNA molecules. It requires a pair of primers (P1 and P2) and three enzymes: avian myeloblastosis virus reverse transcriptase (AMV-RT), RNase H, and T7 DNA dependent RNA polymerase (DdRp) (138). It is similar to Transcription mediated amplification (TMA), but TMA uses a reverse transcriptase with intrinsic RNase H activity (139).

When the target is RNA, it needs a first heating step at 65 °C for the annealing of P1, which contains the T7 promoter (DdRp binding site). P1 will be extended by AMV-RT, giving rise to a cDNA complementary to the starting RNA. Subsequently, P2 binds, and it will be extended by AMV-RT, resulting in a DNA duplex. The temperature is lowered to 41 °C so that the DdRp generates multiple copies of sRNA that will serve as a template for subsequent cycles (Figure 9) (138).

NASBA has been mainly applied to identify RNA viruses (140). For instance, it has recently been used for colorimetric detection of noroviruses (141). Although NASBA is not one of the most explored techniques to discriminate SNV, some examples can be found in the literature. For instance, employing NASBA with a fluorescent molecular beacon to detect the mutation in Factor V Leiden (G1691A) (142). Alternatively, it is utilized in microarray technology to detect SNV in genes involved with breast cancer (*CASP8* (D302H) and *SOD2* (V16A)) (143), or for multiplex recognition of different mRNA y miRNA (144). Also, colorimetric split G-quadruplex DNAzyme assay to differentiate between Shimen and HCLV strains of Classical Swine Fever Virus (145). However, in order to perform this technique from DNA samples, it is also necessary to add a previous denaturalization step and a DNA-polymerase.



**Figure 9.** Mechanism of NASBA amplification. Original figure from (136).

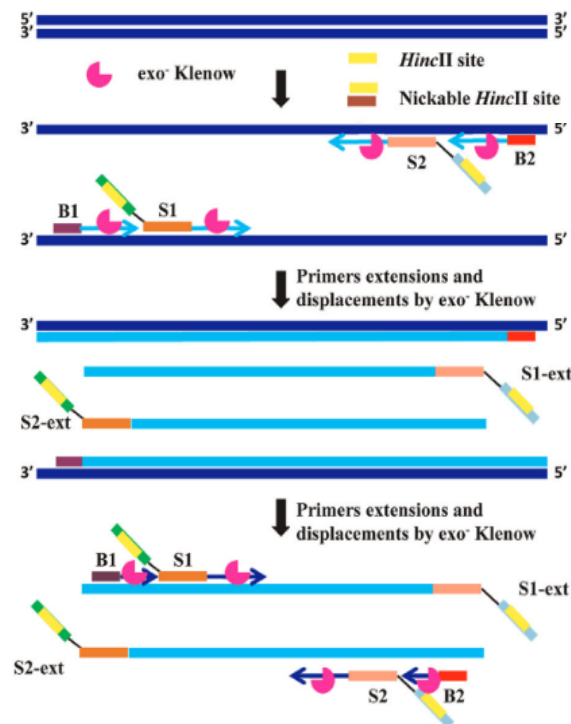
### ■ Strand Displacement Amplification (SDA)

In 1992, SDA was reported by Walter et al. This technique requires: a pair of promoter primers harboring a restriction site that hybridizes to the target sequence, a pair of displacer primers, a DNA polymerase with strand displacement activity (exo Klenow), and a restriction endonuclease.

An initial denaturation step at 95 °C is required. Subsequently, the temperature is lowered to 37 °C so that the forward promoter primer binds to the ssDNA template to extend it. Then, the forward displacer primer binds, and this new strand is synthesized, displacing the previous strand giving rise to the original dsDNA. This exact process occurs with reverse primers so that the amplicon is flanked by the restriction sites. These sites are then recognized and cleaved by the endonuclease leaving the 3' end free. The resulting overhang serves as a template for the following cycles (Figure 10) (146).

SDA has been widely implemented in the diagnosis of human pathogens (147). This technology has been explored more extensively than NASBA for SNV testing. Some authors coupled SDA with a ligation step and different ways of detection, such as DNAzyme-based chemiluminescence to discriminate of *CYP2C19\*2* (148), or magnetic beads and chemiluminescence biosensor to detect SNV in *KRAS* (codon 12-13) (149). Others employed LNA with SDA and an electrochemical approach to identify *p53* mutations in complex samples (150). SDA has also been combined with palindromic molecular beacon via fluorescence for hotspots in *KRAS* (codon 12-13) (151). Moreover, recently, a restriction enzyme-SDA detected p.H1047R in *PIK3CA* by an electrochemical

method using DNA nanostructures (152). Nevertheless, it should be noted that SDA is not very sensitive to the DNA background, and therefore undesirable products may be generated, leading to false-positive results.



**Figure 10.** Mechanism of SDA amplification. Figure adapted from (136).

## ■ Rolling Circle Amplification (RCA)

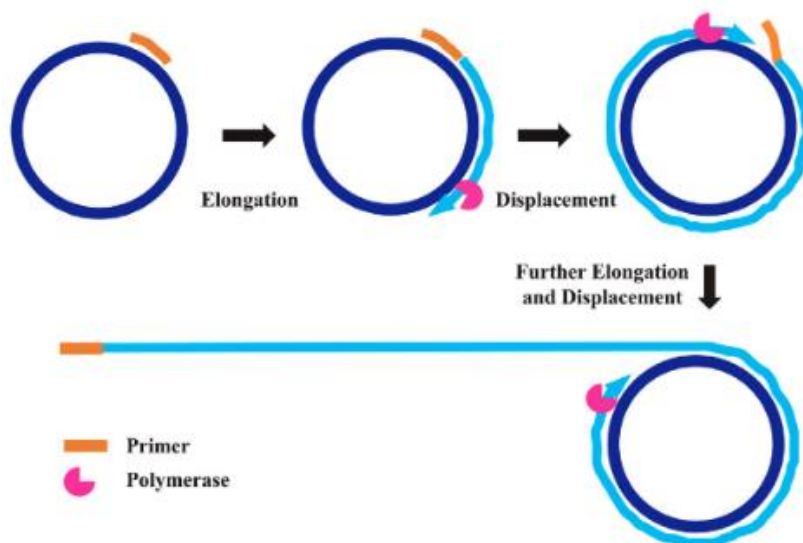
In 1995 Fire and Xu discovered RCA that mimics *in vivo* replication mechanism of some plasmids and viral nucleic acids. Linear amplification is based on a primer and an enzyme with strand displacement activity, such as DNA polymerase  $\phi 29$ , with high processivity and continuously synthesizes a cDNA strand at 30 °C. Thus, the final product is a long strand of ssDNA containing tandem repeats of the cDNA sequence (153).

A second complementary primer can be added to the generated downstream DNA sequences to achieve exponential amplification, allowing strand displacement to occur, serving as a template for subsequent cycles. As a result, hyperbranched anchored from the original cDNA is produced. Even cascade amplification systems have been developed that perform multiple rounds of secondary amplifications at 60 °C. Enzymes can digest the final products into shorter fragments (Figure 11) (154).

RCA has been used since 2001 to genotype SNV combined with ligation and subsequent fluorescence readout (155,156). It has also been used in conjunction with other mechanisms, such as gold nanoparticles to detect mutations in the beta-thalassemia gene (157), or with template enhanced hybridization process and an

electrochemical platform for single-base mismatched oligonucleotides (158). Lately, RCA has combined with time-controlled FRET to distinguish V600E in *BRAF* gene (159), and for the same SNV dual padlock-GLCR with hyperbranched RCA have integrated, allowing real-time fluorescent biosensing (160).

Nevertheless, this technique needs to start from a single cDNA strand. Therefore, for its application to dsDNA, RCA requires a previous step that circularizes the dsDNA by using padlock probes and an additional ligase, which adds complexity and makes its integration in POC more difficult (161).



**Figure 11.** Mechanism of RCA amplification. Original figure from (136).

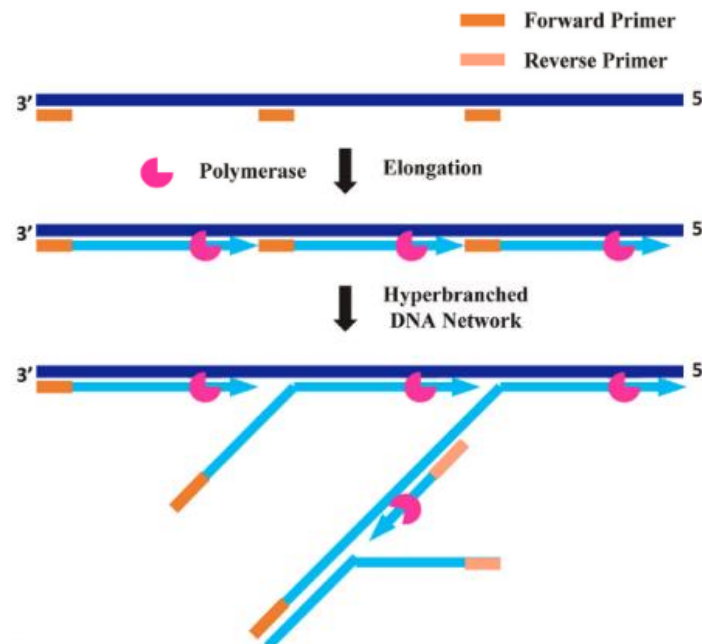
### ■ Multiple Displacement Amplification (MDA)

While the previously described methods amplify specific sequences with defined primers (136), the first example of an isothermal reaction using random primers was **MDA**, described by Deant et al. in 2001. MDA uses cDNA or gDNA as a template and short six base-pairs primers (hexamers) that randomly hybridize to the genome, allowing the DNA-polymerase ( $\phi$ 29) to bind. As a result, simultaneous synthesis of DNA at the multiple hexamer binding sites occurs (162).

Once the polymerase reaches the upstream hexamer, it synthesizes new chains through its continuous strand displacement activity. Consequently, single-chain branches are formed to which new hexamers could attach, resulting in an exponential amplification at 30 °C. Neither denaturation/renaturation cycles nor prior knowledge of the sequence is required (Figure 12) (163).

MDA amplification generates a large number of templates with high-quality enabling SNV genotyping for applications such as whole genome amplification (164), comparative genomic hybridization (165), or short tandem repeats (163). For example,

forty-five SNV in different genomic regions were genotyped using fluorescent minisequencing in a microarray format (166). Although in these cases, balanced amplification of both alleles of each SNV is needed. However, since MDA amplifies all the DNA present in the sample, only 30% of the final product is specific to each SNV (167).



**Figure 12.** Mechanism of MDA amplification. Figure adapted from (136).

### ■ Loop-Mediated Isothermal Amplification (LAMP)

In 2000 Notomi et al. reported a new isothermal approach called LAMP. It employs two pairs of primers, a forward (F1c+F2) and backward (B2+B1c) inner primers plus a forward (F3) and backward (B3) outer primers, which recognize six distinct sequences in the dsDNA template, and a polymerase with strand displacement activity, such as Bst polymerase (168).

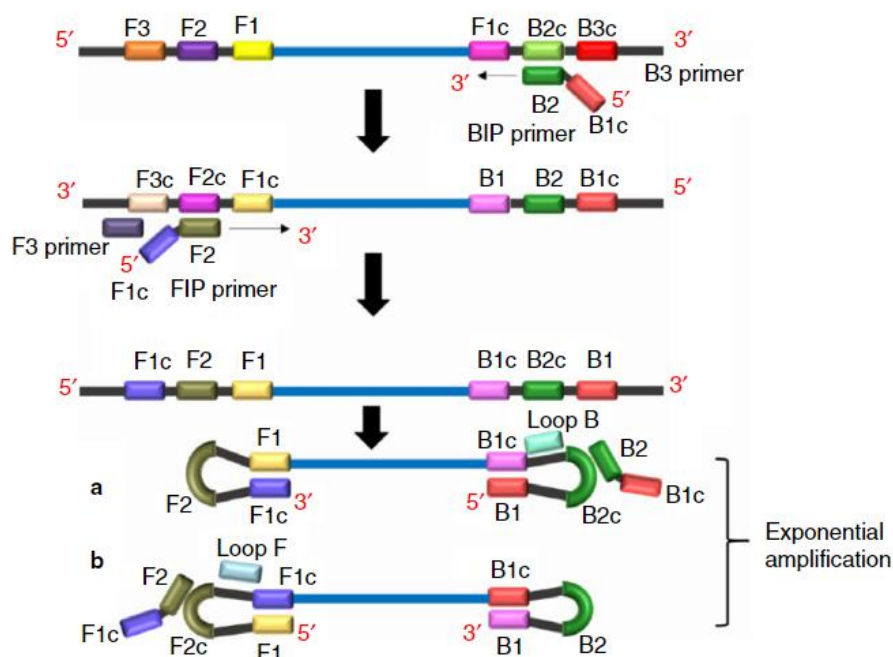
One of the inner primers recognizes its target sequence, which is extended by the polymerase separating the duplex. Next, the outer primer recognizes an upstream target region, and as it is extended, it displaces the first formed product. The end of this product forms a self-hybridizing loop structure due to the inclusion of a reverse complementary sequence in the inner primer (F1c). This annealing and displacement cycle repeats on the opposite end of the target sequence, and the resulting product is a short dumbbell structure that forms a seed for exponential LAMP amplification at 60-65 °C, as it contains multiple sites to initiate the synthesis (Figure 13) (169,170).

The products grow and form long concatamers, each with more sites for initiation, leading to a rapid accumulation of dsDNA with different lengths and multiple loops. As a

result of the reaction, high amounts of pyrophosphate ions are produced, which on reacting with the magnesium in the medium, generate a white precipitate, facilitating visual detection. In addition, the concentration of these ions is correlated with the initial DNA concentration, which allows quantification of the samples without any additional treatment (171).

LAMP is a technique with great POC potential to detect point mutations that have been extensively studied by numerous groups combining it with different strategies. Some examples from recent years will be briefly presented. Coupled with microarrays and electrochemical detection of six SNV simultaneously associated with rheumatoid arthritis in *NAT2*, *MTHFR*, and *SAA1* genes (172). Allele-specific LAMP in conjunction with ASO hybridization to identify SNV in *GRIK4* using a colorimetric approach (173). Probe-enhanced LAMP using fluorescent real-time and visual way for the hotspots mutations in *CYP2C19\*2* and *MDR1* (174). A centrifugal compact disc microfluidic platform employed PNA and LNA blocking probes to colorimetric detection of SNV in *calreticulin* gene (175). LAMP with a molecular beacon to the fluorescent recognition of the V600E in *BRAF* from cDNA (176). Using an RNase-activatable primer for fluorescent detection of G12C mutation in *KRAS* (177). Integrating LAMP with CRISPR technology to fluorescent discrimination of five *Neisseria meningitidis* serotypes virus (178).

Nonetheless, the main limitation of LAMP is false-positive amplification due to cross-contamination of the product. Although the number of primers can increase (6-8) to solve it, this complicates the design and optimization of the reaction (169).



**Figure 13.** Mechanism of LAMP amplification. Figure adapted from (170).

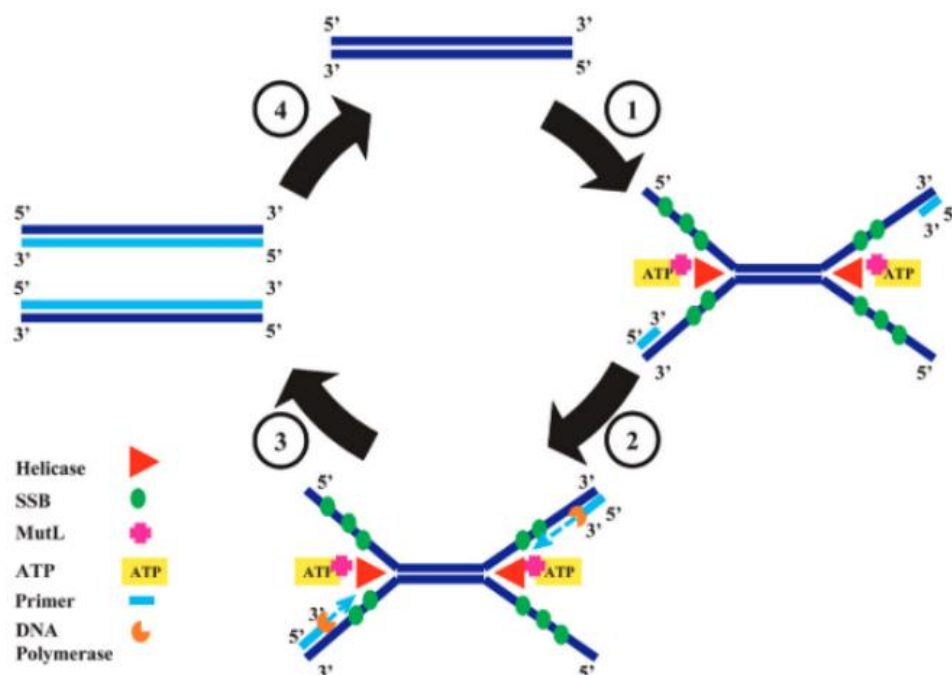


## ■ Helicase Dependent Amplification (HDA)

In 2004, Vicent et al. described HDA based on a design similar to DNA replication in vivo (179). The HDA reaction consists of three steps. First, separation of the duplex by the enzyme helicase UvrD and its cofactors, methyl-directed mismatch repair protein (MutL), and ATP. Second, hybridization of the primers by stabilizing the separated strands with single strand binding proteins (SSB). Last, elongation to form new strands. Two dsDNA are then formed, and the cycle is repeated, resulting in exponential amplification at 60-65 °C (180).

Few applications of HDA have been found in SNV genotyping. For instance, HDA has combined with allele-specific probes and real-time fluorescent to detect the most critical SNV in *VKORC1* and *CYP2C9* (181). Also, HDA with gold nanoparticles and microarray identified hotspots in codons 12-13 of *KRAS* (182). As well as using it with primers containing mismatches at 3' end and gold nanoparticles utilizing real-time fluorescence to *CYP2C9\*2* genotyping (Figure 14) (183).

Although for quantities less than 100 copies of initial DNA, its efficiency and reaction speed decrease considerably due to the low speed and processivity of the helicase and its poor coordination with DNA polymerase (180).

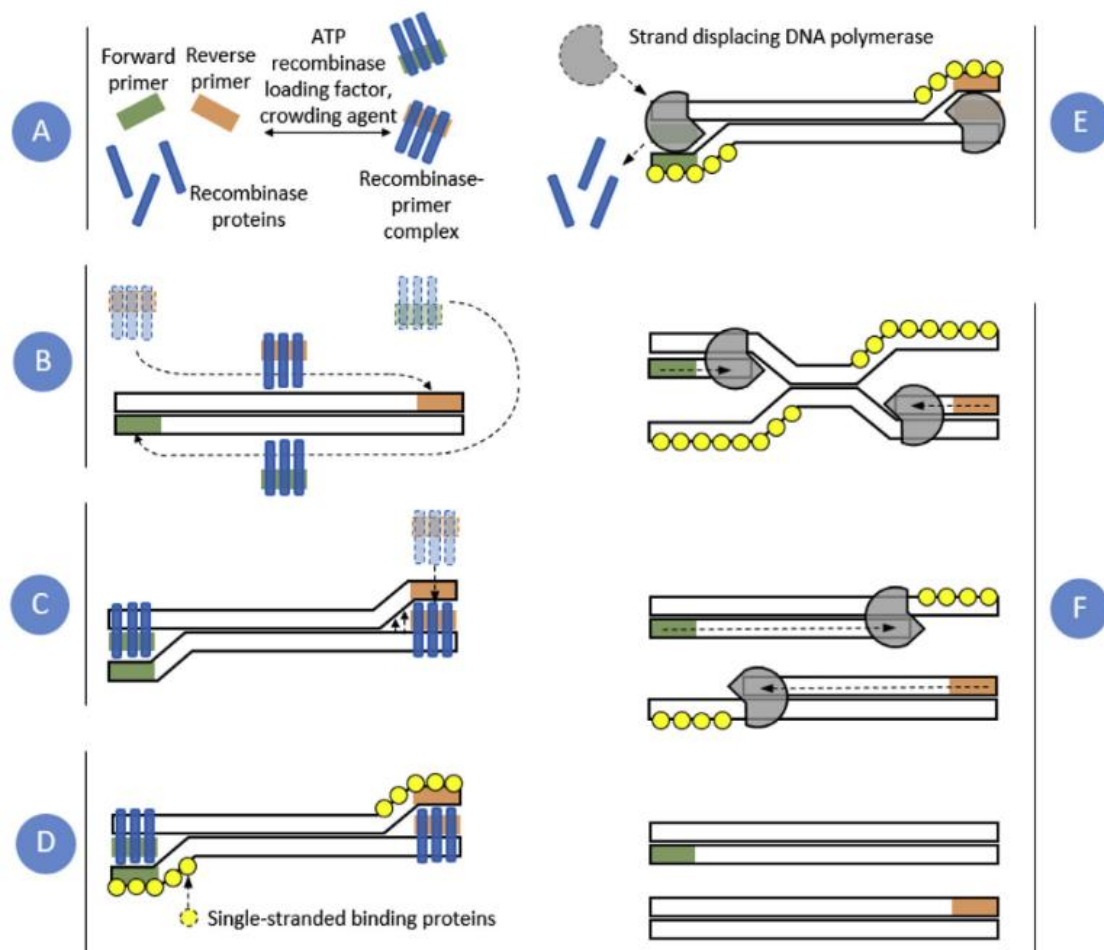


**Figure 14.** Mechanism of HDA amplification. Figure adapted from (136).

## ■ Recombinase Polymerase Amplification (RPA)

In 2006, Piepenburg et al. discovered a method called RPA that, like HDA, does not require an initial heating step of the dsDNA to separate it, since it utilizes an enzyme, in this case, a recombinase (RecA). The mechanism mimics the homologous recombination that occurs in cells to repair dsDNA breaks (184).

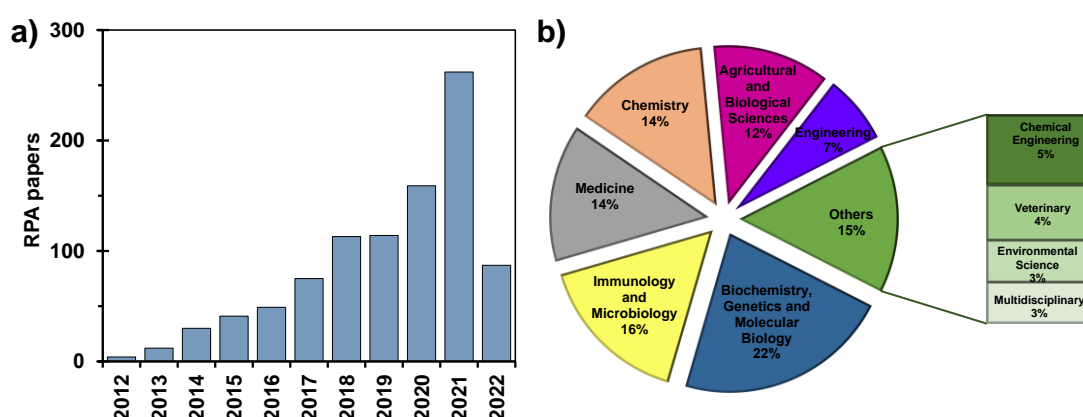
Initially, in the presence of ATP, RecA binds to the primers and a crosslinking agent (high molecular polyethylene glycol) to form nucleoprotein complexes that scan the dsDNA until the homologous sequence is found. Upon joining, recombination occurs, and the other strand of DNA is displaced. The structures obtained are maintained by SSB proteins that bind to the template strand and stabilize it to avoid primer displacement due to strand migration. DNA polymerase with strand displacement activity, starting from the free 3' ends of the primers, initiates the extension and replication of the entire template strand. The cyclic repetition of this whole process leads to exponential amplification of the target sequence (Figure 15) (185). Although since 2010, TwistDx (UK) has patented a formulation for using this technique, in which the enzyme mixture is freeze-dried, "substantially different", and is only available in the TwistAmp kit (186).



**Figure 15.** Mechanism of RPA amplification. Original figure from (185).

Two primers (forward and reverse) are needed per target. Its design has been the subject of study. Primers used for PCR (18-23 bp) are suitable for RPA, although longer primers (> 45 bp) that may form secondary structures that interfere with the recognition process should be avoided. In addition, it has been shown that a reasonable reaction yield is achieved when amplifying short sequences of 80-200 bp. (187).

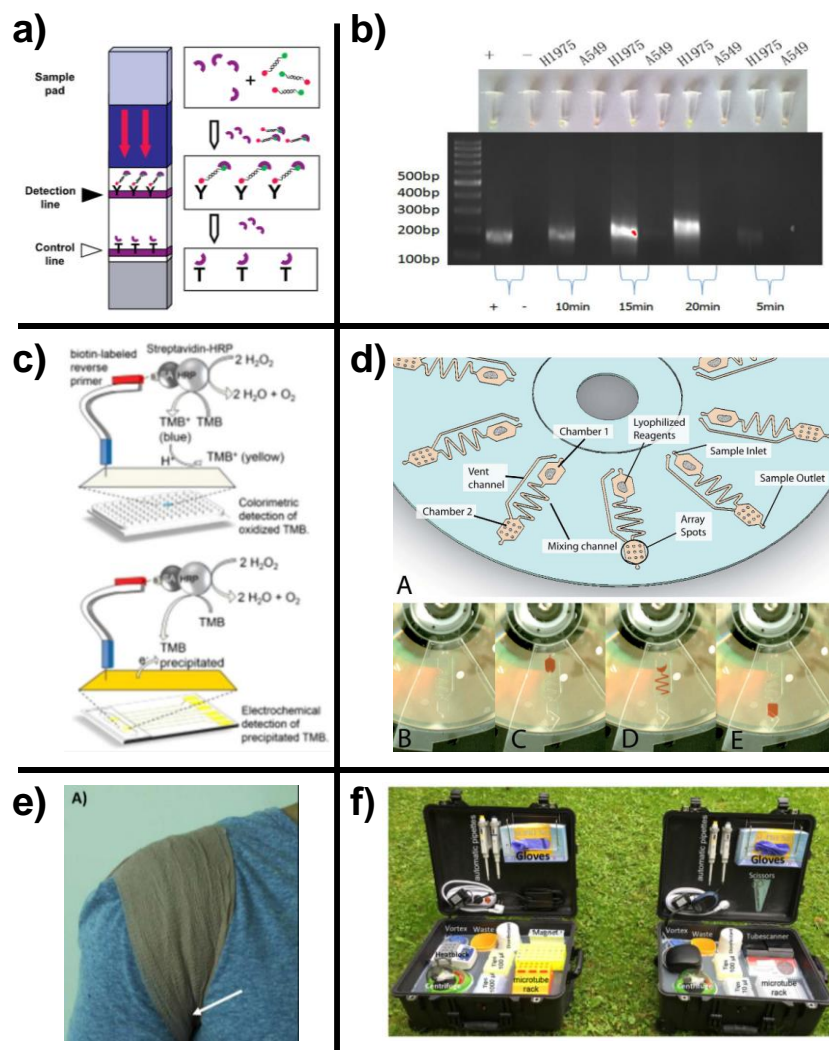
Significantly since 2015, research using RPA has grown exponentially, and exciting studies with practical applications to detect animal, food, and environmental pathogens have been developed (Figure 16) (188–190). In fact, TwistDx has also developed two kits to detect *Salmonella enterica* in 10 and 20 min, TwistGlow® y TwistFlow®, respectively, based on the amplification of the *invA* gene (186).



**Figure 16.** RPA metric studies. a) Time evolution of the number of RPA publications (2012-2022). b) RPA research areas. Source: Web of Science, 2022.

Incubators, heating plates, laboratory ovens, and even regions of the human body, such as the armpit, wrist, or abdomen, have been used to control the reaction temperature (188), making RPA compatible with POC systems (Figure 17). Despite this, its main handicap is undesirably early unspecific amplification at room temperature. This problem can be solved by controlling the reaction conditions (preparation of the mix at 4°C) and adding magnesium acetate as the last reagent since it is a cofactor needed by the recombinase to start working (191).

RPA endpoint detection is usually carried out by gel electrophoresis (192), electrochemistry (193), and Surface-enhanced Raman spectroscopy (194). Although one of the most popular is the lateral strip flow assays, allowing for a cost-effective and straightforward visual detection. However, the generated results are qualitative (195). On the other hand, real-time and quantitative fluorescence measurement is possible using intercalating binding dye or an endonuclease and unique probes (196). Meanwhile, these approaches are not compatible with actual multiplexing. Microarray-based systems have been used in which products that RPA had amplified were hybridized to overcome this limitation (Figure 17) (197).



**Figure 17.** Different detection approaches using RPA. a) Lateral flow test (184). b) Gel electrophoresis (192). c) Electrochemical and colorimetric assay (198). d) Microfluidic assay on DVD disc (197). e) Reactions taking place in the axilla of a person using a bandage from (188). f) Portable laboratory (185). Figures taken from the cited publications.

In particular, for SNV detection, there are few research papers in the literature because RPA has the disadvantage that it can recognize one or more mismatches depending on their location (185). However, it was possible to detect 30% of mutant alleles in *EGFR* by combining RPA with specific alleles, PNA, and fluorescence analysis in clinical FFPE samples (192). However, this strategy is not genuinely isothermal since it requires a preheating and cooling step at 99 °C and 66 °C, respectively. This pre-heating step was necessary in order to hybridize gDNA with PNA. Electrochemical detection of 50 copies of the SNV in the *TMPRSS2-ERG* gene from ctDNA using a nanofluidic platform has also been reported (199). Our group developed an AS-RPA both in the solid phase and in solution achieving a colorimetric detection of 5-10% of SNV involved in pharmacogenetics (rs4680,rs1799971) from buccal swabs (200). In addition,

incorporation of a blocking agent in the RPA and subsequent hybridization of products led to discrimination of 5% mutant in the *PIK3CA* from FFPE (201). Recently, RPA in solution was combined with LAMP for the fluorescent detection of 100 copies of V600E in *BRAF*, though using synthetic samples (202).

On the other hand, RPA has been widely used in recent years as a previous step to detection assay to achieve the required sensitivity and reduce the complication of gDNA favoring the recognition of the target. Current applications of this approach are the combination of RPA with CRISPR/Cas 9 for multiplexed, portable, and ultra-sensitive SNV detection as commented previously (130,133).

RPA has clear advantages over other isothermal techniques using enzymes (Table 11), such as fewer primers, reduced time, low reaction temperature, and simplicity. Indeed, RPA is characterized by high efficiency, specificity, robustness, and affordability ( $\approx 3 \text{ €}/\text{reaction}$ ). Furthermore, RPA does not require complex equipment or qualified personnel, making it suitable to be used in almost any environment (198). All these reasons make this technique an ideal candidate for POC detection of SNV in clinical practice. Therefore, in combination with SNV discrimination strategies, this isothermal approach to amplify the copy number of the targets has been selected in this doctoral thesis.

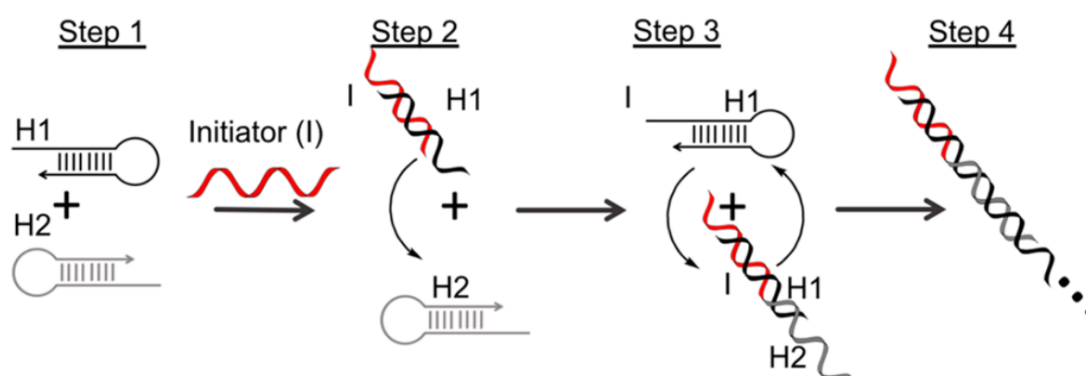
**Table 11.** Comparison between characteristics of enzymatic isothermal amplification techniques.

	<b>NASBA</b>	<b>SDA</b>	<b>RCA</b>	<b>MDA</b>	<b>LAMP</b>	<b>HDA</b>	<b>RPA</b>
<b>Diana</b>	RNA DNA	DNA	RNA DNA	DNA	DNA	DNA	DNA
<b>Denaturation stage</b>	Yes	Yes	No	No	No	No	No
<b>Primers</b>	2	4	1	Random hexamer primers	4-6	2	2
<b>Temperature (°C)</b>	41	30-55	30-65	30	60-65	60-65	37-42
<b>Time (min)</b>	60-180	60-120	60-90	100-300	45-65	30-120	20-40
<b>LOD (copies)</b>	1	10	10	1-10	6	1	1
<b>Yield amplification</b>	$10^6-10^9$	$10^7$	$10^9$	$10^5$	$10^9$	$10^7$	$10^4$ (in 10 min)
<b>Multiplexing</b>	Yes	No	Yes	No	Yes	Yes	Yes
<b>Enzymes</b>	RT and RNApol	DNApol and RE	DNApol and Ligase	DNApol	DNApol	DNApol and Helicase	DNApol and Rec

**RNApol:** RNA polymerase; **RT:** Reverse Transcriptase; **DNApol:** DNA polymerase; **RE:** Restriction Endonuclease; **Rec:** Recombinase

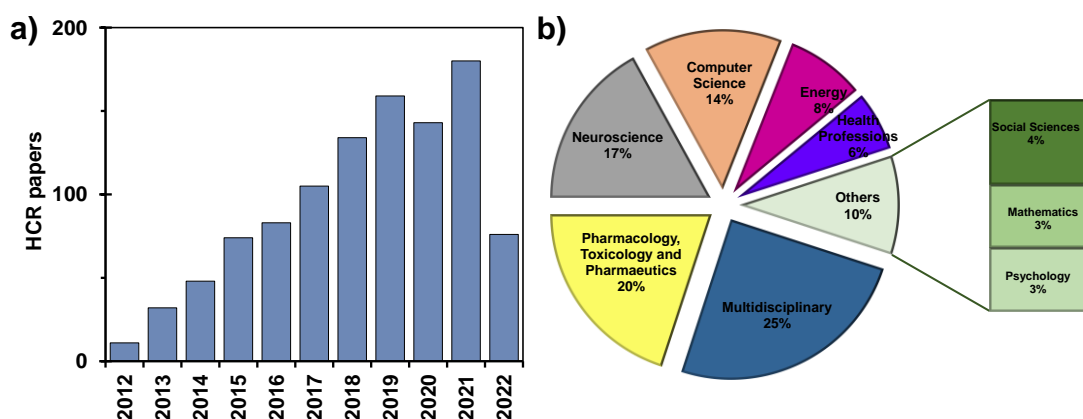
## 1.3.2.3 Non-enzymatic isothermal amplification methods

In 2004, Dirks and Pierce first described **Hybridization Chain Reaction (HCR)**, an isothermal amplification technique that does not require enzymes. HCR is based on the hybridization ability of two oligonucleotides (H1 and H2) hairpin-shaped in solution but have recognition sites to produce cross-hybridization between them at room temperature (25-37 °C). In the presence of the target ssDNA, the hairpins open, allowing the recognition of more annealing sites that hybridize with each other, producing a cascade of strand-displacement reactions (203). As a result, long chains of nicked dsDNA are formed, resulting in DNA superstructures in less than an hour (Figure 18) (204).



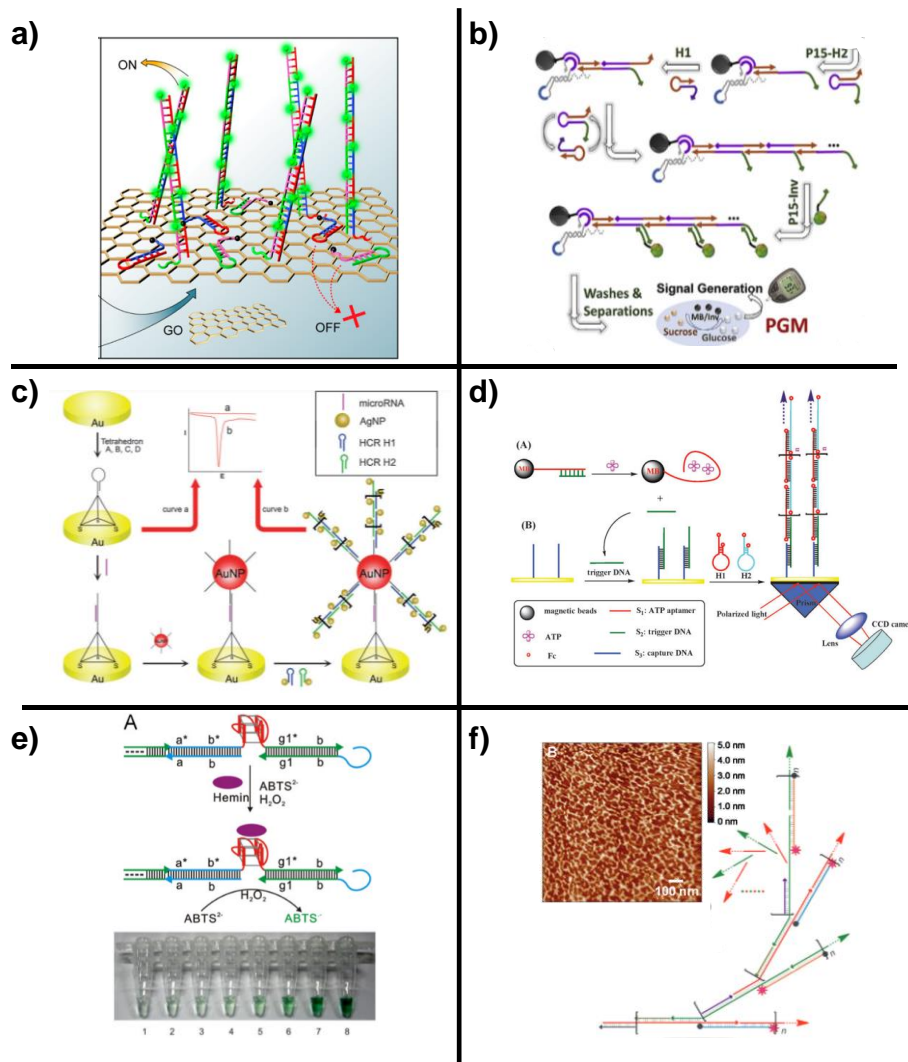
**Figure 18.** Mechanism of HCR amplification. Figure original from (204).

An oligo-initiator called link can also be incorporated to trigger the reaction in the presence of the target DNA so that the set of hairpins is universal for any target, allowing multiplexing. With the rise of programmable DNA nanostructures in recent years, more complicated monomer sets, such as branched HCR and even dendritic systems, were designed, achieving exponential growth mechanisms (205). However, these systems increase the complexity of hairpin set design and increase the leakage possibilities or spontaneous reaction initiation. Indeed, they have higher steric hindrance and lower flexibility (206). Due to the advantages of HCR for its isothermal and enzyme-free nature, high amplification efficiency, few-step protocols, and low cost, it is an ideal POC method for applications in molecular detection, bioimaging, and biomedicine. In fact, in the last years HCR has been used for the sensitive detection of nucleic acids, proteins, enzymatic activities, tumor cells, and small molecules (Figure 19) (204).



**Figure 19.** HCR metric studies a) Time evolution of the number of HCR publications (2012-2022). b) HCR research areas. Source: Web of Science, 2022.

Moreover, HCR amplification alone often cannot provide sufficient sensitivity for trace amounts of nucleic acid. Therefore, it has been combined with different strategies (Table 12), and due to its versatile nature, it can be used both as a target and signal amplifier (207). Furthermore, flexible developer reagents can be used, and HCR has been detected using different signal transduction approaches for nucleic acid detection, as shown in the Figure 20 and Table 12. Although fluorescence has important advantages, for certain uses it is not the most convenient technique (208). The same situation happens with electrochemical techniques (as it will be discussed in section 1.4.1) (209). Instead, colorimetric assays can be an easy operation alternative, especially useful for naked-eye detection and low-cost POC solutions (210).



**Figure 20.** Different detection approaches using HCR. a) Fluorescent homogeneous assays (211). b) Personal glucometer reader (212). c) Electrochemical assay using silver and gold nanoparticles (213). d) Surface plasmon resonance using CCD equipment (214). e) Colorimetric assay using G-quadruplex (215). f) AFM images of hyperbranched products (216). Figures taken from the cited publications.

Even though HCR has been applied to SNV detection, no studies have been reported that can identify them, which is crucial information for diagnosis and prognosis. On the other hand, HCR-based bioanalytical methods have not been incorporated into clinical practice yet. Due to only short strands (microRNAs, short gene sequences, and certain ctDNA) can be amplified by HCR, which seriously limits its application for gDNA (211). Indeed, none of the reported studies identifies a single nucleotide change in gDNA. Overcoming this limitation is an objective of this thesis. Hence, RPA was combined, for the first time, with HCR from clinical samples to achieve the required sensitivity and reduce the structural complexity of gDNA by generating small-sized products, which allow access to the probes and the biorecognition process (130,217), as it will be explained in chapter 3.



**Table 12.** Examples of different HCR sensing approaches classified according to the transduction principle.

Platform	Haipins modifications	Signal developer	HCR coupled to	Target	LOD*	SNP detection	Application to biological samples	Reference
<b>Fluorescence</b>								
Solution	Both with pyrene excimer at the 3'- and 5'-end.	Hairpins	NO	Target DNA	256 fM	NO	Cell media and human serum	(208)
Solution	H2 with pyrene at the end.	Cationic cyclodextrin-tethered polymers	NO	Target DNA	0.1 nM	NO	NO	(218)
Solution	H1 contained a polycytosine loop and H2 a poly-guanine sticky end.	Hairpins	Highly fluorescent Ag nanocluster	let-7a miRNA	0.78 nM	YES	NO	(219)
Solution	Both with fluorescent ATMND at their ends	Hairpins	PCR	Target DNA and genotype 107-mer PCR DNA product of the <i>KRAS</i> gene	0.2 nM	YES	NO	(220)
Solution	Non-modification.	SYBR Green I	Click-chemical ligation-assisted hybridization and MBs separation processes	miR-141 family and other miR-200 members	5 fmol	YES	Fetal bovine serum	(221)
GO surface	Both with FAM fluorophore at the sticky terminal.	Hairpins	GO platform	let-7a miRNA	1 pM- 5 nM	YES	NO	(211)
SA-coated MB surface	Both with biotin at their ends.	HRP catalyzes HPA into bi-p,p'-4-hydroxyphenylacetic	MB	Target DNA	0.8 fM	YES	NO	(222)
Chip surface	4 hairpins with non-modification.	SYBR Green I	Sandwich capture hybridization assay before 2D branched HCR	Malaria 18sRNA and GAPDH mRNA	1 pM	NO	Blood and cell lysate	(223)
Gold substrate	Both with azide at their ends.	NIR fluorescent dyes (DBCO-Conjugated Cy5)	PE-HCR	trace miRNAs-21	1 fM-1 pM	YES	Cell cultures of HepG-2, MCF-7, HeLa, 3T3	(224)
Solid phased	Both with 2'OMe-RNA probes at their ends.	Hairpins	Norther blot	miRNA	25 aM	NO	NO	(225)

(continue)

(continued)

**FRET**

Solution	Hairpins with FAM and Dabcyl on opposite ends.	Hairpins	Target-recycling features of CHA	miR-141	0.3 fM	NO	NO	(226)
Solution	2-aminopurine in the toe-hold region of H1.	Hairpins	Branched	Target genes of HIV and Chlamydia.	50 fM	NO	NO	(227)
Solution	Both with a fluorophore FAMN in one end and a quencher ABkFQ in the other end.	Hairpins	Dendrimers	Target DNA	0.001%	NO	NO	(206)
Solution	Both with a quencher (BHQ-2) and a fluorophore (ROX) at their ends.	Hairpins	HB-HCR	Target DNA	10 <sup>-13</sup> M	NO	NO	(216)
Solution	Four hairpins: H1 and H3 with Tb (FRET donor), and H2 and H4 with Cy5.5 (FRET acceptor) at their ends.	Hairpins	TG-FRET	miRNA-20a and miRNA -21 and their ssDNA analogues	0.88-1.7 pM	NO	Bovine serum albumin	(228)

**Electrochemical**

Au electrode	Both with biotin at their ends.	Avidin-modified HRP and H <sub>2</sub> O <sub>2</sub>	NO	Target DNA and microRNA	100 aM for DNA and 10 aM for RNA	NO	NO	(229)
negatively charged ITO electrode	Both with split G-quadruplex sequences at their ends.	Methylene blue	DNAzyme	let-7a miRNA.	1 pM	YES	NO	(230)
Au electrode	Both with split G-quadruplex sequences at their ends.	G quadruplex- hemin reducen TMB in presence of H <sub>2</sub> O <sub>2</sub>	EXPAR and DNAzymes nanowires	Avian influenza A(H7N9) virus DNA	9.4 fM	YES	HepG2 cell lysates	(231)
Au electrode	Non modification.	AgNPs	Tetrahedral modified with AuNPs	DNA with miRNA-17	2 aM	YES	Cell lysates of HUVEC, HeLa, HK-2 and MCF-7	(213)
Au electrode	Non modification	[Ru(NH <sub>3</sub> ) <sub>6</sub> ] <sup>3+</sup>	Sandwiched hybridization	BRCA1 gene	1 aM	YES	50% human serum	(232)

(continue)

(continued)

Negatively charged ITO electrode	Non modification	Fc-PNA probes	dendrimer DNA/PNA	of	Target DNA	100 fM	YES	NO	(233)
Solution	The first-H1 attached to MBs. H2 had 15-base extension at the 5'end which could hybridize with its complementary DNA labelled with invertase	Re-suspended MBs with invertase hydrolyze sucrose into glucose which can be detected by PGM.	LAMP		Gene segment of contagious norovirus GII	30 copies	NO	Fecal samples	(212)
AuNP biocathode	Non modified	[Ru(NH <sub>3</sub> ) <sub>6</sub> ] <sup>3+</sup>	SDR		p53 gene fragment	20 aM	YES	Cell lysate sample	(234).

**Electrochemiluminescence**

Au electrode	Both with biotin at their ends.	SA-AuNPs catalyze luminol	Amplification of modified-AuNPs	of	DNA sequence derived from the HIV-1	5.0 fM	YES	Human serum	(235)
Au electrode	Non-modified.	[Ru(NH <sub>3</sub> ) <sub>6</sub> ] <sup>3+</sup>	ECL technique		DNA sequence specific to Escherichia coli.	15 fM	YES	NO	(236)
Solution	Both with biotin at their 3'end.	HRP-capped MBs which catalyze the PIP enhanced oxidation of luminol	Branched RCA		DNA methylation	0.52 CFU·mL <sup>-1</sup>	NO	NO	(237)

**Colorimetric**

In solution	Non modified	AuNPs	SPR		miR-10b, miR-21 and miR-141	20 fM	NO	Total RNA extracted from 4T1 breast cancer cell line	(238)
In solution	Sticky end of H1 and H2 is attached to AuNPs	AuNPs	SPR		16S rRNA gene of the bacterial V. cholerae	20 nM	YES	NO	(239).
In solution	Both with ssDNA sticky ends	AuNPs	SPR		Target DNA	50-100 pM	YES	NO	(240)
Solution	H1 with biotin	Avidin-GOD mediated plasmonic triangular Ag nanoprism etching	SPR		p53gene	6.0 fM	YES	NO	(241)

(continue)

(continued)

Plastic sheet	H1 with G-quadruplex in the loop and in 1/3 of the stem.	DNAzyme by catalyzing the oxidation of ABTS after binding to hemin with H <sub>2</sub> O <sub>2</sub> .	G-quadruplex, PCR and Lambda exonuclease.	Target DNA and sRNA.	7.5 nM	YES	NO	(215)
Au chip	Both with Fc at both ends.	Hairpins	SPR	Target DNA and ATP	0.3 fM	NO	NO	(214)
Solution	H1 with GOx and H2 with HRP.	GOx/HRP for the oxidation of ABTS <sup>2-</sup> by H <sub>2</sub> O <sub>2</sub>	With a multiple enzymatic cascade reaction.	Target DNA	5.2 fM	YES	NO	(242)
Solution	Both with structures include three-fourths and one-fourth of the HRP-mimicking DNAzyme	Hemin/G-quadruplex (HRP) nanowires catalyze the oxidation of ABTS <sup>2-</sup> by H <sub>2</sub> O <sub>2</sub> to ABTS <sup>-</sup> and H <sub>2</sub> O	Hemin/G-quadruplex HRP-mimicking DNAzyme.	BRCA1 oncogene	100 fM	YES	NO	(243)
Solution	Both with sticky ends, but non modified	Split G-quadruplex/hemin with high peroxidase-like activity, catalyzed the oxidation of ABTS <sup>2-</sup> by H <sub>2</sub> O <sub>2</sub> to ABTS <sup>-</sup> and H <sub>2</sub> O.	DNAzymes activity through triplex DNA formation.	ctDNA	0.1 pM	YES	Dilute blood samples from healthy patients with ctDNA	(244)
Solution	Non modification	Peroxidase mimics of Fe <sub>3</sub> O <sub>4</sub> NPs and AuNPs	Fe <sub>3</sub> O <sub>4</sub> nanoparticles	Yersinia pestis-relevant DNA sequence	100 pM	YES	NO	(245)
Solution	G-quadruplex structure in the stem of H1	Using the peroxidase activity induced by the binding between G-quadruplex and hemin	DNA zyme	miR-21, miR-125b, KRAS-Q61K and BRAF-V600E	5.67 nM	YES	Diluted sample FBS	(246)

\***LOD**: Limit of detection expressed as percentage (%), concentration (CFU·mL<sup>-1</sup>) or amount (aM, fM, pM, nM), according to the results of each work.

**Target DNA**: secuencia sintética corta de oligonucleotidos de single cadena; **MBs**: Magnetic beads; **GO**: graphene oxide; **Ag**: silver; **Au**: gold; **SA**: Streptavidin; **ATMND**: amine molecule, 5,6,7-trimethyl-1,8-naphthyridin-2-ylamine; **HPA**: hydroxyphenylacetic acid; **NIR**: near-infrared; **PE-HCR**: plasmon-enhanced hybridization chain reaction; **TG-FRET**: time-gated FRET; **CHA**: catalytic hairpin assembly; **HB-HCR**: Hiperbranched HCR; **ITO**: indium tin oxide; **EXPAR**: exponential amplification reaction; **AgNP**: Silver nanoparticles; **AuNPs**: Gold nanoparticles; **Fc**: ferrocene; **PGM**: personal glucose meter; **SDR**: strand displacement reaction; **SA-AuNPs**: streptavidin coated AuNPs; **avidin-GOD**: avidin-tagged glucose oxidase; **SPR**: surface plasmon resonance; **PIP**: p-iodophenol; **HRP**: horseradish peroxidase; **Gox**: glucose oxidase; **ABTS<sup>2-</sup>2,2'** -azino-bis; **ctDNA**: circulating tumor DNA

## 1.4 NOVEL DETECTION STRATEGIES

### 1.4.1 Transduction modes

In recent years, research activity has increased in the search for POC devices that offer tools for biomarker identification. In this section, we discuss approaches that follow these premises to achieve technologies accessible to small and decentralized laboratories in which precision medicine can be implemented. There are several classifications of these detection methods according to the criteria used, among which we distinguish:

- Depending on the **assay format** there are two types based on the use or not of labels. **Label-free** techniques do not use tags to provide a direct signal since are measured changes in physicochemical properties during the interaction between probe and target, such as mass, refractive index, luminescent waves, or electrical impedance (247). While the **label techniques** need a marker to generate or amplify the signal, so their detection is indirect. The labeling techniques are the most widely used because, although label-free techniques simplify the process, their high manufacturing cost and complex assembly make them poor candidates for sensitive and selective POC detection of nucleic acid (47).

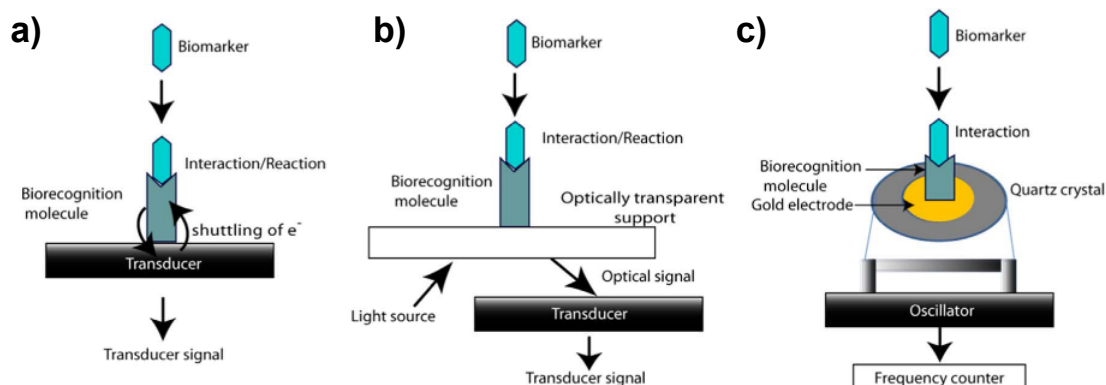
- Depending on the **phase** in which the biorecognition occurs, hybridization assays can be classified as **homogeneous** or **heterogeneous**. More than one phase is involved in the heterogeneous format since the analyte and receptor are in different phases. The biorecognition occurs at the interface between the solid support and the solution. Whereas in the homogeneous, the biorecognition and detection occur in the same phase, usually solution, being a fast and straightforward approach. Although heterogeneous hybridization is slower, it is often the best option for the analysis of multiple complex samples since a previous target separation step is usually required before detection (248).

- Depending on the **nature** of the biorecognition detection system, **electrochemical**, **piezoelectric**, and **optical** ones can be distinguished, as they are typically used in the DNA field (Figure 21). They make it possible to convert the information corresponding to variations of physicochemical properties into quantifiable electrical magnitudes, which can be amplified, stored, and recorded (249).

**Electrochemical** methods measure changes in current, potential, or impedance induced by recognition between the bioreceptor and the analyte. These approaches are inexpensive, fast, and easily miniaturized, so they have been used in multiple works based on the oxidation of DNA bases (direct or enzyme-catalyzed) or in charge transport

reactions. Although their lifetime is limited because electrochemical species may be easily reduced or oxidized due to their low stability.

In contrast, **piezoelectric or mass** systems measure mass changes induced by the formation of the bioreceptor-target complex. Their main advantage is that they allow direct detection without labeling with high sensitivity. However, their multiplexing capacity is limited (250).



**Figure 21.** Different transduction modes of the biorecognition process: a) Electrochemical. b) Optical. c) Mass-based transducers. Original figure from (249).

In this thesis, devices with **optical detection** are considered since they are effortless to use, have high response speed, portability, sensitivity, versatility, robustness, multiplexing capacity, and, especially, present fewer interferences than electrochemical methods. In addition, they take advantage of the multiple properties of light such as transmission, reflection, scattering, fluorescence, intensity, phase, polarization, interferometry, or plasmon resonance. Optical labeling techniques include:

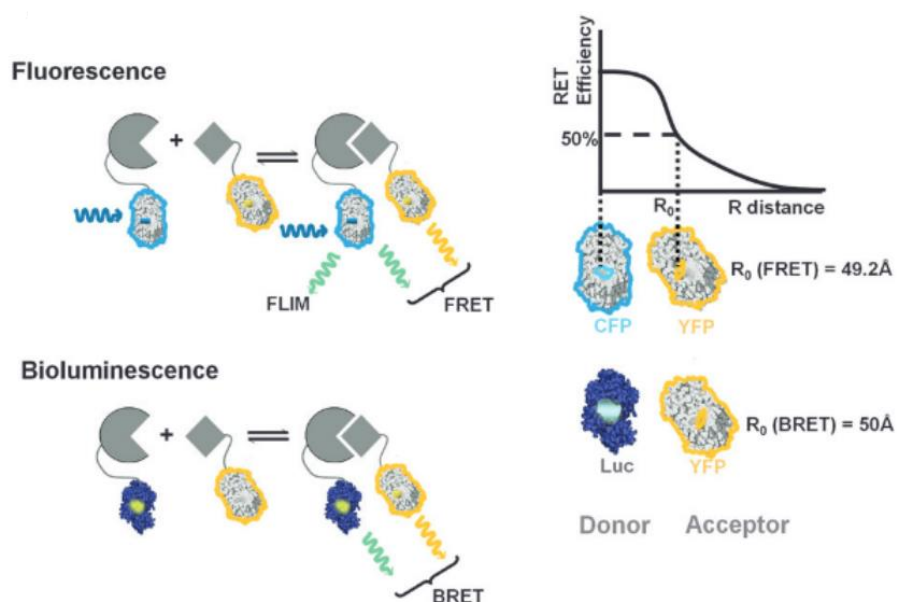
**Fluorescence** techniques are one of the most widespread transduction approaches. The fluorophore is usually incubated directly in the sample, and the emission intensity and decay time are measured. Many fluorophores are available for labeling, which allows multiplexing by displaying a different color in the same assay when hybridizing to different target sequences. The main fluorophores used are the cyanins Cy3 ( $\lambda_{exc} = 550 \text{ nm}$ ;  $\lambda_{em} = 570 \text{ nm}$ ), and Cy5 ( $\lambda_{exc} = 650 \text{ nm}$ ;  $\lambda_{em} = 670 \text{ nm}$ ), and AlexaFluor 46 tags. In this sense, it is possible to use fluorochromes that are nearby and present the emission spectrum of the donor overlapping with the excitation spectrum of the acceptor so that the excitation of one is transferred to the other, leading to the appearance of fluorescence, which is known as **Förster Resonance Energy Transfer (FRET)**.

Fluorescent methods are characterized by easy handling, reasonable sensitivity, and a wide dynamic range of measurement. They are generally less expensive and complicated than radioactive and chemiluminescent labeling but require expensive

instrumentation, and suffer drawbacks like relatively weak intensity, photobleaching, short lifetime and instability, requiring immediate measurement (251).

An emerging alternative is the use of bioluminescence. This detection strategy involves the application of proteins such as **luciferases**. Luciferases produce light without external excitation by the oxidation of a substrate. Bioluminescent proteins can be coupled to an acceptor to enable **Bioluminescent Resonance Energy Transfer (BRET)**, similar to FRET-based approaches (Figure 22). This principle has already been employed to create analogs of fluorescent detection methods, such as bioluminescent molecular beacons (252).

The BRET phenomenon has been used for the detection of nucleic acids using Cy3-labeled beacon probes (253) and of methylated DNA with two modified luciferases and two intercalating agents (BOBO 1 and BOBO 3) (254) but has not been used for SNV discrimination.



**Figure 22.** Comparison between FRET and BRET approaches. Original figure from (252).

**Colorimetric** techniques are a clear alternative to fluorescence detection since they provide a direct visual signal and naked eye that can be quantified with inexpensive equipment, such as a standard document scanner. One of the most used procedures is based on an additional detection biomolecule, such as an antibody, which recognizes the hybridized product and is usually labeled with enzymes or gold nanoparticles. One of the most popular group of enzymes utilized is peroxidases as the horseradish peroxidase enzyme (HRP) due to its high stability, ease of conjugation, and availability of commercial substrates. Subsequently, the reaction developer occurs upon the catalysis of the substrate by the enzyme. In a heterogeneous format, this produces a dark-colored insoluble precipitate that remains adhered to the assay area where the

biorecognition has occurred. TMB (3,3',5,5'-tetramethylbenzidine) is one of the substrates used for HRP. Related to gold nanoparticles, silver is the reference developer. Thus gold nanoparticles conjugated to biomolecules act as a nucleation point, catalyzing the reduction of silver cations and generating a deposit of metallic silver with a grayish appearance adsorbed on the substrate surface, significantly improving the detection limits (255,256).

Another colorimetric approach is the conjugation of specific oligonucleotide probes with gold nanoparticles (AuNPs) since these have a molar extinction coefficient of three to four orders of magnitude higher ( $\epsilon = 2,7 \times 10^8 \text{ M}^{-1} \text{ cm}^{-1}$  for 13 nm AuNP) than most organic molecules. This detection is based on **Surface Plasmon Resonance (SPR)**. When visible light strikes, a specific resonant wavelength is absorbed by this to induce oscillations of the surface electrons. The resulting oscillation generates electromagnetic radiation at the same frequency as the oscillating electrons. When the distance between them is smaller than their size, the resonance of the individual particles starts to hybridize, and, consequently, the maximum absorption peak shifts to a longer wavelength. So that its maximum absorbance peak shifts from around 500-550 nm to 650 nm, and the AuNPs change color from red to purple or blue. In addition to being visible to the naked eye, this color change can also be monitored by UV-visible spectroscopy or scanometric measurements to obtain quantitative or semi-quantitative results (257). This principle has been applied to detect SNV such as variants in the *KRAS* gene from cell lysate samples (258), in *BRCA1* (259), and two mutations in the  $\beta$ -globin gene, associated with  $\beta$ -thalassemia, through an aggregation system mediated by allele-specific DNA ligation (260). Also, for the identification of SNV associated with obesity, thrombotic diseases, or autoimmune diseases.

#### 1.4.2 Consumer electronic devices

In an attempt to overcome the long times and high complexity of genetic testing in a standard laboratory, work has been carried out in various fields over the last decades. Solutions are being sought that use software, hardware, and consumable equipment in the same system, such as smartphones, touch screens, conventional scanners, or compact discs. These commercial products of daily use, being widespread technologies at all levels, have achieved mass production, lowered costs, and increased quality standards. In addition, their elements have high reproducibility, robustness, reliability, versatility, and minimum maintenance (261). For these reasons, consumer electronic devices are proposed as a POC test platform with great potential for detecting SNV.



#### 1.4.2.1 Optical digital disc technology

Among the consumer electronic devices are those that incorporate plastics (discussed in point 2 in microarray platforms) with electronic elements of interest, such as optical digital media, the compact disc (CD), the digital versatile disc (DVD), and the Blu-Ray disc (BD), designed to store any type of information (audio, video, applications, or documents).

In 1979, in collaboration Philips and Sony developed a robust CD audio system. The DVD emerged in 1996 from the need to have a medium capable of storing audiovisual content that improved in quality and had greater capacity than the CD. Finally, to reproduce all types of audiovisual documents in high definition, Sony, Philips, and Pioneer officially announced in 2002 the creation of the BD disc, which had a capacity five times higher than that of DVD.

At first glance, the three discs look the same since they have a circular shape of 120 mm in diameter, a thickness of 1.2 mm, a central hole of 15 mm and a functional surface of approximately 95 cm<sup>2</sup>, which allows them to perform hundreds of tests on a single disc (262). However, the layers that compose them are different (Figure 23a). CDs have the fewest layers compared to DVDs and BDs, which have 7. Though, all of them have a metallic layer, usually aluminum, silver, or gold, covering the internal structure where the information is stored, called **track**, to reflect the laser beam of the reader. And a protective layer of polymeric lacquer. CDs and DVDs are composed of a transparent plastic substrate, usually PC, an ideal material for the immobilization of target probes, as we already discussed. Instead, the BD's underside is protected by a compound known as **Durabis** (patented by *Tokyo Denkikagaku Kogyo KK*), which is chemically inert and highly hydrophobic. This gives them properties such as reduced background signal, hydrophobicity, durability, anti-adhesion, and scratch resistance, allowing smaller and more regular dots to be printed when used for micro-arrays.

In addition, the track, which consists of a spiral-shaped groove from the inner part of the disk to the outer part, and whose size depends on the storage capacity, also has apparent differences. Each type of disc has this structure in different positions. In CDs, the structure is in the upper part of the disc, while DVDs, composed of two 0.6 mm thick PC discs, have it in the central part. BD discs have the structure in the lower part of the disc, protected by Durabis (263).

Disc readers have many elements in common. They all have a rotating shaft and a motor where the disc is coupled to start rotating. A head that integrates both the laser source and all the optical elements (photodiode) necessary to collect the reflected laser beam and separate it from the incident one. Also, a rail with a motor in charge of moving

the laser horizontally, allowing scanning of all disc radii. The scanning bases on the laser incidence on its surface, which passes through PC and focuses on the reflective metallic layer, where the internal groove of the disk is located. The laser follows this groove from inside to outside in a levorotatory direction while the disk rotates, thus scanning the entire surface. The laser light is reflected back to the reader's optical head, which redirects it to the photodiode, continuously recording, converting, and amplifying the reflected laser signal into an analog signal (video, audio, or images) (Figure 23b) (264).

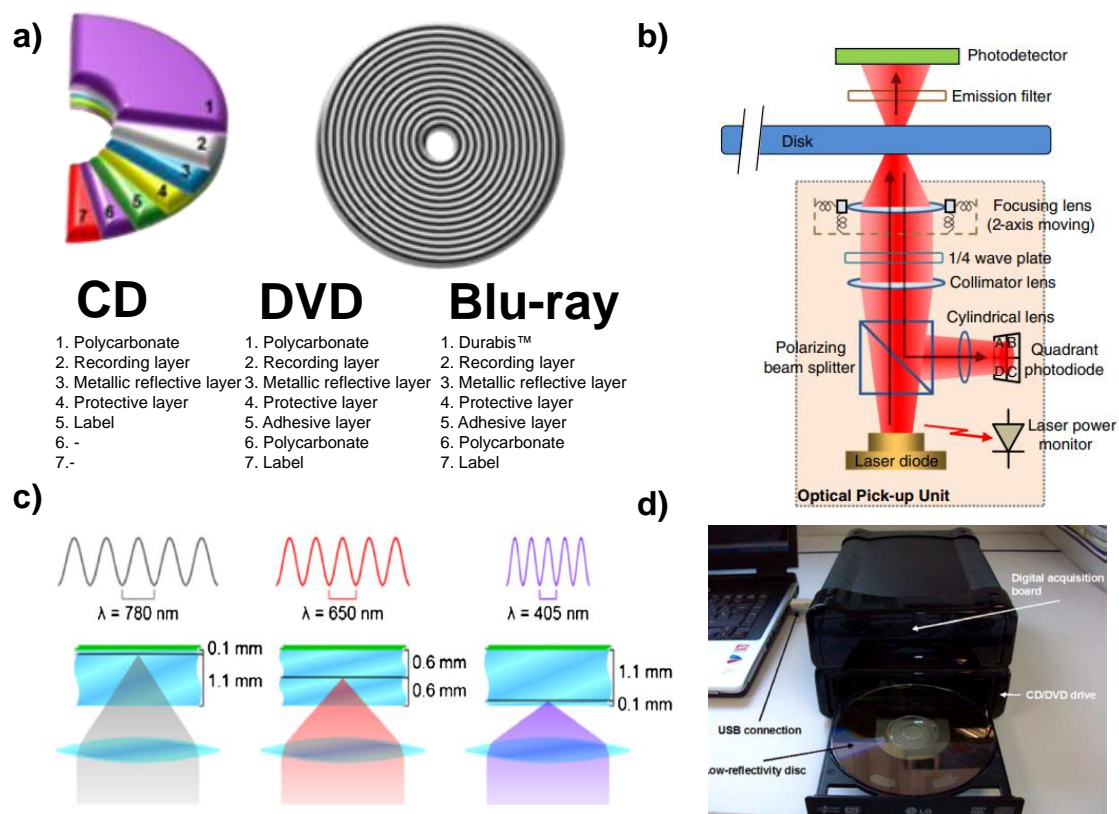
Nevertheless, the differences between the laser diode wavelength, objective lenses, and optical characteristics change depending on the disc (Figure 23c).

In a pioneering work (265), the potential of optical disc technology for the fluorescent detection of pesticides by inhibition immunoassays was demonstrated for the first time, being the starting of the use of commercial optical disc platforms in the field of bioanalytics. Since this work, many publications have been reported using discs as analytical platforms, which can be grouped according to two different approaches.

One option is the use of compact discs with **micro-reactors using microfluidics**. This technology allows the miniaturization and automation of analytical systems to develop portable platforms. Microfluidic devices, also called "laboratories on chips", have a size in the order of microns, formed by different channels, reservoirs, and valves that communicate the parts of the disk, and in which one or several steps of the analytical process are integrated (sample treatment, amplification, or detection). In this way, it is allowed the reduction of reagent consumption, sample, and time. They use centrifugal force as a driving system to move the fluids and perform the tests, and the results are read with conventional detectors (photometers, density meters, or fluorimeters) adapted to a circular geometry. Compared to conventional microfluidics, those applied to POC have the added difficulty of extreme simplicity (266,267). For SNV genotyping, microfluidic fluorescent detection of PCR products has been combined with AS hybridization in endpoint (268) or real-time (269) measurement. However, further research is needed because its manufacturing process is generally expensive and complex, it is often irreproducible, and sample handling remains critical (267).

Another option is to use **audio-video technology** both for the performing of the tests (discs) and for their optical reading (reader/recorder) since the disc reader scans the entire surface of the test with high precision (270). In this case, assay cost is much lower since a conventional disk reader is around 50-100 € and the disks 0.10-0.25 €. In addition, this alternative also makes data acquisition easier, as the readers are compatible with the computer hardware. The use of this technology allows sensitive, rapid assays of multiple analytes in a microarray format. Since 2002, the conventional

disk reader has been used as a screening tool for discs-based assays of hybridized nucleic acids or molecular screening (Figure 23d) (271).



**Figure 23.** Outline of the most common compact discs. a) Set of layers forming a compact disc. b) Diagram of the interior of an optical pickup unit. c) Different positions for each disc where the reader's laser is focused and its features. d) Disc drive prototype reader. Figures adapted from (262).

The basis of the detection principle using compact disc technology lies in the variation of the optical properties of the discs due to molecular recognition events. The reading of the discs can be performed either by reading error analysis or by analyzing the attenuation of the reflected laser. Read error analysis can be performed with a commercial disk reader without any modification and with commercial software. However, this strategy shows a high background signal, providing a low signal-to-noise ratio (SNR) and, therefore, a low sensitivity. The analysis of the reflected laser attenuation is based on the fact that the laser beam generates reading errors when hitting more minor elements than this one, while the larger ones generate an attenuation of the reflected laser intensity. This strategy shows better sensitivities than the previous one, although it requires a slight modification of the disk reader (272). In this case, a data acquisition card is added to the reader, which collects the measurement signals and digitizes them for storage, analysis, and display on a computer. In this way, an image is obtained with the reflection intensity values of the entire test area (262).

In fact, since the emergence of this technology, our group has developed numerous advances using optical disks as an analytical platform and the readers as detectors for the development of assays based on probe-target molecular recognition of proteins and nucleic acids with different applications. For example, in 2014, they developed a high-density microarray on a BD disk for mass screening of different compounds (273). Additionally, in 2015, isothermal amplification technologies, RPA and MDA, were combined with the subsequent detection of their products using DVD technology applied to the duplex detection of pathogenic bacteria *Salmonella spp.* and *Cronobacter spp.* (274). Even 28 SNV associated with psychiatric pharmacogenomics have been identified by combining ligation with a universal PCR and hybridizing the products in microarray in DVDs (275).

All these results demonstrate that optical disc technology is a potent tool to be employed as a platform for performing genetic analysis with an integrated POC approach.

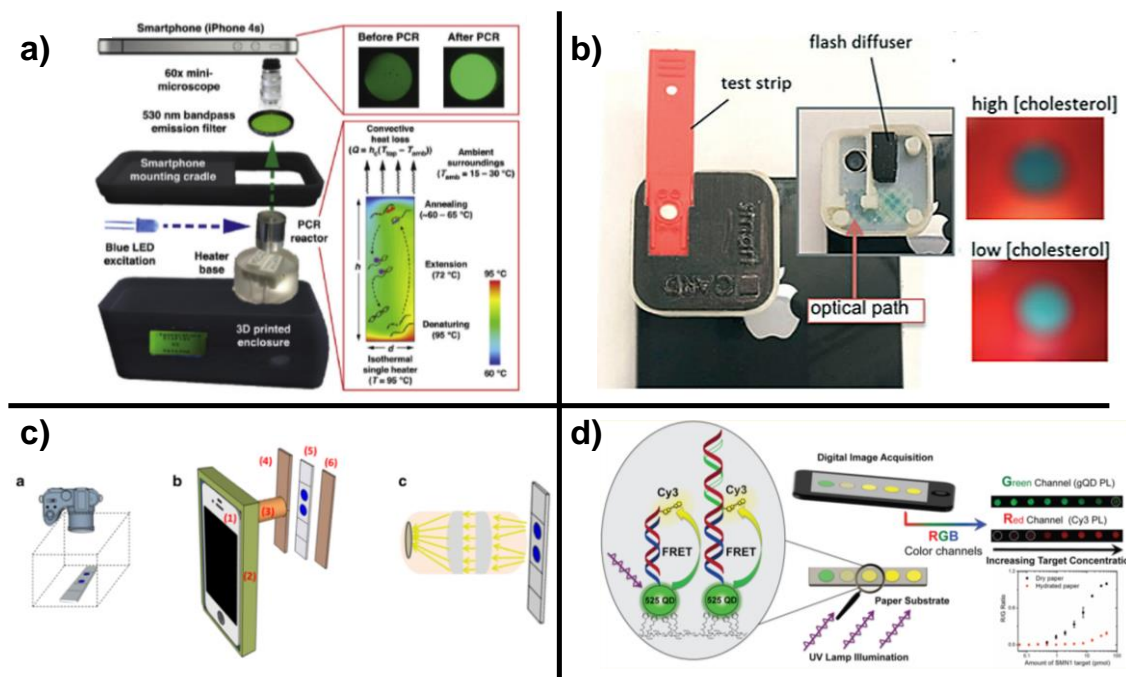
#### 1.4.2.2 Smartphone technology

The smartphone, a wireless electronic device, has changed our lives in many ways and is widely accepted by the population. Smartphones use a high-performance operating system that manages the device's resources, stores data, and optimizes different functions. Their applications go beyond simple communication since they currently have: multiple complementary metal-oxide-semiconductor (CMOS sensor) cameras, lantern mode with LED flash, connection to the Internet, and different networks using wireless technologies such as Infrared, Bluetooth, Wi-Fi, global positioning system (GPS), among others, and use touch screens, so they are generally tools with enough computational power similar to miniature computers (261).

The quality of CMOS sensor chips embedded in phone cameras is increasing exponentially, resulting in devices with high-quality camera lenses. Their characteristics are continuously being improved, offering good imaging characteristics (recently reaching a resolution of more than 40 megapixels), as well as new apps are becoming widely available, making them ideal detectors for cost-effective assays (276).

Their main advantages are low-cost, portable, ubiquitous, small size, and user-friendly, as they do not require experts for their use. Also, they have a battery that allows them to operate with autonomy for days without connecting to a charger or external power source. Therefore, due to their unique characteristics, smartphones have been studied in the last years as an alternative to traditional laboratory detection equipment, such as microscopes or spectrophotometers, as they can serve both as detector and bioassay data processors for different applications (health, food safety, environmental

monitoring, biosecurity) especially in remote areas and tests performed outside the laboratory. In addition, smartphones are compatible with accessories such as lenses, filters, alternative LEDs, and diffraction gratings, making them versatile detectors combined with almost all-optical techniques (255). Considering POC for DNA diagnostic, investigations have been reported that combine isothermal amplification, such as LAMP, RPA, and RCA, with smartphone detection, especially of pathogens (277). Although the most widely used has been LAMP, in which the reaction products are labeled with fluorescence using different supports (278,279).



**Figure 24.** Examples of applications that use smartphones as readers of the biorecognition process. a) Lab-on-a-drone system of PCR reactions (261). b) Colorimetric test strip assays (277). c) Paper-based chemiluminescence (280). d) Paper-based solid-phase FRET (281). Figures taken from the cited publications.

Indeed, studies using smartphone-based devices to photograph the results obtained have also been reported to apply them to SNV genotyping. Regarding those combined with colorimetric assays, the first was in 2008 for SNV without reported clinical significance (rs2910190, rs3811999, rs1862391) and was based on the anchoring of AS-LNA primers in array format and the addition of biotin-dUTPs. When the primers were extended, the new nascent chains were recognized by streptavidin-alkaline phosphatase complexes producing a visual signal (282). Another approach to detecting SNV in the *SEMA4D* gene (rs11526468 and rs12979860) used T7 endonuclease to cleave the PCR-generated products and detect them by HRP-mediated immunoassay (283). Another example in which the G > C transversion in the *TP53* gene was discriminated from gDNA captured silver ions by activating urease, so its presence could be distinguished with a pH strip (284). Recently, our group has combined LAMP amplification of the SNV in the

*GRIK4* gene in solid and liquid phases with the smartphone and an external LED light source (285).

Smartphone-detected fluorescent assays for SNV genotyping have also been described, mainly using microarray technology. For example, in one study, they labeled the products formed by the RCA, the wild type *KRAS* with Cy3, and the c.34G > A mutants with Cy5 so that they obtained dual-color images using a smartphone that has an integrated microscope with lasers for the fluorophores. In another, using a dry paper substrate immobilized green quantum dots that served as FRET donors and Cy3 as acceptors when perfect-match hybridization of synthetic sequences occurred and was detected using the RGB (Red-Green-Blue) channels of the mobile camera with an external excitation source (281). Another example is FRET detection in a solution of SNV in microRNA using molecular beacons with and smartphone camera as a fluorescence spectrometer (286).

Chemiluminometric has also been employed for SNV genotyping, such as, for example, the detection of C677T of the *MTHFR* gene. The products amplified by PCR and by first extension reaction were hybridized, and their signal reported by the SA-HRP conjugates and chemiluminogenic substrate in lateral flow strips that fits to a fabricated lens holder connected to the smartphone (280).

## 1.5. METHODS COMPARISON

### 1.5.1 Recapitulation

After reviewing the literature, we conclude that primary prevention programs, early diagnosis, and therapeutic measures based on the detection of molecular biomarkers are the key elements in the fight against diseases. According to the state-of-art, there are several technologies available for SNV genotyping as showed in Table 13. The currently available platforms that can be organized into three main categories: **Screening techniques**. They use very sophisticated equipment, such as sequencers, spectrophotometers, chromatographs, or genetic analyzers, requiring bioinformatics tools and genetic expertise to process and interpret the data. These are high-throughput platforms capable of detecting point mutations when they are present between 1% and 25% in a native matrix. Among them, sequencing stands out as the gold standard since it can identify all SNV presented in the sample, but it has long analysis times, requires a large number of mutant samples, and the cost of analysis is high. On the one hand, these difficulties can be overcome with MALDI-TOFF-MS or dHPLC technologies, which require fewer mutant alleles and guarantee a primary evaluation of the amplified product without risk of contamination. However, they do not identify specific mutations, and the consecutive application of a direct method of confirmation of the analytical results is necessary. SNaPshot, on the other hand, can analyze up to ten nucleotide substitutions in a single reaction, specifically characterizing the mutation, but does not investigate less frequent or rare mutations.

- **Techniques based on PCR amplification.** They require a thermal cycler simple or accoupled to a detector. A relevant group is real-time PCR methods using specific probes, such as Scorpion ARMS, TaqMan, AS-PCR, or HRM (post-PCR). They are highly sensitive technologies (0.1-5%) with a more elevated cost-effectiveness ratio than sequencing, but only capable of identifying known mutations. On the other hand, PCR-based enrichment techniques, such as clamp-PCR or COLD-PCR, improve the selectivity of the analysis (0.01-1%) because they rely on targeted amplification to increase the concentration of mutated versus native DNA. However, they require complicated assay optimization processes. In particular, ddPCR achieves the highest sensitivity (up to 0.001%) because the enrichment of mutant molecules per reaction is proportional to the number of partitions in the sample; nevertheless, it requires specialized equipment and high-cost specific reagents.

- **Hybridization-based techniques.** They include microarrays and nuclease assays (PCR-RFLP), highlighted for their high selectivity and multiplexing. In this sense, microarrays allow rapid and specific identification of the mutated base with low-cost

equipment. Regardless, these technologies are dependent on a previous PCR step to achieve adequate sensitivity.

**Table 13.** Techniques currently available for SNV genotyping.

Techniques	LOD	Time	Multiplexing (samples/run)	Equipment	Comments	Ref.
<b>Sanger sequencing</b>	10-25%	2 days	1-10	Sequencers	Requires a high amount of mutated DNA	(42)
<b>Pyrosequencing</b>	3-7%	1 day	up to 96	Sequencers	High error rate in homopolymer readout (>5 ntd)	(35)
<b>MALDI-TOF-MS</b>	5-10%	4-5 hours	up to 96	Mass spectrometer	Needs multiple preparation steps and high purity samples	(287)
<b>dHPLC</b>	3-5%	Few hours	1	HPLC equipment	Cannot detect homozygous mutations directly nor determine the mutation type	(288)
<b>SNaPshot</b>	1-5%	7 hours	10	Capillary genetic analyzer	Multi-step process and each assay run separately	(289)
<b>Scorpion ARMS</b>	1%	8 hours	up to 96	Specific real-time thermal cycler	Scorpion probe design is very complex and laborious	(110)
<b>AS-PCR</b>	0.1-1%	3 hours	up to 96	Real-time thermal cycler	May lead to false-positives results	(108)
<b>HRM</b>	1-5%	3 hours	up to 96	Real-time thermal cycler	Requires a previous PCR amplification step	(62)
<b>Clamp PCR</b>	0.1-1%	2 hours	up to 96	Real-time thermal cycler	Tedious blocker design	(115)
<b>COLD-PCR</b>	0.01-1%	3 hours	up to 96	Real-time thermal cycler	Requires precise control of working temperatures	(120)
<b>Droplet digital PCR</b>	0.001-0.1%	2 hours	up to 96	Sophisticated thermal cycler	Absolute quantification. Costly for both the equipment and the reagents	(76)
<b>PCR-RFLP</b>	5%	4-5 hours	up to 96	Thermal cycler	Limit to targets containing sequences recognizable by the endonucleases	(91)
<b>Microarraying</b>	0.1-1%	3 hours	hundreds	Chip reader	Challenging to optimize probes design and hybridization conditions. Low-cost.	(52)



All these techniques share a high accuracy in determining the sequence of bases in the genome and a high working capacity. Therefore, they are handy for genotyping thousands of SNV and large numbers of samples, i.e., on a large scale. Nonetheless, they require high instrument complexity, high costs, and lengthy analysis times, demanding specialized equipment and human resources.

### **1.5.2 Unsolved challenges**

Despite the wide range of available technologies, this is not a complete scientific-technical field. Many of these technologies are not incorporated into all healthcare systems for reasons such as cost, time, required resources, limited analytical performances, sophisticated equipment and specialized personnel, being only accessible to large hospitals and centralized laboratories. Thus, there is a clear necessity for new strategies that can be implemented in any clinical scenario. The investigation in alternative genotyping technologies will help to implement personalized medicine by overcoming the limitations of current clinical practice.

This doctoral thesis aims to contribute to this issue with novel analytical solutions. The working hypothesis is based on developing innovative assays capable of detecting and identifying a small group of genetic biomarkers, allowing an individualized molecular diagnosis. These new strategies should stand out for their applicability in any environment, simplicity, high speed, and low cost in contrast to current high-throughput platforms.

To this end, one of the solutions is to investigate methods based on PCR, the most widespread and established technique, but modifying the working conditions to obtain better analytical performances. Thus, a reasonable approach is to develop a variant of blocked PCR with real-time fluorescent detection. Another interesting approach for environments without access to sophisticated technologies is alternative reactions, such as isothermal amplification techniques. An appealing proposition is to combine isothermal techniques with discriminatory reactions and non-enzymatic amplification to increase sensitivity and selectivity, entitling recognition of the sequence of interest. On the other hand, to reduce the complexity and costs of detection, the study of different optical systems based on innovative transduction principles can be useful. Mainly, certain platforms and assay formats can be coupled with consumer electronic devices, allowing POC detection. Through a single assay, these strategies will predict and classify patients in the most appropriate population group to receive a personalized treatment.

## 1.6 REFERENCES

1. Indicadores del desarrollo mundial [Internet]. Google Public Data Explorer. 2021 [cited 2021 Oct 5]. Available from: [https://www.google.com/publicdata/explore?ds=d5bncppjof8f9\\_&ctype=c&strail=false&bc=d&nslm=s&met\\_y=sp\\_dyn\\_le00\\_in&scale\\_y=lin&ind\\_y=false&idim=world:Earth&ifdim=world&tunit=Y&pit=1494212400000&hl=es&dl=es&ind=false&icfg](https://www.google.com/publicdata/explore?ds=d5bncppjof8f9_&ctype=c&strail=false&bc=d&nslm=s&met_y=sp_dyn_le00_in&scale_y=lin&ind_y=false&idim=world:Earth&ifdim=world&tunit=Y&pit=1494212400000&hl=es&dl=es&ind=false&icfg)
2. Organización Mundial de la Salud [Internet]. 2021 [cited 2021 Oct 5]. Available from: <https://www.who.int/es>
3. Atkinson AJ, Colburn WA, DeGruttola VG, DeMets DL, Downing GJ, Hoth DF, et al. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001;69(3):89–95.
4. Jain KK. *The Handbook of Biomarkers* [Internet]. New York: Humana Press; 2017. 760p. doi: 10.1007/978-1-4939-7431-3
5. FDA-NIH Biomarker Working Group. BEST ( Biomarkers , EndpointS , and other Tools ) Resource [Internet]. Silver Spring: US Food and Drug Administration; 2021. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK326791/> Co-published by National Institutes of Health (US), Bethesda (MD).
6. Online Biomarker Database | Efficacy | Prognosis | Diagnosis [Internet]. 2021 [cited 2021 Oct 5]. Available from: <https://gobiomdbplus.com/>
7. Ziegler A, Koch A, Krockenberger K, Großhennig A. Personalized medicine using DNA biomarkers: A review. *Hum Genet.* 2012;131(10):1627-38.
8. Craig Venter J, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. *Science.* 2001;291(5507):1304–51.
9. Collins FS, Morgan M, Patrinos A. The Human Genome Project: Lessons from large-scale biology. *Science.* 2003;300(5617):286–90.
10. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet.* 2006;7(2):85–97.
11. Feero G., Guttmacher A, Collins F. Genomic medicine—an updated primer. *N Engl J Med.* 2010;362(21):2001-11.
12. Karki R, Pandya D, Elston RC, Ferlini C. Defining “mutation” and “polymorphism” in the era of personal genomics. *BMC Med Genomics.* 2015;8(1);1-7.
13. Belmont JW, Hardenbol P, Willis TD, Yu F, Yang H, Ch’Ang LY, et al. The international HapMap project. *Nature.* 2003;426(6968):789–96.
14. Shastry BS. SNPs: Impact on Gene Function and Phenotype [Internet]. *Methods Mol Biol.* 2009;578:3–22.
15. Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med.* 2010;363(2):166–76.
16. Siegel RL, Miller KD, Jemal A. Cancer statistics. *CA Cancer J Clin.* 2020;70(1):7–30.
17. Falcomatà C, Schneider G, Saur D. Personalizing KRAS-mutant allele-specific therapies. *Cancer Discov.* 2020;10(1):23–5.
18. Chen CC, Er TK, Liu YY, Hwang JK, Barrio MJ, Rodrigo M, et al. Computational Analysis of KRAS Mutations: Implications for Different Effects on the KRAS p.G12D and p.G13D Mutations. *PLoS One.* 2013;8(2):e55793. doi: 10.1371/journal.pone.0055793
19. Karnoub AE, Weinberg RA. Ras oncogenes: Split personalities. *Nat Rev Mol Cell Biol.* 2008;9(7):517–31.
20. Allegra CJ, Rumble RB, Hamilton SR, Mangu PB, Roach N, Hantel A, et al. Extended RAS gene mutation testing in metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy: American society of clinical oncology provisional clinical opinion update 2015. *J Clin Oncol.* 2016;34(2):179–85.
21. Fiore F Di, Sesboü R, Michel P, Sabourin JC, Frebourg T. Molecular determinants of anti-EGFR sensitivity and resistance in metastatic colorectal cancer. *Br J Cancer.* 2010;103(12):1765–72.
22. Das V, Kalita J, Pal M. Predictive and prognostic biomarkers in colorectal cancer: A systematic review of recent advances and challenges. *Biomed Pharmacother.* 2017;87:8–19.
23. Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer special articles. *Ann Oncol.* 2016;27(8):1386–422.

24. Relling M V., Evans WE. Pharmacogenomics in the clinic. *Nature*. 2015;526(7573):343–50.
25. Ufer M. Comparative Pharmacokinetics of Vitamin K Antagonists. *Clin Pharmacokinet*. 2005;44(12):1227–46.
26. Kamali F, Wynne H. Pharmacogenetics of warfarin. *Annu Rev Med*. 2010;61:63–75.
27. Rieder MJ, Reiner AP, Gage BF, Nickerson DA, Eby CS, McLeod HL, et al. Effect of VKORC1 Haplotypes on Transcriptional Regulation and Warfarin Dose. *N Engl J Med*. 2005;352(22):2285–93.
28. Yin T, Miyata T. Warfarin dose and the pharmacogenomics of CYP2C9 and VKORC1 - Rationale and perspectives. *Thromb Res*. 2007;120(1):1–10.
29. Perk J, De Backer G, Gohlke H, Graham I, Reiner Ž, Verschuren M, et al. European Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J*. 2012;33(13):1635–701.
30. Chan IS, Ginsburg GS. Personalized medicine: Progress and promise. *Annu Rev Genomics Hum Genet*. 2011;12:217–44.
31. Schwarzenbach H, Hoon DSB, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*. 2011;11(6):426–37.
32. Sanger F, Nicklen S, Coulson A. DNA sequencing with chain-terminating. *Proc Natl Acad Sci USA*. 1977;74(12):5463–7.
33. Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, et al. Fluorescence detection in automated DNA sequence analysis. *Nature*. 1986;321(6071):674–9.
34. Liu L, Li Y, Li S, Hu N, He Y, Pong R, et al. Comparison of Next-Generation Sequencing Systems. 2012;2012.
35. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next- generation sequencing for detecting rare and subclonal mutations. *Nat Publ Gr*. 2018;19(5):269–85.
36. Gagan J, Van Allen EM. Next-generation sequencing to guide cancer therapy. *Genome Med*. 2015;7(1):1–10.
37. Ronaghi M. Pyrosequencing sheds light on DNA sequencing. *Genome Res*. 2001;11(1):3–11.
38. Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends Genet*. 2008;24(3):133–41.
39. Goodwin S, Mcpherson JD, McCombie WR. Coming of age : ten years of next- generation sequencing technologies. *Nat Rev Genet*. 2016;17(6):333-51.
40. Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011;475(7356):348–52.
41. Gupta PK. Single-molecule DNA sequencing technologies for future genomics research. *Trends Biotechnol*. 2008;26(11):602–11.
42. Bairi K El, Chapter B. Illuminating Colorectal Cancer Genomics by Next-Generation Sequencing [Internet]. Switzerland: Springer Nature; 2020. 191p. doi: 10.1007/978-3-030-53821-7
43. Clarke J, Wu HC, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol*. 2009;4(4):265–70.
44. Pirrung MC. How to make a DNA chip. *Angew Chem Int Ed*. 2002;41(8):1276–89.
45. Khodakov D, Wang C, Zhang DY. Diagnostics based on nucleic acid sequence variant profiling: PCR, hybridization, and NGS approaches. *Adv Drug Deliv Rev*. 2016;105:3–19.
46. Bañuls MJ, Morais SB, Tortajada-Genaro LA, Maqueira Á. Microarray Developed on Plastic Substrates. In: Li P, Sedighi A, Wang L, editores. *Microarray Technology*. New York: Humana Press; 2016.p. 37-51. doi: 10.1007/978-1-4939-3136-1
47. Ventimiglia G, Petralia S. Recent Advances in DNA Microarray Technology: An Overview on Production Strategies and Detection Methods. *Bionanoscience*. 2013;3(4):428–50.
48. Bumgarner R. Overview of dna microarrays: Types, applications, and their future. *Curr Protoc Mol Biol*. 2013;101(1):1–11.
49. Prieto-Simon B, Campas M, Marty J-L. Biomolecule Immobilization in Biosensor Development: Tailored Strategies Based on Affinity Interactions. *Protein Pept Lett*. 2008;15(8):757–63.
50. Wilchek M, Bayer EA, Livnah O. Essentials of biorecognition: The (strept)avidin-biotin system as a model for protein-protein and protein-ligand interaction. *Immunol Lett*. 2006;103(1):27–32.
51. Desmet C, Marquette CA. Surface Functionalization for Immobilization of Probes on

- Microarrays. In: Li P, Sedighi A, Wang L, editores. *Microarray Technology*. New York: Humana Press; 2016.p. 7-23. doi: 10.1007/978-1-4939-3136-1
52. Sahebrao Batule B. DNA Microarray-Based Technologies to Genotype Single Nucleotide Polymorphisms. In: Chang HN, editores. *Emerging Areas in Bioengineering*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2018.p. 531-56. doi: 10.1002/9783527803293
  53. Letowski J, Brousseau R, Masson L. Designing better probes: Effect of probe size, mismatch position and number on hybridization in DNA oligonucleotide microarrays. *J Microbiol Methods*. 2004;57(2):269–78.
  54. Marzancola MG, Sedighi A, Li PCH. DNA Microarray-Based Diagnostics. In: Li P, Sedighi A, Wang L, editores. *Microarray Technology*. New York: Humana Press; 2016.p.161-78. doi: 10.1007/978-1-4939-3136-1
  55. Traeger JC, Schwartz DK. Surface-Mediated DNA Hybridization: Effects of DNA Conformation, Surface Chemistry, and Electrostatics. *Langmuir*. 2017;33(44):12651–9.
  56. Yilmaz LS, Loy A, Wright ES, Wagner M, Noguera DR. Modeling formamide denaturation of probe-target hybrids for improved microarray probe design in microbial diagnostics. *PLoS One*. 2012;7(8). doi: 10.1371/journal.pone.0043862
  57. Bielec K, Sozanski K, Seynen M, Dziekan Z, Ten Wolde PR, Holyst R. Kinetics and equilibrium constants of oligonucleotides at low concentrations. Hybridization and melting study. *Phys Chem Chem Phys*. 2019;21(20):10798–807.
  58. Mullis KB, Faloona FA. Specific Synthesis of DNA in Vitro via a Polymerase-Catalyzed Chain Reaction. *Methods Enzymol*. 1987;155:335–50.
  59. Bartlett JMS, Stirling D. A Short History of the Polymerase Chain Reaction. In: John M. S. Bartlett JMS, Stirling D, editores. *PCR Protocols*. New York: Humana Press; 2003.p. 3-6. doi: 10.1385/1-59259-384-4
  60. Saik RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, et al. Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. *Science*. 1987;239(29):487–91.
  61. Jeffrey S, Chamberlain JS, Gibbs RA, Ranier JE, Nguyen PN, Thomas C. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res*. 1988;16(23):11141–56.
  62. Ihle MA, Fassunke J, König K, Grünwald I, Schlaak M, Kreuzberg N, et al. Comparison of high resolution melting analysis, pyrosequencing, next generation sequencing and immunohistochemistry to conventional Sanger sequencing for the detection of p.V600E and non-p.V600E BRAF mutations. *BMC Cancer*. 2014;14(1):1-13.
  63. Crockett AO, Wittwer CT. Fluorescein-labeled oligonucleotides for real-time PCR: Using the inherent quenching of deoxyguanosine nucleotides. *Anal Biochem*. 2001;290(1):89–97.
  64. Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ. High-resolution genotyping by amplicon melting analysis using LCGreen. *Clin Chem*. 2003;49(6):853–60.
  65. Liew M, Pryor R, Palais R, Meadows C, Erali M, Lyon E, et al. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clin Chem*. 2004;50(7):1156–64.
  66. Zhou L, Myers AN, Vandersteen JG, Wang L, Wittwer CT. Closed-tube genotyping with unlabeled oligonucleotide probes and a saturating DNA dye. *Clin Chem*. 2004;50(8):1328–35.
  67. Chou LS, Meadows C, Wittwer CT, Lyon E. Unlabeled oligonucleotide probes modified with locked nucleic acids for improved mismatch discrimination in genotyping by melting analysis. *Biotechniques*. 2005;39(5):644–7.
  68. Zhou L, Palais RA, Ye F, Chen J, Montgomery JL, Wittwer CT. Symmetric snapback primers for scanning and genotyping of the cystic fibrosis transmembrane conductance regulator gene. *Clin Chem*. 2013;59(7):1052–61.
  69. Sundberg SO, Wittwer CT, Howell RM, Huuskonen J, Pryor RJ, Farrar JS, et al. Microfluidic genotyping by rapid serial PCR and high-speed melting analysis. *Clin Chem*. 2014;60(10):1306–13.
  70. Arya M, Shergill IS, Williamson M, Gommersall L, Arya N, Patel HRH. Basic principles of real-time quantitative PCR. *Expert Rev Mol Diagn*. 2005;5(2):209–19.
  71. Morrison TB, Weis JJ, Wittwer CT. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques*. 1998;24(6):954–8.
  72. Navarro E, Serrano-Heras G, Castaño MJ, Solera J. Real-time PCR detection chemistry. *Clin Chim Acta*. 2015;439:231–50.

73. Vogelstein B, Kinzler KW. Digital PCR. *Proc Natl Acad Sci U S A*. 1999;96(16):9236–41.
74. Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc Natl Acad Sci U S A*. 2003;100(15):8817–22.
75. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem*. 2011;83(22):8604–10.
76. McEvoy AC, Wood BA, Ardakani NM, Pereira MR, Pearce R, Cowell L, et al. Droplet Digital PCR for Mutation Detection in Formalin-Fixed, Paraffin-Embedded Melanoma Tissues: A Comparison with Sanger Sequencing and Pyrosequencing. *J Mol Diagnostics*. 2018;20(2):240–52.
77. Barany F. The ligase chain reaction in a PCR world. *Genome Res*. 1991;1(1):5–16.
78. Kim S, Misra A. SNP genotyping: Technologies and biomedical applications. *Annu Rev Biomed Eng*. 2007;9(1):289–320.
79. Wiedmann M, Wilson WI, Luo J, Barany F, Batt A. Ligase Chain Reaction (LCR) -Overview and Applications. *PCR Methods Appl*. 1994;3:S51–64.
80. Gibriel AA, Adel O. Advances in ligase chain reaction and ligation-based amplifications for genotyping assays: detection and applications. *Mutat Res Rev Mutat Res*. 2017;773:66–90.
81. Hamada M, Shimase K, Tsukagoshi K, Hashimoto M. Discriminative detection of low-abundance point mutations using a PCR/ligase detection reaction/capillary gel electrophoresis method and fluorescence dual-channel monitoring. *Electrophoresis*. 2014;35(8):1204–10.
82. Osiowy C. Sensitive detection of HBsAG mutants by a gap ligase chain reaction assay. *J Clin Microbiol*. 2002;40(7):2566–71.
83. Yi P, Jiang H, Li L, Dai F, Zheng Y, Han J, et al. A New Genotyping Method for Detecting Low Abundance Single Nucleotide Mutations Based on Gap Ligase Chain Reaction and Quantitative PCR Assay. *Cell Biochem Biophys*. 2012;62(1):161–7.
84. Marshall RL, Laffler TG, Cerney MB, Sustachek JC, Kratochvil JD, Morgan RL. Detection of HCV RNA by the asymmetric gap ligase chain reaction. *PCR Methods Appl*. 1994;4(2):80–4.
85. Tobler AR, Short S, Andersen R, Paner TM, Jason C, Lambert SM, et al. The SNPlex Genotyping System: A Flexible and Scalable Platform for SNP Genotyping. *J Biomol Tech*. 2005;16(4):396–404.
86. Pindo M, Vezzulli S, Coppola G, Cartwright DA, Zharkikh A, Velasco R, et al. SNP high-throughput screening in grapevine using the SNPlex genotyping system. *BMC Plant Biol*. 2008;8(1):1–6.
87. Schouten JP, Mcelgunn CJ, Waaijjer R, Zwiijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res*. 2002;30(12):e57. doi: 10.1093/nar/gnf056
88. Tortajada-Genaro LA, Niñosoles R, Mena S, Maquieira Á. Digital versatile discs as platforms for multiplexed genotyping based on selective ligation and universal microarray detection. *Analyst*. 2019;144(2):707–15.
89. Chan SH, Stoddard BL, Xu SY. Natural and engineered nicking endonucleases - From cleavage mechanism to engineering of strand-specificity. *Nucleic Acids Res*. 2011;39(1):1–18.
90. Li WM, Hu TT, Zhou LL, Feng YM, Wang YY, Fang J. Highly sensitive detection of the PIK3CA H1047R mutation in colorectal cancer using a novel PCR-RFLP method. *BMC Cancer*. 2016;16(1):1–11.
91. Hashemi SA, Khoshi A, Ghasemzadeh-Moghaddam H, Ghafouri M, Taghavi M, Namdar-Ahmadabad H, et al. Development of a PCR-RFLP method for detection of D614G mutation in SARS-CoV-2. *Infect Genet Evol*. 2020;86:104625.
92. Abi A, Safavi A. Targeted Detection of Single-Nucleotide Variations: Progress and Promise. *ACS Sens*. 2019;4(4):792–807.
93. Lyamichev V, Mast AL, Hall JG, Prudent JR, Kaiser MW, Takova T, et al. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat Biotechnol*. 1999;17(3):292–6.
94. Chen Z, Miao L, Liu Y, Dong T, Ma X, Guan X, et al. A universal genotyping-microarray constructed by ligating a universal fluorescence-probe with SNP-encoded flaps cleaved from multiplex invasive reactions. *Chem Commun*. 2017;53(96):12922–5.

95. Zou B, Cao X, Wu H, Song Q, Wang J, Kajiyama T, et al. Sensitive and specific colorimetric DNA detection by invasive reaction coupled with nicking endonuclease-assisted nanoparticles amplification. *Biosens Bioelectron.* 2015;66:50–4.
96. Liu YL, Wu HP, Zhou Q, Song QX, Rui JZ, Guan XX, et al. Controllable extension of hairpin-structured flaps to allow low-background cascade invasive reaction for a sensitive DNA logic sensor for mutation detection. *Chem Sci.* 2018;9(6):1666–73.
97. Shen W, Tian Y, Ran T, Gao Z. Genotyping and quantification techniques for single-nucleotide polymorphisms. *Trends Analyt Chem.* 2015;69:1–13.
98. Petralia S, Conoci S. PCR technologies for point of care testing: Progress and perspectives. *ACS Sens.* 2017;2(7):876–91.
99. Matsuda K. PCR-Based Detection Methods for Single-Nucleotide Polymorphism or Mutation: Real-Time PCR and Its Substantial Contribution Toward Technological Refinement. In: Makowski GS, editores. *Advances in Clinical Chemistry*. Amsterdam: Elsevier Inc; 2017. p. 45–72. doi: 10.1016/bs.acc.2016.11.002
100. Wu DY, Ugozzoli L, Pal BK, Wallace RB. Allele-specific enzymatic amplification of  $\beta$ -globin genomic DNA for diagnosis of sickle cell anemia. *Proc Natl Acad Sci U S A.* 1989;86(8):2757–60.
101. Newton CR, Graham A, Heptinstall LE, Powell SJ, Summers C, Kalsheker N, et al. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.* 1989;17(10):3889–97.
102. Sommer SS, Groszbach AR, Bottema CD. PCR amplification of specific alleles (PASA) is a general method for rapidly detecting known single-base changes. *Biotechniques.* 1992;12(1):82–7.
103. Orum H. PCR clamping. *Curr Issues Mol Biol.* 2000;2(1):27–30.
104. Briones C, Moreno M. Applications of peptide nucleic acids (PNAs) and locked nucleic acids (LNAs) in biosensor development. *Anal Bioanal Chem.* 2012;402(10):3071–89.
105. Fouz MF, Appella DH. PNA clamping in nucleic acid amplification protocols to detect single nucleotide mutations related to cancer. *Molecules.* 2020;25(4):786.
106. Milbury CA, Li J, Makrigiorgos GM. PCR-based methods for the enrichment of minority alleles and mutations. *Clin Chem.* 2009;55(4):632–40.
107. Yu M, Cao Y, Ji Y. The principle and application of new PCR Technologies. *IOP Conf Ser Earth Environ Sci.* 2017;100(1):012065. doi: 10.1088/1755-1315/100/1/012065
108. Barbano R, Pasculli B, Coco M, Fontana A, Copetti M, Rendina M, et al. Competitive allele-specific TaqMan PCR (Cast-PCR) is a sensitive, specific and fast method for BRAF V600 mutation detection in Melanoma patients. *Sci Rep.* 2015;5(1):1-11.
109. Vannucchi AM, Pancrazzi A, Bogani C, Antonioli E, Guglielmelli P. A quantitative assay for JAK2V617F mutation in myeloproliferative disorders by ARMS-PCR and capillary electrophoresis. *Leukemia.* 2006;20(6):1055–60.
110. Horiike A, Kimura H, Nishio K, Ohyanagi F, Satoh Y, Okumura S, et al. Detection of epidermal growth factor receptor mutation in transbronchial needle aspirates of non-small cell lung cancer. *Chest.* 2007;131(6):1628–34.
111. Cha RS, Zarbl H, Keohavong P, Thilly WG. Mismatch amplification mutation assay (MAMA): Application to the c-H-ras gene. *Genome Res.* 1992;2(1):14–20.
112. Glaab WE, Skopek TR. A novel assay for allelic discrimination that combines the fluorogenic 5' nuclease polymerase chain reaction (TaqMan®) and mismatch amplification mutation assay. *Mutat Res - Fundam Mol Mech Mutagen.* 1999;430(1):1–12.
113. Liu Q, Sommer SS. Pyrophosphorolysis-activated polymerization (PAP): Application to allele-specific amplification. *Biotechniques.* 2000;29(5):1072–83.
114. Shi J, Liu Q, Sommer SS. Detection of Ultrarare Somatic Mutation in the Human TP53 Gene by Bidirectional Pyrophosphorolysis-Activated Polymerization Allele-Specific Amplification. *Hum Mutat.* 2007;28(2):131–6.
115. Kim HR, Lee SY, Hyun DS, Lee MK, Lee HK, Choi CM, et al. Detection of EGFR mutations in circulating free DNA by PNA-mediated PCR clamping. *J Exp Clin Cancer Res.* 2013;32(1):1-8.
116. Ugozzoli LA, Latorra D, Pucket R, Arar K, Hamby K. Real-time genotyping with oligonucleotide probes containing locked nucleic acids. *Anal Biochem.* 2004;324(1):143–52.
117. Zhang S, Chen Z, Huang C, Ding C, Li C, Chen J, et al. Ultrasensitive and quantitative detection of: EGFR mutations in plasma samples from patients with non-small-cell lung cancer using a dual PNA clamping-mediated LNA-PNA PCR clamp. *Analyst.*

- 2019;144(5):1718–24.
118. Jia Y, Sanchez JA, Wangh LJ. Kinetic hairpin oligonucleotide blockers for selective amplification of rare mutations. *Sci Rep.* 2014;4(1):1–8.
  119. Li J, Wang L, Mamon H, Kulke MH, Berbeco R, Makrigiorgos GM. Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat Med.* 2008;14(5):579–84.
  120. Castellanos-Rizaldos E, Milbury CA, Guha M, Makrigiorgos GM. COLD-PCR Enriches Low-Level Variant DNA Sequences and Increases the Sensitivity of Genetic Testing. In Thurin M, Marincola FM, editores. *Molecular Diagnostics for Melanoma: Methods and Protocols, Methods in Molecular Biology.* Totowa: Humana Press; 2014.p. 623–39. doi: 10.1007/978-1-62703-727-3\_33
  121. Mauger F, How-Kit A, Tost J. COLD-PCR Technologies in the Area of Personalized Medicine: Methodology and Applications. *Mol Diagnosis Ther.* 2017;21(3):269–83.
  122. Milbury CA, Li J, Liu P, Makrigiorgos GM. COLD-PCR: Improving the sensitivity of molecular diagnostics assays. *Expert Rev Mol Diagn.* 2011;11(2):159–69.
  123. Milbury CA, Li J, Makrigiorgos GM. Ice-COLD-PCR enables rapid amplification and robust enrichment for low-abundance unknown DNA mutations. *Nucleic Acids Res.* 2011;39(1):1–10.
  124. Tost J. The clinical potential of Enhanced-ice-COLD-PCR. *Expert Rev Mol Diagn.* 2016;16(3):265–8.
  125. Pumford EA, Lu J, Spaczai I, Prasetyo ME, Zheng EM, Zhang H, et al. Developments in integrating nucleic acid isothermal amplification and detection systems for point-of-care diagnostics. *Biosens Bioelectron.* 2020;170:112674. doi: 10.1016/j.bios.2020.112674
  126. Drain PK, Hyle EP, Noubary F, Freedberg KA, Wilson D, Bishai WR, et al. Diagnostic point-of-care tests in resource-limited settings. *Lancet Infect Dis.* 2014;14(3):239–49.
  127. Horvath P, Barrangou R. CRISPR/Cas, the immune system of Bacteria and Archaea. *Science.* 2010;327(5962):167–70.
  128. Li Y, Li S, Wang J, Liu G. CRISPR/Cas Systems towards Next-Generation Biosensing. *Trends Biotechnol.* 2019;37(7):730–43.
  129. Zuo X, Fan C, Chen HY. Biosensing: CRISPR-powered diagnostics. *Nat Biomed Eng.* 2017;1(6):1–2.
  130. Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science.* 2017;356(6336):438–42.
  131. Myhrvold C, Freije CA, Gootenberg JS, Abudayyeh OO, Metsky HC, Durbin AF, et al. Field-deployable viral diagnostics using CRISPR-Cas13. *Science.* 2018;360(6387):444–8.
  132. Chen JS, Ma E, Harrington LB, Costa M Da, Tian X, Palefsky JM, et al. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science.* 2018;360(6387):436–9.
  133. Gootenberg JS, Abudayyeh OO, Kellner MJ, Joung J, Collins JJ, Zhang F. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a and Csm6. *Science.* 2018;360(6387):439–44.
  134. Ackerman CM, Myhrvold C, Thakku SG, Freije CA, Metsky HC, Yang DK, et al. Massively multiplexed nucleic acid detection with Cas13. *Nature.* 2020;582(7811):277–82.
  135. Lee SH, Park YH, Jin YB, Kim SU, Hur JK. CRISPR Diagnosis and Therapeutics with Single Base Pair Precision. *Trends Mol Med.* 2020;26(3):337–50.
  136. Li J, Macdonald J. Advances in isothermal amplification: Novel strategies inspired by biological processes. *Biosens Bioelectron.* 2015;64:196–211.
  137. Zhao Y, Chen F, Li Q, Wang L, Fan C. Isothermal Amplification of Nucleic Acids. Vol. 115, *Chem Rev.* 2015;115(22):12497–545.
  138. Compton J. Nucleic acid sequence-based amplification. *Nature.* 1991;350:91–2.
  139. Gill P, Ghaemi A. Nucleic acid isothermal amplification technologies - A review. *Nucleosides Nucleotides Nucleic Acids.* 2008;27(3):224–43.
  140. Abdolazadeh A, Dolgosheina E V, Unrau PJ. RNA detection with high specificity and sensitivity using nested fluorogenic Mango NASBA. *RNA.* 2019;25(12):1806–13.
  141. Sun Q, Cao M, Zhang X, Wang M, Ma Y, Wang J. A simple and low-cost paper-based colorimetric method for detecting and distinguishing the GII.4 and GII.17 genotypes of norovirus. *Talanta.* 2021;225:121978. doi: 10.1016/j.talanta.2020.121978
  142. Berard C, Cazalis MA, Leissner P, Mouglin B. DNA nucleic acid sequence-based amplification-based genotyping for polymorphism analysis. *Biotechniques.*

- 2004;37(4):680–6.
143. Mader A, Riehle U, Brandstetter T, Stickeler E, Zur Hausen A, Rühle J. Microarray-based amplification and detection of RNA by nucleic acid sequence based amplification. *Anal Bioanal Chem.* 2010;397(8):3533–41.
  144. Mader A, Riehle U, Brandstetter T, Stickeler E, Rühle J. Universal nucleic acid sequence-based amplification for simultaneous amplification of messengerRNAs and microRNAs. *Anal Chim Acta.* 2012;754:1–7.
  145. Lu X, Shi X, Wu G, Wu T, Qin R, Wang Y. Visual detection and differentiation of Classic Swine Fever Virus strains using nucleic acid sequence-based amplification (NASBA) and G-quadruplex DNAzyme assay. *Sci Rep.* 2017;7(1):1-9.
  146. Walker GT, Little MC, Nadeau JG, Shank DD. Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. *Proc Natl Acad Sci U S A.* 1992;89(1):392–6.
  147. Little MC, Andrews J, Moore R, Bustos S, Jones L, Embres C, et al. Strand displacement amplification and homogeneous real-time detection incorporated in a second-generation DNA probe system, BDProbeTecET. *Clin Chem.* 1999;45(6):777–84.
  148. Wang HQ, Liu WY, Wu Z, Tang LJ, Xu XM, Yu RQ, et al. Homogeneous label-free genotyping of single nucleotide polymorphism using ligation-mediated strand displacement amplification with DNAzyme-based chemiluminescence detection. *Anal Chem.* 2011;83:1883–9.
  149. Shi C, Ge Y, Gu H, Ma C. Highly sensitive chemiluminescent point mutation detection by circular strand-displacement amplification reaction. *Biosens Bioelectron.* 2011;26(12):4697–701.
  150. Gao ZF, Ling Y, Lu L, Chen NY, Luo HQ, Li NB. Detection of single-nucleotide polymorphisms using an ON-OFF switching of regenerated biosensor based on a locked nucleic acid-integrated and toehold-mediated strand displacement reaction. *Anal Chem.* 2014;86(5):2543–8.
  151. Li H, Tang Y, Zhao W, Wu Z, Wang S, Yu R. Palindromic molecular beacon-based intramolecular strand-displacement amplification strategy for ultrasensitive detection of K-ras gene. *Anal Chim Acta.* 2019;1065:98–106.
  152. Wang T, Peng Q, Guo B, Zhang D, Zhao M, Que H, et al. An integrated electrochemical biosensor based on target-triggered strand displacement amplification and “four-way” DNA junction towards ultrasensitive detection of PIK3CA gene mutation. *Biosens Bioelectron.* 2020;150:111954. doi: 10.1016/j.bios.2019.111954
  153. Fire A, Xu SQ. Rolling replication of short DNA circles. *Proc Natl Acad Sci U S A.* 1995;92(10):4641–5.
  154. Banér J, Nilsson M, Mendel-Hartvig M, Landegren U. Signal amplification of padlock probes by rolling circle replication. *Nucleic Acids Res.* 1998;26(22):5073–8.
  155. Qi X, Bakht S, Devos KM, Gale MD, Osbourn A. L-RCA (ligation-rolling circle amplification): a general method for genotyping of single nucleotide polymorphisms (SNPs). *Nucleic Acids Res.* 2001;29(22):e116. doi: 10.1093/nar/29.22.e116
  156. Cheng Y, Zhao J, Jia H, Yuan Z, Li Z. Ligase chain reaction coupled with rolling circle amplification for high sensitivity detection of single nucleotide polymorphisms. *Analyst.* 2013;138(10):2958–63.
  157. Li J, Deng T, Chu X, Yang R, Jiang J, Shen G, et al. Rolling circle amplification combined with gold nanoparticle aggregates for highly sensitive identification of single-nucleotide polymorphisms. *Anal Chem.* 2010;82(7):2811–6.
  158. Ji H, Yan F, Lei J, Ju H. Ultrasensitive electrochemical detection of nucleic acids by template enhanced hybridization followed with rolling circle amplification. *Anal Chem.* 2012;84(16):7166–71.
  159. Dekaliuk M, Qiu X, Troalen F, Busson P, Hildebrandt N. Discrimination of the V600E Mutation in BRAF by Rolling Circle Amplification and Förster Resonance Energy Transfer. *ACS Sens.* 2019;4(10):2786–93.
  160. Zhang L, Zhang Y, Huang L, Zhang Y, Li Y, Ding S, et al. Ultrasensitive biosensing of low abundance BRAF V600E mutation in real samples by coupling dual padlock-gap-ligase chain reaction with hyperbranched rolling circle amplification. *Sens Actuators B Chem.* 2019;287:111–7.
  161. Xu L, Duan J, Chen J, Ding S, Cheng W. Recent advances in rolling circle amplification-based biosensing strategies-A review. *Anal Chim Acta.* 2021;1148:238187.
  162. Dean FB, Nelson JR, Giesler TL, Lasken RS. Rapid amplification of plasmid and phage



- DNA using Phi29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* 2001;11(6):1095–9.
163. Spits C, Le Caignec C, De Rycke M, Van Haute L, Van Steirteghem A, Liebaers I, et al. Whole-genome multiple displacement amplification from single cells. *Nat Protoc.* 2006;1(4):1965–70.
  164. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, et al. Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A.* 2002;99(8):5261–6.
  165. Le Caignec C, Spits C, Sermon K, De Rycke M, Thienpont B, Debrock S, et al. Single-cell chromosomal imbalances detection by array CGH. *Nucleic Acids Res.* 2006;34(9):e68. doi: 10.1093/nar/gkl336
  166. Lovmar L, Fredriksson M, Liljedahl U, Sigurdsson S, Syvänen AC. Quantitative evaluation by minisequencing and microarrays reveals accurate multiplexed SNP genotyping of whole genome amplified DNA. *Nucleic Acids Res.* 2003;31(21):e129. doi: 10.1093/nar/gng129
  167. Raghunathan A, Ferguson HR, Bornarth CJ, Song W, Driscoll M, Lasken RS. Genomic DNA amplification from a single bacterium. *Appl Environ Microbiol.* 2005;71(6):3342–7.
  168. Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, Hase T. Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res.* 2000;28(12):e63. doi: 10.1093/nar/28.12.e63
  169. Notomi T, Mori Y, Tomita N, Kanda H. Loop-mediated isothermal amplification (LAMP): principle, features, and future prospects. *J Microbiol.* 2015;53(1):1–5.
  170. Wong Y, Othman S, Lau Y, Radu S, Chee H, Hui-Yee Chee C, et al. Loop-mediated isothermal amplification (LAMP): a versatile technique for detection of micro-organisms. *J Appl Microbiol.* 2018;124(3):626–43.
  171. Mori Y, Nagamine K, Tomita N, Notomi T. Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. *Biochem Biophys Res Commun.* 2001;289(1):150–4.
  172. Nakamura N, Ito K, Takahashi M, Hashimoto K, Kawamoto M, Yamanaka M, et al. Detection of six single-nucleotide polymorphisms associated with rheumatoid arthritis by a loop-mediated isothermal amplification method and an electrochemical DNA chip. *Anal Chem.* 2007;79(24):9484–93.
  173. Yamanaka ES, Tortajada-Genaro LA, Pastor N, Maquieira Á. Polymorphism genotyping based on loop-mediated isothermal amplification and smartphone detection. *Biosens Bioelectron.* 2018;109:177–83.
  174. Ding S, Chen G, Chen G, Li M, Wang J, Zou J, et al. One-step colorimetric genotyping of single nucleotide polymorphism using probe-enhanced loop-mediated isothermal amplification (PE-LAMP). *Theranostics.* 2019;9(13):3723–31.
  175. Cao G, Kong J, Xing Z, Tang Y, Zhang X, Xu X, et al. Rapid detection of CALR type 1 and type 2 mutations using PNA-LNA clamping loop-mediated isothermal amplification on a CD-like microfluidic chip. *Anal Chim Acta.* 2018;1024:123–35.
  176. Varona M, Eitzmann DR, Pagariya D, Anand RK, Anderson JL. Solid-Phase Microextraction Enables Isolation of BRAF V600E Circulating Tumor DNA from Human Plasma for Detection with a Molecular Beacon Loop-Mediated Isothermal Amplification Assay. *Anal Chem.* 2020;92(4):3346–53.
  177. Du WF, Ge JH, Li JJ, Tang LJ, Yu RQ, Jiang JH. Single-step, high-specificity detection of single nucleotide mutation by primer-activatable loop-mediated isothermal amplification (PA-LAMP). *Anal Chim Acta.* 2019;1050:132–8.
  178. Bao Y, Jiang Y, Xiong E, Tian T, Zhang Z, Lv J, et al. CUT-LAMP: Contamination-Free Loop-Mediated Isothermal Amplification Based on the CRISPR/Cas9 Cleavage. *ACS Sens.* 2020;5(4):1082–91.
  179. Vincent M, Xu Y, Kong H. Helicase-dependent isothermal DNA amplification. *EMBO Rep.* 2004;5(8):795–800.
  180. An L, Tang W, Ranalli TA, Kim HJ, Wytiaz J, Kong H. Characterization of a thermostable UvrD helicase and its participation in helicase-dependent amplification. *J Biol Chem.* 2005;280(32):28952–8.
  181. Li Y, Jortani SA, Ramey-Hartung B, Hudson E, Lemieux B, Kong H. Genotyping three SNPs affecting warfarin drug response by isothermal real-time HDA assays. *Clin Chim Acta.* 2011;412(1–2):79–85.
  182. Sedighi A, Oberc C, Whitehall V, Li PCH. NanoHDA: A nanoparticle-assisted isothermal

- amplification technique for genotyping assays. *Nano Res.* 2017;10(1):12–21.
183. Chen F, Zhao Y, Fan C, Zhao Y. Mismatch Extension of DNA Polymerases and High-Accuracy Single Nucleotide Polymorphism Diagnostics by Gold Nanoparticle-Improved Isothermal Amplification. *Anal Chem.* 2015;87(17):8718–23.
  184. Piepenburg O, Williams CH, Stemple DL, Armes NA. DNA detection using recombination proteins. *PLoS Biol.* 2006;4(7):e204. doi: 10.1371/journal.pbio.0040204
  185. Lobato IM, O'Sullivan CK. Recombinase polymerase amplification: Basics, applications and recent advances. *Trends Analyt Chem.* 2018;98:19–35.
  186. Versatile Isothermal DNA/RNA Amplification by TwistDx [Internet]. [cited 2021 Sep 28]. Available from: <https://www.twistdx.co.uk/>
  187. Mayboroda O, Benito AG, Del Rio JS, Svobodova M, Julich S, Tomaso H, et al. Isothermal solid-phase amplification system for detection of *Yersinia pestis*. *Anal Bioanal Chem.* 2016;408(3):671–6.
  188. Crannell ZA, Rohman B, Richards-Kortum R. Equipment-free incubation of recombinase polymerase amplification reactions using body heat. *PLoS One.* 2014;9(11):1–7.
  189. Krölov K, Frolova J, Tudoran O, Suhorutsenko J, Lehto T, Sibul H, et al. Sensitive and rapid detection of chlamydia trachomatis by recombinase polymerase amplification directly from urine samples. *J Mol Diagnostics.* 2014;16(1):127–35.
  190. Abd El Wahed A, Sanabani SS, Faye O, Pessôa R, Patriota JV, Giorgi RR, et al. Rapid Molecular Detection of Zika Virus in Acute-Phase Urine Samples Using the Recombinase Polymerase Amplification Assay. *PLoS Curr.* 2017;9:1–9.
  191. Qi H, Yue S, Bi S, Ding C, Song W. Isothermal exponential amplification techniques: From basic principles to applications in electrochemical biosensors. *Biosens Bioelectron.* 2018;110:207–17.
  192. Liu Y, Lei T, Liu Z, Kuang Y, Lyu J, Wang Q. A novel technique to detect EGFR mutations in lung cancer. *Int J Mol Sci.* 2016;17(5):792. doi: 10.3390/ijms17050792
  193. del Rio JS, Lobato IM, Mayboroda O, Katakis I, O'Sullivan CK. Enhanced solid-phase recombinase polymerase amplification and electrochemical detection. *Anal Bioanal Chem.* 2017;409(12):3261–9.
  194. Wang J, Koo KM, Wee EJJ, Wang Y, Trau M. A nanoplasmonic label-free surface-enhanced Raman scattering strategy for non-invasive cancer genetic subtyping in patient samples. *Nanoscale.* 2017;9(10):3496–503.
  195. Jiang W, Ren Y, Han X, Xue J, Shan T, Chen Z, et al. Recombinase polymerase amplification-lateral flow (RPA-LF) assay combined with immunomagnetic separation for rapid visual detection of *Vibrio parahaemolyticus* in raw oysters. *Anal Bioanal Chem.* 2020;412(12):2903–14.
  196. Silva G, Bömer M, Nkere C, Lava Kumar P, Seal SE. Rapid and specific detection of Yam mosaic virus by reverse-transcription recombinase polymerase amplification. *J Virol Methods.* 2015;222:138–44.
  197. Tortajada-Genaro LA, Santiago-Felipe S, Amasia M, Russom A, Maquieira A. Isothermal solid-phase recombinase polymerase amplification on microfluidic digital versatile discs (DVDs). *RSC Adv.* 2015;5(38):29987–95.
  198. Qi H, Yue S, Bi S, Ding C, Song W. Isothermal exponential amplification techniques: From basic principles to applications in electrochemical biosensors. *Biosens Bioelectron.* 2018;110:207–17.
  199. Koo KM, Dey S, Trau M. A Sample-to-Targeted Gene Analysis Biochip for Nanofluidic Manipulation of Solid-Phase Circulating Tumor Nucleic Acid Amplification in Liquid Biopsies. *ACS Sens.* 2018;3(12):2597–603.
  200. Yamanaka ES, Tortajada-Genaro LA, Maquieira Á. Low-cost genotyping method based on allele-specific recombinase polymerase amplification and colorimetric microarray detection. *Microchim Acta.* 2017;184(5):1453–62.
  201. Martorell S, Palanca S, Maquieira Á, Tortajada-Genaro LA. Blocked recombinase polymerase amplification for mutation analysis of PIK3CA gene. *Anal Biochem.* 2018;544:49–56.
  202. Chen G, Chen R, Ding S, Li M, Wang J, Zou J, et al. Recombinase assisted loop-mediated isothermal DNA amplification. *Analyst.* 2020;145(2):440–4.
  203. Dirks RM, Pierce NA. Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci U S A.* 2004;101(43):15275–8.
  204. Augspurger EE, Rana M, Yigit M V. Chemical and Biological Sensing Using Hybridization Chain Reaction. *ACS Sens.* 2018;3(5):878–902.

205. Bi S, Yue S, Zhang S. Hybridization chain reaction: A versatile molecular tool for biosensing, bioimaging, and biomedicine. *Chem Soc Rev.* 2017;46(14):4281–98.
206. Xuan F, Hsing IM. Triggering hairpin-free chain-branching growth of fluorescent DNA dendrimers for nonlinear hybridization chain reaction. *J Am Chem Soc.* 2014;136(28):9810–3.
207. Ikbal J, Lim GS, Gao Z. The hybridization chain reaction in the development of ultrasensitive nucleic acid assays. *Trends Analyt Chem.* 2015;64:86–99.
208. Huang J, Wu Y, Chen Y, Zhu Z, Yang X, Yang CJ, et al. Pyrene-excimer probes based on the hybridization chain reaction for the detection of nucleic acids in complex biological fluids. *Angew Chem Int Ed Engl.* 2011;50(2):401–4.
209. Wang Y, Jiang L, Leng Q, Wu Y, He X, Wang K. Electrochemical sensor for glutathione detection based on mercury ion triggered hybridization chain reaction signal amplification. *Biosens Bioelectron.* 2016;77:914–20.
210. Song Y, Wei W, Qu X. Colorimetric biosensing using smart materials. *Adv Mater.* 2011;23(37):4215–36.
211. Yang L, Liu C, Ren W, Li Z. Graphene Surface-Anchored Fluorescence Sensor for Sensitive Detection of MicroRNA Coupled with Enzyme-Free Signal Amplification of Hybridization Chain Reaction. *ACS Appl Mater Interfaces.* 2012;4(12):6450–3.
212. Dong Q, Liu Q, Guo L, Li D, Shang X, Li B, et al. A signal-flexible gene diagnostic strategy coupling loop-mediated isothermal amplification with hybridization chain reaction. *Anal Chim Acta.* 2019;1079:171–9.
213. Miao P, Tang Y, Yin J. MicroRNA detection based on analyte triggered nanoparticle localization on a tetrahedral DNA modified electrode followed by hybridization chain reaction dual amplification. *Chem Commun.* 2015;51(86):15629–32.
214. Li X, Wang Y, Wang L, Wei Q. A surface plasmon resonance assay coupled with a hybridization chain reaction for amplified detection of DNA and small molecules. *Chem Commun.* 2014;50(39):5049–52.
215. Dong J, Cui X, Deng Y, Tang Z. Amplified detection of nucleic acid by G-quadruplex based hybridization chain reaction. *Biosens Bioelectron.* 2012;38(1):258–63.
216. Bi S, Chen M, Jia X, Dong Y, Wang Z. Hyperbranched Hybridization Chain Reaction for Triggered Signal Amplification and Concatenated Logic Circuits. *Angew Chem Int Ed.* 2015;54(28):8144–8.
217. Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, Palefsky JM, et al. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science (80- ).* 2018;360(6387):436–9.
218. Song C, Li B, Yang X, Wang K, Wang Q, Liu J, et al. Use of  $\beta$ -cyclodextrin-tethered cationic polymer based fluorescence enhancement of pyrene and hybridization chain reaction for the enzyme-free amplified detection of DNA. *Analyst.* 2017;142(1):224–8.
219. Qiu X, Wang P, Cao Z. Hybridization chain reaction modulated DNA-hosted silver nanoclusters for fluorescent identification of single nucleotide polymorphisms in the let-7 miRNA family. *Biosens Bioelectron.* 2014;60:351–7.
220. Song L, Zhang Y, Li J, Gao Q, Qi H, Zhang C. Non-covalent fluorescent labeling of hairpin DNA probe coupled with hybridization chain reaction for sensitive DNA detection. *Appl Spectrosc.* 2016;70(4):688–94.
221. Oishi M. Enzyme-free and isothermal detection of microRNA based on click-chemical ligation-assisted hybridization coupled with hybridization chain reaction signal amplification. *Anal Bioanal Chem.* 2015;407(14):4165–72.
222. Niu S, Jiang Y, Zhang S. Fluorescence detection for DNA using hybridization chain reaction with enzyme-amplification. *Chem Commun.* 2010;46(18):3089–91.
223. Xu Y, Zheng Z. Direct RNA detection without nucleic acid purification and PCR: Combining sandwich hybridization with signal amplification based on branched hybridization chain reaction. *Biosens Bioelectron.* 2016;79:593–9.
224. Yin F, Liu H, Li Q, Gao X, Yin Y, Liu D. Trace MicroRNA Quantification by Means of Plasmon-Enhanced Hybridization Chain Reaction. *Anal Chem.* 2016;88(9):4600–4.
225. Schwarzkopf M, Pierce NA. Multiplexed miRNA northern blots via hybridization chain reaction. *Nucleic Acids Res.* 2016;44(15):e129. doi: 10.1093/nar/gkw503
226. Wei Y, Zhou W, Li X, Chai Y, Yuan R, Xiang Y. Coupling hybridization chain reaction with catalytic hairpin assembly enables non-enzymatic and sensitive fluorescent detection of microRNA cancer biomarkers. *Biosens Bioelectron.* 2016;77:416–20.
227. Chandran H, Rangnekar A, Shetty G, Schultes EA, Reif JH, Labean TH. An autonomously

- self-assembling dendritic DNA nanostructure for target DNA detection. *Biotechnol J.* 2013;8(2):221–7.
228. Guo J, Mingoos C, Qiu X, Hildebrandt N. Simple, Amplified, and Multiplexed Detection of MicroRNAs Using Time-Gated FRET and Hybridization Chain Reaction. *Anal Chem.* 2019;91(4):3101–9.
229. Ge Z, Lin M, Wang P, Pei H, Yan J, Shi J, et al. Hybridization chain reaction amplification of microRNA detection with a tetrahedral DNA nanostructure-based electrochemical biosensor. *Anal Chem.* 2014;86(4):2124–30.
230. Hou T, Li W, Liu X, Li F. Label-Free and Enzyme-Free Homogeneous Electrochemical Biosensing Strategy Based on Hybridization Chain Reaction: A Facile, Sensitive, and Highly Specific MicroRNA Assay. *Anal Chem.* 2015;87(22):11368–74.
231. Yu Y, Chen Z, Jian W, Sun D, Zhang B, Li X, et al. Ultrasensitive electrochemical detection of avian influenza A (H7N9) virus DNA based on isothermal exponential amplification coupled with hybridization chain reaction of DNAzyme nanowires. *Biosens Bioelectron.* 2015;64:566–71.
232. Yang H, Gao Y, Wang S, Qin Y, Xu L, Jin D, et al. In situ hybridization chain reaction mediated ultrasensitive enzyme-free and conjugation-free electrochemical genosensor for BRCA-1 gene in complex matrices. *Biosens Bioelectron.* 2016;80:450–5.
233. Xuan F, Fan TW, Hsing IM. Electrochemical interrogation of kinetically-controlled dendritic DNA/PNA assembly for immobilization-free and enzyme-free nucleic acids sensing. *ACS Nano.* 2015;9(5):5027–33.
234. Gu C, Kong X, Liu X, Gai P, Li F. Enzymatic Biofuel-Cell-Based Self-Powered Biosensor Integrated with DNA Amplification Strategy for Ultrasensitive Detection of Single-Nucleotide Polymorphism. *Anal Chem.* 2019;66:37–9.
235. Wang X, Ge L, Yu Y, Dong S, Li F. Highly sensitive electrogenerated chemiluminescence biosensor based on hybridization chain reaction and amplification of gold nanoparticles for DNA detection. *Sens Actuators B Chem.* 2015;220:942–8.
236. Chen Y, Xu J, Su J, Xiang Y, Yuan R, Chai Y. In Situ Hybridization Chain Reaction Amplification for Universal and Highly Sensitive Electrochemiluminescent Detection of DNA. *Anal Chem.* 2012;84(18):7750–5.
237. Bi S, Zhao T, Luo B, Zhu J-J. Hybridization chain reaction-based branched rolling circle amplification for chemiluminescence detection of DNA methylation. *Chem Commun.* 2013;49(61):6906–8.
238. Rana M, Balcioglu M, Kovach M, Hizir MS, Robertson NM, Khan I, et al. Reprogrammable multiplexed detection of circulating oncomiRs using hybridization chain reaction. *Chem Commun.* 2016;52(17):3524–7.
239. Ma C, Wang W, Li Z, Cao L, Wang Q. Simple colorimetric DNA detection based on hairpin assembly reaction and target-catalytic circuits for signal amplification. *Anal Biochem.* 2012;429(2):99–102.
240. Liu P, Yang X, Sun S, Wang Q, Wang K, Huang J, et al. Enzyme-free colorimetric detection of DNA by using gold nanoparticles and hybridization chain reaction amplification. *Anal Chem.* 2013;85(16):7689–95.
241. Yang X, Yu Y, Gao Z. A highly sensitive plasmonic DNA assay based on triangular silver nanoprism etching. *ACS Nano.* 2014;8(5):4902–7.
242. Lu S, Hu T, Wang S, Sun J, Yang X. Ultra-Sensitive Colorimetric Assay System Based on the Hybridization Chain Reaction-Triggered Enzyme Cascade Amplification. *ACS Appl Mater Interfaces.* 2017;9(1):167–75.
243. Shimron S, Wang F, Orbach R, Willner I. Amplified detection of DNA through the enzyme-free autonomous assembly of hemin/G-quadruplex DNAzyme nanowires. *Anal Chem.* 2012;84(2):1042–8.
244. Li R, Zou L, Luo Y, Zhang M, Ling L. Ultrasensitive colorimetric detection of circulating tumor DNA using hybridization chain reaction and the pivot of triplex DNA. *Sci Rep.* 2017;7(1):1–10.
245. Zeng C, Lu N, Wen Y, Liu G, Zhang R, Zhang J, et al. Engineering Nanozymes Using DNA for Catalytic Regulation. *ACS Appl Mater Interfaces.* 2019;11(2):1790–9.
246. Park CR, Rhee WJ, Kim KW, Hwang BH. Colorimetric biosensor using dual-amplification of enzyme-free reaction through universal hybridization chain reaction system. *Biotechnol Bioeng.* 2019;116(7):1567–74.
247. Khansili N, Rattu G, Krishna PM. Label-free optical biosensors for food and biological sensor applications. *Sens Actuators B Chem.* 2018;265:35–49.

248. Ravan H, Kashanian S, Sanadgol N, Badoei-Dalfard A, Karami Z. Strategies for optimizing DNA hybridization on surfaces. *Anal Biochem.* 2014;444(1):41–6.
249. Jayanthi VSA, Das AB, Saxena U. Recent advances in biosensor development for the detection of cancer biomarkers. *Biosens Bioelectron.* 2017;91:15–23.
250. Sadighbayan D, Sadighbayan K, Tohid-kia MR, Khosroushahi AY, Hasanzadeh M. Development of electrochemical biosensors for tumor marker determination towards cancer diagnosis: Recent progress. *Trends Analyt Chem.* 2019;118:73–88.
251. Chen C, Wang J. Optical biosensors: An exhaustive and comprehensive review. *Analyst.* 2020;145(5):1605–28.
252. Kim J, Lee J, Kwon D, Lee H, Grailhe R. A comparative analysis of resonance energy transfer methods for Alzheimer related protein-protein interactions in living cells. *Mol Biosyst.* 2011;7(11):2991–6.
253. Engelen W, Van de Wiel KM, Meijer LHH, Saha B, Merckx M. Nucleic acid detection using BRET-beacons based on bioluminescent protein-DNA hybrids. *Chem Commun.* 2017;53(19):2862–5.
254. Baba Y, Yamamoto K, Yoshida W. Multicolor bioluminescence resonance energy transfer assay for quantification of global DNA methylation. *Anal Bioanal Chem.* 2019;411(19):4765–73.
255. Song Y, Wei W, Qu X. Colorimetric Biosensing Using Smart Materials. *Adv Mater.* 2011;23(37):4215–36.
256. Zhu H, Fohlerová Z, Pekárek J, Basova E, Neužil P. Recent advances in lab-on-a-chip technologies for viral diagnosis. *Biosens Bioelectron.* 2020;153:112041. doi: 10.1016/j.bios.2020.112041
257. Masson JF. Surface Plasmon Resonance Clinical Biosensors for Medical Diagnostics. *ACS Sens.* 2017;2(1):16–30.
258. Cui L, Ke G, Zhang WY, Yang CJ. A universal platform for sensitive and selective colorimetric DNA detection based on Exo III assisted signal amplification. *Biosens Bioelectron.* 2011;26(5):2796–800.
259. Jung YL, Jung C, Parab H, Cho DY, Park HG. Colorimetric SNP Genotyping Based on Allele-Specific PCR by Using a Thiol-Labeled Primer. *Chembiochem.* 2011;12(9):1387–90.
260. Li J, Jiang JH, Xu XM, Chu X, Jiang C, Shen G, et al. Simultaneous identification of point mutations via DNA ligase-mediated gold nanoparticle assembly. *Analyst.* 2008;133(7):939–45.
261. Zarei M. Portable biosensing devices for point-of-care diagnostics: Recent developments and applications. *Trends Analyt Chem.* 2017;91:26–41.
262. Morais S, Puchades R, Maquieira Á. Disc-based microarrays: principles and analytical applications. *Anal Bioanal Chem.* 2016;408(17):4523–34.
263. Morais S, Tortajada-Genaro L, Maquieira Á. Array-on-a-disk? How Blu-ray technology can be applied to molecular diagnostics. *Expert Rev Mol Diagn.* 2014;14(7):773–5.
264. Kumagai O, Ikeda M, Yamamoto M. Application of laser diodes to optical disk systems: The history of laser diode development and mass production in three generations of optical disk systems. *Proc IEEE.* 2013;101(10):2243–54.
265. Kido H, Maquieira A, Hammock BD. Disc-based immunoassay microarrays. *Anal Chim Acta.* 2000;411(1–2):1–11.
266. Gorkin R, Park J, Siegrist J, Amasia M, Lee BS, Park JM, et al. Centrifugal microfluidics for biomedical applications. *Lab Chip.* 2010;10(14):1758–73.
267. Wang L, Li PCH. Microfluidic DNA microarray analysis: A review. *Anal Chim Acta.* 2011;687(1):12–27.
268. Li PC, Sedighi A. High-Throughput DNA Array for SNP Detection of KRAS Gene Using a Centrifugal Microfluidic Device. In Li P, Sedighi A, Wang L, editores. *Microarray Technology.* New York: Humana Press; 2016.p. 99-109. doi: 10.1007/978-1-4939-3136-1\_10
269. Strohmeier O, Laßmann S, Riedel B, Mark D, Roth G, Werner M, et al. Multiplex genotyping of KRAS point mutations in tumor cell DNA by allele-specific real-time PCR on a centrifugal microfluidic disk segment. *Microchim Acta.* 2014;181(13–14):1681–8.
270. Li X, Weng S, Ge B, Yao Z, Yu HZ. DVD technology-based molecular diagnosis platform: Quantitative pregnancy test on a disc. *Lab Chip.* 2014;14(10):1686–94.
271. Hwu EET, Boisen A. Hacking CD/DVD/Blu-ray for biosensing. *ACS Sens.* 2018;3(7):1222–32.

272. Li Y, Ou LML, Yu HZ. Digitized molecular diagnostics: Reading disk-based bioassays with standard computer drives. *Anal Chem.* 2008;80(21):8216–23.
273. Morais S, Puchades R, Arnandis-Chover T, González-Martínez MÁ, Maquieira Á. High density microarrays on Blu-ray discs for massive screening. *Biosens Bioelectron.* 2014;51:109–14.
274. Santiago-Felipe S, Tortajada-Genaro LA, Morais S, Puchades R, Maquieira Á. Analytical Methods Isothermal DNA amplification strategies for duplex microorganism detection. *Food Chem.* 2015;174:509–15.
275. Tortajada-Genaro LA, Niñoles R, Mena S, Maquieira Á. Digital versatile discs as platforms for multiplexed genotyping based on selective ligation and universal microarray detection. *Analyst.* 2019;144(2):707–15.
276. Ozcan A. Mobile phones democratize and cultivate next-generation imaging, diagnostics and measurement tools. *Lab Chip.* 2014;14(17):3187–94.
277. Roda A, Michelini E, Zangheri M, Di Fusco M, Calabria D, Simoni P. Smartphone-based biosensors: A critical review and perspectives. *Trends Analyt Chem.* 2016;79:317–25.
278. Liao SC, Peng J, Mauk MG, Awasthi S, Song J, Friedman H, et al. Smart cup: A minimally-instrumented, smartphone-based point-of-care molecular diagnostic device. *Sens Actuators B Chem.* 2016;229:232–8.
279. Sayad A, Ibrahim F, Mukim Uddin S, Cho J, Madou M, Thong KL. A microdevice for rapid, monoplex and colorimetric detection of foodborne pathogens using a centrifugal microfluidic platform. *Biosens Bioelectron.* 2018;100:96–104.
280. Spyrou EM, Kalogianni DP, Tragoulias SS, Ioannou PC, Christopoulos TK. Digital camera and smartphone as detectors in paper-based chemiluminometric genotyping of single nucleotide polymorphisms. *Anal Bioanal Chem.* 2016;408(26):7393–402.
281. Noor MO, Krull UJ. Camera-based ratiometric fluorescence transduction of nucleic acid hybridization with reagentless signal amplification on a paper-based platform using immobilized quantum dots as donors. *Anal Chem.* 2014;86(20):10331–9.
282. Michikawa Y, Suga T, Ohtsuka Y, Matsumoto I, Ishikawa A, Ishikawa K, et al. Visible genotype sensor array. *Sensors.* 2008;8(4):2722–35.
283. Zhang X, Yang L, Wang F, Liu Z, Liu R, Ying Q, et al. Development of a simple and cost-effective method based on T7 endonuclease cleavage for detection of single nucleotide polymorphisms. *Genet Test Mol Biomarkers.* 2018;22(12):719–23.
284. Wolfe MG, Ali MM, Brennan JD. Enzymatic Litmus Test for Selective Colorimetric Detection of C-C Single Nucleotide Polymorphisms. *Anal Chem.* 2019;91(7):4735–40.
285. Yamanaka ES, Tortajada-Genaro LA, Pastor N, Maquieira Á. Polymorphism genotyping based on loop-mediated isothermal amplification and smartphone detection. *Biosens Bioelectron.* 2018;109:177–83.
286. Yu H, Tan Y, Cunningham BT. Smartphone fluorescence spectroscopy. *Anal Chem.* 2014;86(17):8805–13.
287. Sherwood JL, Müller S, Orr MCM, Ratcliffe MJ, Walker J. Panel Based MALDI-TOF Tumour Profiling Is a Sensitive Method for Detecting Mutations in Clinical Non Small Cell Lung Cancer Tumour. *PLoS One.* 2014;9(6):e100566. doi: 10.1371/journal.pone.0100566
288. Norman RL, Singh R, Langridge JI, Ng LL, Jones DJL. The measurement of KRAS G12 mutants using multiplexed selected reaction monitoring and ion mobility mass spectrometry. *Rapid Commun Mass Spectrom.* 2020;34(S4):e8657. doi: 10.1002/rcm.8657
289. Penson RT, Sales E, Sullivan L, Borger DR, Krasner CN, Goodman AK, et al. A SNaPshot of potentially personalized care: Molecular diagnostics in gynecologic cancer. *Gynecol Oncol.* 2016;141(1):108–12.

## **2. OBJECTIVES**





The main **objective** of this doctoral thesis is to develop new biorecognition tools with competitive analytical performance for the identification and quantification of clinically relevant genetic biomarkers. These approaches were materialized as discriminatory bioanalytical strategies combined with optical detection systems satisfying the demands of clinical laboratories focused on the identification of minority alleles from complex human samples.

In order to achieve this goal, a series of **partial objectives** are proposed.

1. To design and develop alternative biorecognition events based on allele-specific reactions and hybridization assays that distinguish and determine target genetic variants, including SNV in low proportions.
2. To integrate biorecognition strategies in versatile analytical platforms with simple optical transduction modes (colorimetry and emission), enabling the detection with consumer electronic devices and its massive application considering POC requirements.
3. To evaluate the analytical performance of the developed methods by genotyping different genetic biomarkers associated with colorectal cancer and pharmacogenetic treatment from clinical patient samples.
4. To compare them with reference methods and develop solutions that integrate the successes of the above milestones to facilitate the implementation of personalized medicine.



# **3. EXPERIMENTAL SECTION AND RESULTS**





## **CHAPTER 1:**

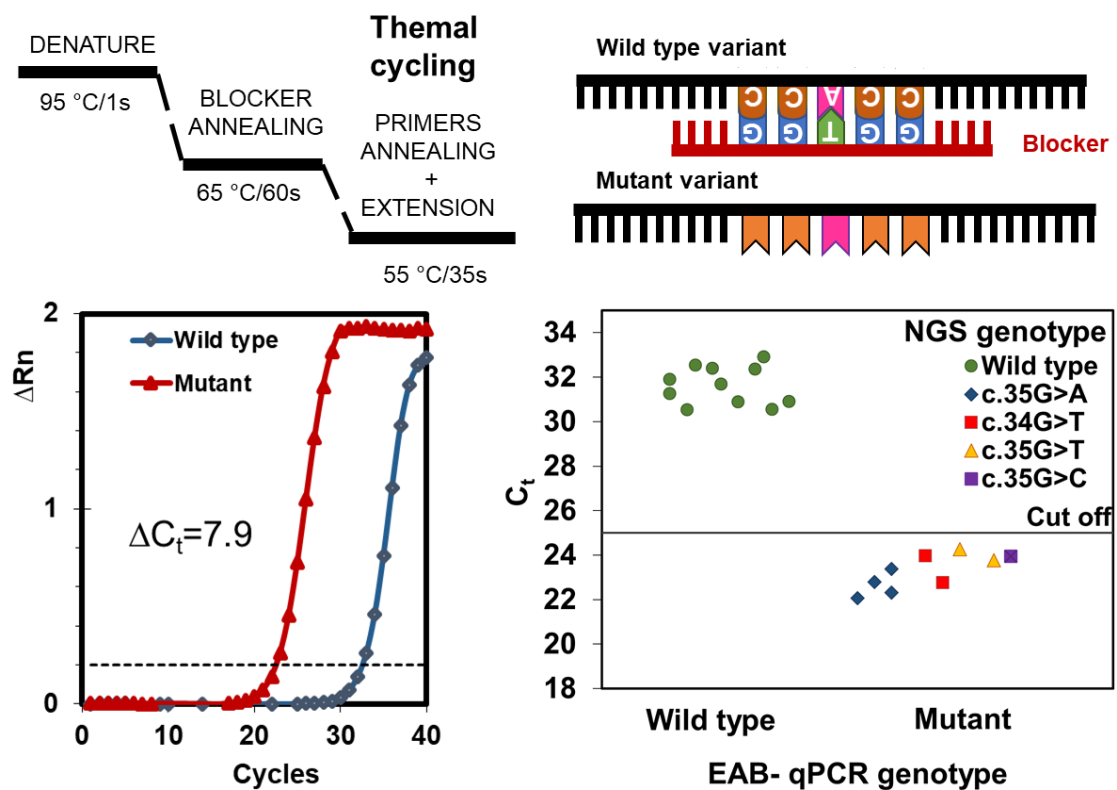
**Enhanced asymmetric blocked  
qPCR method for affordable  
detection of point mutations in  
*KRAS* oncogene**



### 3.1 CHAPTER 1: ENHANCED ASYMMETRIC BLOCKED qPCR METHOD FOR AFFORDABLE DETECTION OF POINT MUTATIONS IN *KRAS* ONCOGENE

In this chapter, we use a standard diagnostic tool in clinical laboratories, such as a thermal cycler, to perform a blocked amplification in order to improve the enrichment of minority mutant alleles. The novelty is the combination of asymmetric PCR with a specific blocking agent of the native variants and the addition of a new thermocycling step before primer annealing. The challenge is to achieve sensitive real-time detection of clinically relevant single nucleotide mutations, overcoming molecular similarity and low concentrations of rare mutations.

In this study, the assay principle is demonstrated and the best reaction conditions to favor the formation of the blocker/native variant complex versus primer/native variant are identified. Thus, the primer annealing, and polymerase extension step selectively trigger the amplification of mutant variants against the native variants. The monitoring of product formation is performed by real-time fluorescence, enabling a sensitive detection. As a proof of concept, it is applied to the mutational analysis of the *KRAS* gene (codons 12-13) in metastatic colorectal carcinoma from biopsied tissues, and the results are compared with two reference methods (conventional blocked qPCR and NGS).



Summary figure. Overview chapter 1



## Enhanced asymmetric blocked qPCR method for affordable detection of point mutations in KRAS oncogene

Ana Lázaro<sup>1</sup> · Luis A. Tortajada-Genaro<sup>1,2</sup> · Ángel Maquieira<sup>1,2</sup>

### 3.1.1 Abstract

An accurate genetic diagnostic is key for adequate patient management and the suitability of healthcare systems. The scientific challenge lies in developing methods to discriminate those patients with certain genetic variations present in tumor cells at low-concentrations. We report a method called enhanced asymmetric blocked qPCR (EAB-qPCR) that promotes the blocker annealing against the primer-template hybrid controlling thermal cycling and reaction conditions with nonmodified oligonucleotides. Real-time fluorescent amplification curves of wild-type alleles were delayed by about eight cycles for EAB-qPCR, compared to conventional blocked qPCR approaches. This method reduced the amplification of native DNA variants (blocking percentage 99.7%) and enabled the effective enrichment of low-level DNA mutations. Excellent performance was estimated for the detection of mutated alleles in sensitivity (up to 0.5% mutant/total DNA) and reproducibility terms, with a relative standard deviation below 2.8%. The method was successfully applied to the mutational analysis of metastatic colorectal carcinoma from biopsied tissues. The determined single-nucleotide mutations in the *KRAS* oncogene (codon 12-13) totally agreed with those obtained from next-generation sequencing. EAB-qPCR is an accurate cheap method and can be easily incorporated into daily routine to detect mutant alleles. Hence these features are especially interesting to facilitate the diagnosis and prognosis of several clinical diseases.

**Keywords:** bioanalytical methods; allele-selective qPCR; *KRAS* oncogene; mutation genotyping; DNA variant detection.

### 3.1.2 Introduction

In the precision medicine era, the detection of minority alleles is crucial because it may affect clinical decisions in the fields of cancer, prenatal diagnosis, or infectious diseases [1, 2]. In fact the ability to distinguish single-nucleotide mutations is becoming essential for selecting correct treatment according to patients' individual characteristics



[3]. However given the heterogeneous nature of tumors, the mutated DNA from cancer cells must be detected when non mutated DNA from normal cells are abundant and present [4]. One relevant example is the genotyping of mutations in the *KRAS* oncogene, before the treatment based on monoclonal antibodies such as cetuximab and panitumumab. Wild-type patients better respond to antibody-based therapeutic medicines and have higher survival rates [5, 6].

The detection of mutated variants when excess wild-type DNA is present requires high-performance assays. Thus routine applications in diagnostics require accurate, selective, easy-to-implement and cost-effective techniques [7]. To date, the most useful methods for detecting single-nucleotide mutations can be classified into two categories; sequencing methods and minority allele enrichment strategies [8]. The main advantage of sequencing methods is they identify the specific mutation, although Sanger sequencing shows limited sensitivity, a high contamination risk and low throughput [9]. Likewise, the expense associated with pyrosequencing and next-generation sequencing (NGS) techniques is currently high for instruments (up to €10<sup>5</sup>) and for running costs (up to €10<sup>3</sup>) [7]. In several clinical scenarios, PCR methods for enriching minority alleles are the key alternative [10, 11]. The first approaches were allele-specific PCR [12], amplification refractory mutation system PCR (ARMS) [13] and restriction fragment length polymorphism PCR (RFLP) [14]. Droplet digital PCR (ddPCR) offers very high selectivity (10<sup>-3</sup> to 10<sup>-8</sup>), although ddPCR is still time-consuming and expensive (instrument up to €10<sup>5</sup> and up to €20 per sample) [15].

In recent years, considerable research has focused on methods supported by standard qPCR equipment because it is frequently available in laboratories thanks to its robustness, affordable price and general reagents. In this category, an interesting method is co-amplification that operates at lower denaturation temperature PCR (COLD-PCR) [16]. Despite its high sensitivity, mutation enrichment depends on the sequence context and, thus, certain mutations in a DNA sequence may be more difficult to detect than others [17]. Another strategy is based on the improvement of the blocked qPCR method by incorporating modified oligonucleotides, such as peptide nucleic acid (PNA), locked nucleic acid (LNA) and LNA/DNA chimeras [18, 19]. Their function selectively inhibits the amplification of wild-type sequences to produce a selective hybrid between the target and the blocker [20, 21]. However, these modifications are expensive.

Modified PCR methods has been described to avoid the plateau phase of PCR and to improve amplification specificity [22]. A relevant method is the linear one after exponential PCR (LATE-PCR) because the preferential enrichment of mutant sequences is achieved applying a specific reaction sequence [23]. The stages include a linear pre-amplification (4 steps, 10 cycles), the conversion of single-strand DNA into double-strand

DNA (4 steps, a few cycles) and the exponential amplification of mutants (4 steps, 50 cycles). In each stage, the preferential hybridization of the blocker onto the wild-type template strand is improved because the reaction is paused at the optimal annealing temperature of the blocker. The main drawbacks are the large number of amplification cycles and the stringent working conditions required to open the stem of the blocker. Nowadays novel assay principles to improve the reliability of PCR-based methods are still necessary for extensive use [24].

Herein a novel approach, called enhanced asymmetric blocked qPCR (EAB-qPCR), is reported. The combination of asymmetric PCR with a specific blocking agent and the addition of a new thermal cycling stage enriches minority DNA variants. Blocker annealing is favored by minimizing nonspecific recognition and maximizing the inhibition of perfect-match amplification. In this way, EAB-qPCR was designed for the discrimination of one base pair mismatch to enable the detection of mutant variants.

### **3.1.3 Materials and Methods**

**Target.** Single-nucleotide mutations in the *KRAS* gene (codons 12-13) were selected as the model given their high prevalence and clinical significance [6]. Nucleotide sequences were obtained from the National Biotechnology Information Center database (NCBI Gene 3845). Specific primers and blockers were designed for the EAB-qPCR method, as described in Supplementary Information (Table SI.1 and SI.2). All the oligonucleotides, purified by HPLC, were ordered from Eurofins Genomics (Germany).

**Cell lines and patients.** Human SK-N-AS cells with a wild-type variant for the target region and HCT116 cells with mutant c.38G>A (*KRAS* p.G13D) were purchased from the American Type Culture Collection (ATCC, USA) and were used for method optimization purposes. Formalin-fixed paraffin-embedded (FFPE) biopsy tissues were obtained from the Oncological Service of the Hospital Clínico Universitario La Fe (Spain). Carcinomas were sampled in the infiltrating area of the growth, avoiding the necrotic center. Tissues corresponded to 20 patients with metastatic colorectal cancer who had been pathologically confirmed. Samples were fixed in less than 24 h and stored at 4°C until DNA extraction. All the experimental protocols were conducted according to the ethics and the Declaration of Helsinki, including informed consents obtained from each patient.

**DNA extraction.** The genomic DNA of the cell lines was extracted using the PureLink Genomic DNA kit (Invitrogen, USA). For the genomics of the metastatic colorectal cancer samples, extraction was performed by the QIAamp DNA Investigator kit (Qiagen, Germany). The quality and concentration of the extracted DNA (ng/μL) were

determined by spectrophotometry (Nanodrop 2000, Thermo Fisher Scientific, USA). Extracts were stored at -20 °C until processed.

**EAB-qPCR method.** Reactions were performed in a total volume of 12.5  $\mu$ L containing 1x TB Green Premix Ex Taq (Takara, Gallini, Spain), 1x ROX reference dye II (Takara, Gallini, Spain), 300 nM of the forward primer, 150 nM of the reverse primer, 150 nM of the blocker agent and 1  $\mu$ L of each DNA extract (4 ng/ $\mu$ L, equivalent to 1300 copies). The reagents were loaded in 96-well microplates (Axygen PCR, Fischer Scientific, Spain), covered with ultra-pressure sealing film (Thermo Fisher Scientific, USA). Amplification and detection were carried out by the ViiA 7 Real-Time PCR System instrument (Applied Biosystem, USA). Thermal cycling was: 2 min at 50 °C, 10 min at 95 °C, followed by 40 cycles of amplification of 1 s at 95 °C (denaturation), 60 s at 65 °C (blocker annealing) and 35 s at 55 °C (primer annealing and extension, fluorescence acquisition). Reactions were run in duplicate and the experiment included one negative control and no template control. Optionally, a melting curve analysis was acquired from 60 °C to 95 °C at a thermal transition rate of 0.5 °C per second.

**Data interpretation.** The data were analyzed with software included in a ViiA 7 Real-Time PCR System. The  $C_q$  value, defined as the cycle number at which a significant increase in fluorescence is detected, was recorded. The detection threshold was set at  $\Delta R_n = 0.2$ , calculated from the signal increment between both fluorophores. For genotyping, a discrimination factor was calculated as the delay of the wild-type amplification in relation to mutant amplification ( $\Delta C_q = C_{q,wild-type} - C_{q,mutant}$ ). For the discrimination of patients, a logic gate was defined on the basis of quantification  $C_q$  ( $C_{q, cut-off}$ ). Samples were declared as mutants or wild type if the measured  $C_q$  was lower or higher than 25, respectively.

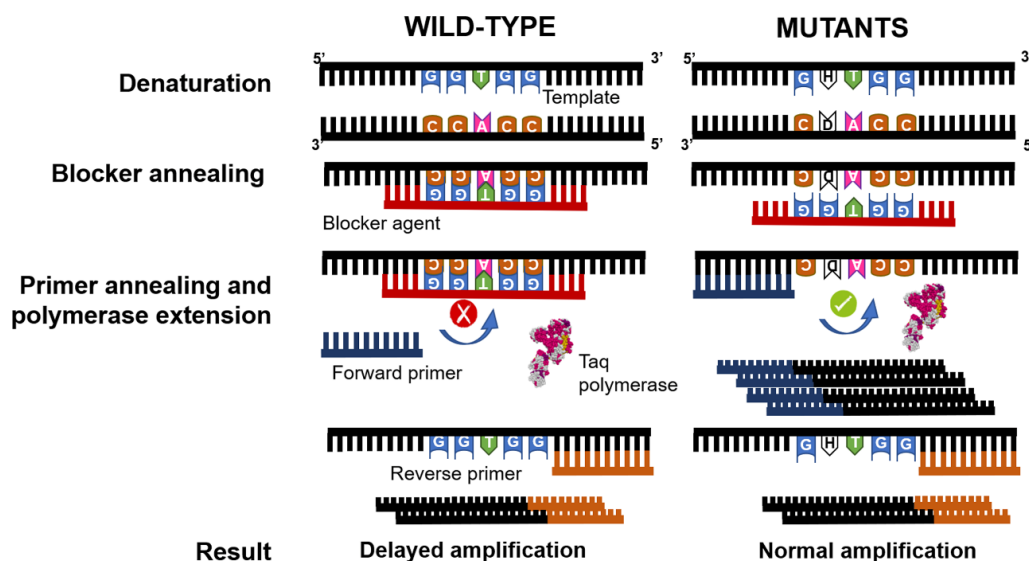
**Reference methods.** Conventional blocked qPCR and Next Generation Sequencing were also applied to tumor samples, see Supplementary Material.

### 3.1.4 Results

#### Principle of selective enrichment EAB-qPCR.

Figure 1 presents the scheme of the EAB-qPCR mechanism, enabling the enrichment of minority alleles, including the discrimination of mutant variants, and even alteration only involves a single-nucleotide change. The method can be considered an enhanced variant of blocked-qPCR based on promoting the wild-type template/blocker hybrid against the template/primer hybrid. After DNA denaturation occurs, an intermediate step is included for the selective annealing of the blocker to the wild-type template. With the correct selection of reaction conditions, the base pair mismatch between the blocker and mutant DNA suffices to prevent the formation of the mutant

template/blocker hybrid. In the next reaction step, the annealing of the primer to the template is targeted. The primer elongation of the blocked sequence by polymerase would not occur, while the effective exponential replication of mutant DNA is possible. This effect is enhanced under asymmetric conditions because the residual undesired production of the complementary strand reduces (linear growth). In the qPCR plot, the expected result is a delay in amplification curves and minority strands are specifically detected (low  $C_q$ ), despite the initial presence of wild-type alleles in high proportions.



**Figure 1.** Scheme of the mechanism of the EAB-qPCR method applied to a wild-type allele (left) and mutant alleles (right). The unfilled white squares correspond to the different mutant genotypes that may exist H: A, C or T.

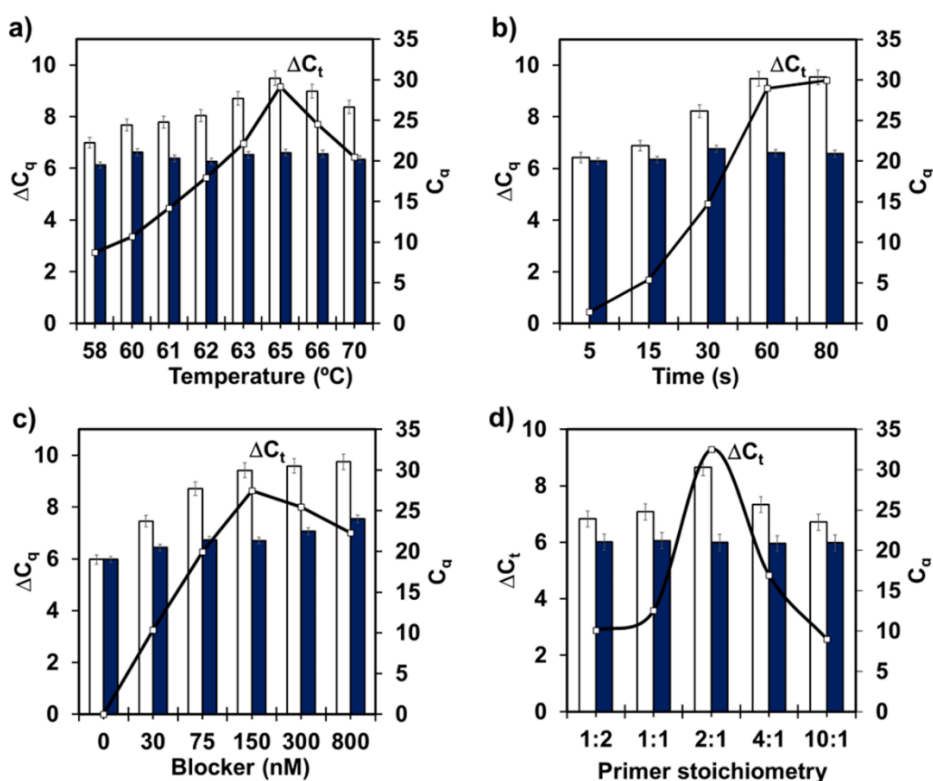
### Selective enrichment method.

The reaction conditions of the EAB-qPCR method were examined, studying the selective enrichment of the mutant *KRAS* variants. For correct genotyping, the discrimination factor ( $\Delta C_q$ ) was chosen as selection criteria. The preliminary experiments ended with high amplification yields, obtained when the primer annealing/extension step was constant at 55 °C and 35 s (Fig SI.1). Excellent results were obtained with a simple 3'-end capped oligonucleotide, which avoided using expensive molecules, such as peptide nucleic acids (PNA) and locked nucleic acids (LNA) among others [18].

Concerning the blocker annealing step, a wide operational window (temperature and time) was established from the estimated stability of the blocker/template and primer/template hybrid and the compatible conditions with the later elongation action of polymerase. Temperature variation (56-70 °C) gave a maximum curve value in the measured  $\Delta C_q$  (Fig. 2a). The discrimination factors suggested that low temperatures did not avoid the undesired primer annealing in the template strand, while high temperatures produced unstable hybrids for both the primer and blocker. These results proved that

competition between the primer and blocker for the template strands could be modulated. The time effect (5-80 s) fitted a saturation curve (Fig. 2b), as expected when considering the conventional hybridization kinetics between two perfect-match oligonucleotides [25]. The greater discrimination took place at 65 °C and 60 s.

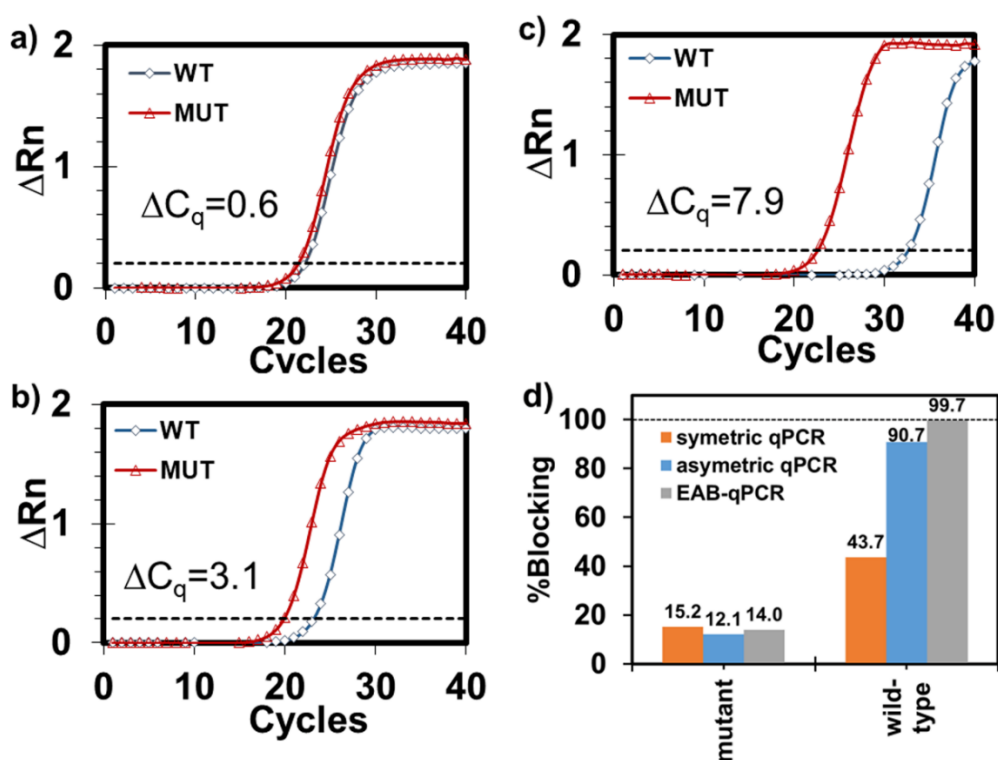
In order to improve enrichment, the blocker concentration was varied (Fig. 2c). At low concentrations, the amplification of all the variants was similar ( $\Delta C_q < 1$ ). By increasing the amount, the quantification cycle ( $C_q$ ) was nearly constant for mutants and higher for the wild type. The maximum difference without significantly reducing amplification yields was achieved at 150 nM and corresponded to half the reverse primer concentration. Therefore, adequate concentration selection favored the enrichment of mutant alleles. Different stoichiometric ratios between primers were also studied to improve the discriminant effect (Fig. 2d). When lowering the reverse primer concentration, the amplification delay of the native variant increased (higher  $C_q$ ). The shift of the wild-type curve can be interpreted based on the residual availability of the template strands to be replicated. Although the blocker was bound to the anti-sense native strand, the sense strand was still available. The maximum discrimination ( $\Delta C_q$ ) was reached using 300 nM of the forward primer and 150 nM of the reverse primer (ratio 2:1). These conditions reduced the linear residual amplification of the native variant.



**Figure 2.** Discrimination effect depending on the EAB-qPCR conditions. a) temperature, b) time, c) blocker concentration, and d) stoichiometric ratios between primers (forward:reverse). Target: the *KRAS* gene (codon 12-13). Mutant: p.G13D (c.38G>A). Template: Cell lines at  $10^5$  copies.

### Comparison to conventional blocked qPCR.

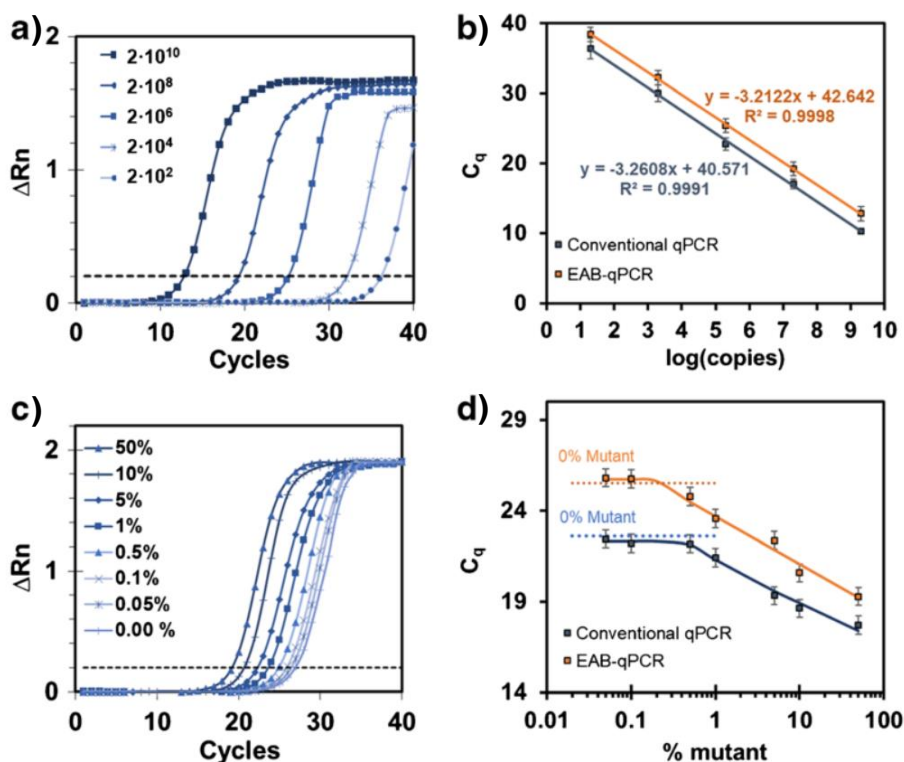
The enhanced genotyping capability of the EAB-qPCR method was experimentally confirmed. Conventional blocked qPCR was chosen as a control because it allowed to evaluate the effect of the blocker hybridization in the amplification yield and the assay selectivity, keeping all other conditions unchanged. The amplification curves showed the blocker annealing step favored the inhibition of the wild-type allele more than the mutated variants by displacing curves to higher cycles (Fig. SI.2). Discrimination capability was also compared to conventional blocked qPCR under symmetric and asymmetric conditions. Both methods resulted in a long wild-type curve delay (Fig. 3) and the calculated discrimination factors ( $\Delta C_q$ ) were 0.6 and 3.1 for conventional blocked qPCR approaches, compared to 7.9 for EAB-qPCR. Thus, our novel method more effectively inhibited the replication of the wild-type allele. Also, the blocking percentage was estimated from the delay data and the amplification efficiency equation (Table SI.3). Although mutant strands were also recognized (up to 15%), the blocker mainly hybridized to the wild-type strands, being blocking percentage 43.7-90.7% and 99.7% for blocked qPCR and EAB-qPCR, respectively.



**Figure 3.** Amplification curves: a) conventional blocked qPCR in symmetric format, b) conventional blocked qPCR in asymmetric format and c) EAB-qPCR. d) Percentage of blocking estimated from efficiency calculations. WT: wild-type template. MUT: mutant template. Target: the *KRAS* gene (codon 12-13). Template: Cell lines at  $10^7$  copies. Mutant: p.G13D (c.38G>A).

### Analytical performances.

The amplification efficiency of the EAB-qPCR method was evaluated from serial dilutions of the mutant template (c.34G>T). Fig. 4a shows a quantitative response according to the template copies. The measured  $C_q$  values matched a linear behavior from  $2 \times 10^2$  to  $2 \times 10^{10}$  copies per reaction, with a slope of -3.21 and a regression coefficient of 0.995 (Fig. 4b). From the calibration slope, good amplification efficiency was estimated with 102.6%. Comparable results were observed for conventional blocked qPCR (Fig. SI.3a), and the calculated values were -3.26, 0.995 and 104.8%, respectively. Reproducibility was determined from triplicate assays and expressed as relative standard deviation, with values going from 2.2% to 2.8%. The high consistency among the parallel results confirmed the robustness of our proposed method. Enrichment capability was estimated from the mixtures of the mutant (*KRAS* c.34G>T) and wild-type DNA, and total DNA remained at  $10^7$  copies. By lowering the mutant percentage, a longer amplification delay was recorded for EAB-qPCR (Fig. 4c) and the curve displacement was the equivalent to a reduction in the initial template copies. As expected, the quantification detection cycles showed a linear correlation with the logarithm of the mutant percentage (Fig. 4d). The estimated detection limits were 1.5% for blocked qPCR and 0.5% for EAB-qPCR. Therefore, the novel method provided 3-fold better enrichment capability.



**Figure 4.** Assay sensitivity of EAB-qPCR. a) Effect of DNA copy number for the mutant template. b) Correlation between  $C_q$  and DNA copy number. c) Effect of the mutant template %. d) Correlation between  $C_q$  and the mutant template percentage. Mutant: *KRAS* c.34G>T.

**Application: detection of mutant variants in clinical samples.**

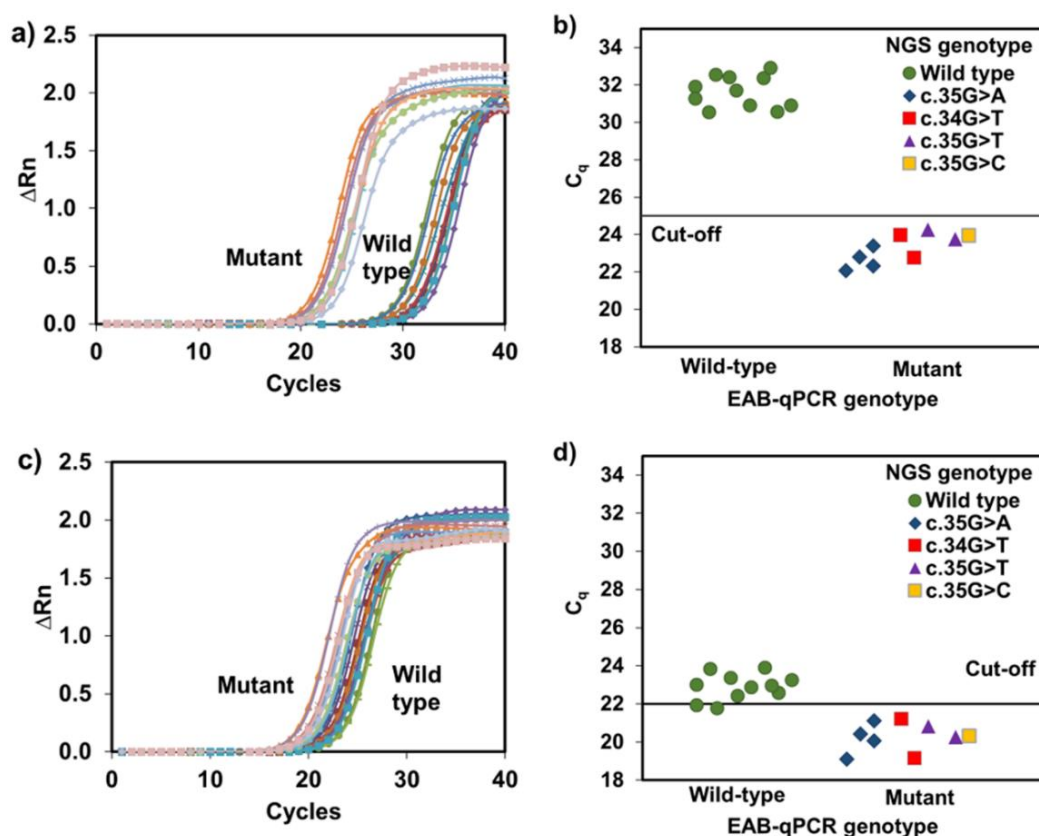
The capability of EAB-qPCR as a diagnostic tool in metastatic colon cancer was examined. In a double-blind study, biopsy tumor tissues were classified depending on *KRAS* genotype by EAB-qPCR and two reference methods (conventional blocked qPCR and NGS).

In all the patients, the EAB-qPCR provided a positive response, although conservation (formalin-fixed and paraffin-embedded) could lead to DNA degradation (Fig. 5a). Likewise, replicate assays yielded precise results ( $C_q$  variation below 0.8). Two groups of amplification curves were distinguished: one with an average  $C_q$  of 23.2, another with an average  $C_q$  of 31.6 (average delay of 8.4 cycles). By defining  $C_{q,cut-off}$ , a binary classification criterion was applied for genotyping purposes. The DNA samples with  $C_q$  over 25 were classified as wild type, whereas the samples with  $C_q$  below 25 were considered mutants (Fig. 5b). Therefore, there were 11 wild-type patients (55%) and nine *KRAS* mutants (45%). According to the oncologic guidelines, only those patients assigned to the wild-type group would be good candidates to receive monoclonal antibody therapy. On the contrary, the mutated group should develop resistance and present shorter progression-free survival [26].

These results were compared to those obtained by conventional blocked qPCR. In this method, the delay of wild-type curves was shorter (average delay of 2.6 cycles) and the classification window for genotyping, defined between two patient groups, was narrower (Fig. 5c and 5c). Thus, uncontrolled variations in the DNA template amount might lead to false-positive or false-negative assignments. In fact, the conventional method yielded an uncertain identification with two patients' samples.

Accuracy was estimated by independently sequencing patients' samples by applying NGS (Table SI.4). A total agreement of the assigned mutant genotypes validated the developed method. The estimated clinical sensitivity and selectivity were 100%. Regardless of the mutation's type and position, EAB-qPCR was capable of detecting all the studied variants in codon 12 and 13 of the *KRAS* oncogene (c.35G>A, c.34G>T, c.35G>T and c.35G>C). The mutation percentage in biopsied tissue samples correlated with the measured  $C_q$  values, estimating a detection limit about 0.05 % (Fig. SI.4). Therefore, EAB-qPCR can be considered a reliable method, although the large wild-type DNA amount in tumor tissue can hinder the detection of mutant alleles.





**Figure 5.** Mutational analysis of cancer patients. a) Amplification curves of EAB-qPCR. b) Discrimination map of EAB-qPCR. c) Amplification curves of conventional blocked qPCR. d) Discrimination map of conventional blocked qPCR. Specific mutant variants were determined by Ion Torrent sequencing technology.

### 3.1.5 Discussion

In the last decades, various qPCR-based methods have been reported for the detection of single-nucleotide changes [8, 10, 11]. A common drawback of blocked approaches is guaranteeing effective primer/blocker competence [27]. The developed method, called EAB-qPCR, shows excellent amplification performances for low-abundant mutant variants in complex samples. The novelty involves the combination of asymmetric qPCR, a blocking agent and a proper thermal cycling. Our strategy minimizes undesired DNA replications because the blocker's recognition process is promoted by controlling the thermodynamic conditions. For EAB-qPCR method, only three wild-type strands per 1000 copies escaped from the blocker's action, yielding a high selective enrichment of minority alleles.

Among the current genotyping techniques, EAB-qPCR can be classified as a high-moderate sensitive method for mutational analyses (0.1–1%) [7], and was only overcome by ddPCR technology [15] and ice-COLD-PCR [28] (0.001-0.1%). As the main difference lies in the thermal cycle, operational EAB-qPCR features were similar to other PCR-based methods, such as instrument (e.g. fluorescent thermal cycler), auxiliary

equipment or material [29]. The assay cost is lower (about 2.5 € per assay) than the approaches that use modified oligonucleotides as blocking agents (i.e. PNA, LNA) [20, 21] or fluorescent markers (i.e. COLD-PCR) [4, 17]. In addition, the oligonucleotide design is easier to be implemented than COLD-PCR approaches. Compared to ARMS-PCR, our approach avoids the use of allele specific primers, that require a laborious process for optimization, and improves the detection capability because ARMS-PCR reports false positives when the mutant content is below 1% [30].

Concerning to assay time, EAB-qPCR was generally slightly longer than other qPCR variants (1 min per cycle) and shorter than LATE-PCR (20 cycles less) [23]. Similar sample requirements were estimated given the quality and amount of DNA (4 ng/ $\mu$ L, equivalent to 1300 copies). EAB-qPCR did not entail substantial additional requisites compared to qPCR-based genetic testing that is currently performed routinely in laboratories [24, 31]. However, the discrimination capability of EAB-qPCR was several times higher than several PCR approaches and, consequently, enhanced enrichment extends potential clinical applications. Therefore, most of current genotyping techniques are expensive, tedious and complex, and require specialized techniques compared to EAB-qPCR.

Achieved sensitivity (0.5% mutant percentage) and reliability (high accuracy and reproducibility) enabled the detection of single-nucleotide mutations in clinical human tissues (solid biopsies), as we demonstrated with colorectal cancer subjects. In fact, the amplification efficiency in paraffin-embedded biopsied tissues from patients had not been hampered by some interfering factors, such as presence of inhibitors. The validation study performed by NGS evidenced that our method can detect mutant alleles in tissues, even those with low percentages of tumor cells. The accurate discrimination was achieved independently on the kind of mutated base. In most of clinical scenarios, this detection capability is enough to choose the proper treatment or patient classification. Using SYBR Green as a detection dye makes EAB-qPCR simple and universal for the detection of single-base mutations. For the determination of the specific genotype, EAB-qPCR can be improved combining it with Taqman probes (or similar probes) or adding discrimination steps, such as fast hybridization assays [32], although these approaches would increment the method complexity.

The amplification performances of EAB-qPCR proved that reliability can be applied to more situations where minority alleles can be detected. Furthermore, extending this method to detect other DNA alterations is relatively easy. One potential example is prenatal diagnosis because the enrichment of fetal DNA sequences in the presence of excess maternal DNA requires sensitive solutions. In the infectious diseases field, the detection of a few copy numbers of microorganisms is crucial. For that, the

requirements are clearly identified. Primers should be chosen for a selective amplification of the target region with a high amplification yield following the standard design algorithms for qPCR methods based on thermodynamic data (e.g. GC percentage, length, melting temperature, absence of secondary structures). The other requirements relate to the blocker. First, the blocker must strongly hybridize to the native template (wide variation in free energy,  $\Delta G$ ). To minimize the undesired inhibition of mutant variants, mutations must be in a central position given a greater destabilization of mismatched complexes (low  $\Delta G$ ). Second, the blocker/template hybrid must be stabler than the primer/template hybrid to establish the intermediate step of the thermal cycle. Third, the blocker should partially overlap the forward primer. This clamp strategy induces greater competition at the binding site by destabilizing the formation of primer/blocker/template complexes. Fourth, the 3'-end must be functionalized to avoid blocker elongation by polymerase activity during the thermal cycling. In short, the EAB-qPCR method requires a blocker oligonucleotide with stronger hybrids for the wild-type template than the mutant template ( $\Delta G_{\text{Blocker,wild-type}} > \Delta G_{\text{Blocker,mutant}}$ ).

### 3.1.6 Conclusions

EAB-qPCR turned out to be an accurate cost-effective approach for extensive use in clinical laboratory settings, because the assay is performed with oligonucleotide without modifications such as LNA or PNA. Indeed, we demonstrated that EAB-qPCR enables an accurate profiling of DNA variants to make genomic analyses more affordable and economical. Therefore, the proposed strategy has the potential to become a powerful biosensing tool to support patient prognosis and classifications in appropriate population groups for diagnostics or for receiving personalized treatment.

**Conflict of interest.** The authors declare that they have no conflict of interest.

**Ethical approval.** Research involving human subjects complied with all relevant national regulations, institution policies and is in accordance with the tenets of the Helsinki Declaration (as revised in 2013), and has been approved by the authors' Institutional Review Board. Informed consent was obtained from all individuals included in this study.

**Funding.** The authors acknowledge the financial support received from the Generalitat Valenciana (GVA-FPI-2017 PhD grant), the Spanish Ministry of Economy and Competitiveness (MINECO project CTQ2016-75749-R) and European Regional Development Fund (ERDF).

### 3.1.7 References

1. Chandler NJ, Ahlfors H, Drury S, Mellis R, Hill M, McKay FJ, et al. Noninvasive prenatal diagnosis for cystic fibrosis: Implementation, uptake, outcome, and implications. *Clin Chem*. 2020;66:207–16.
2. Schmidt RLJ, Simon A, Popow-Kraupp T, Laggner A, Haslacher H, Fritzer-Szekeres M, et al. A novel PCR-based point-of-care method facilitates rapid, efficient, and sensitive diagnosis of influenza virus infection. *Clin Microbiol Infect*. 2019;25:1032–37.
3. Reck M, Rabe KF. Precision diagnosis and treatment for advanced non-small-cell lung cancer. *N Engl J Med*. 2017;377:849–61.
4. Vargas DY, Kramer FR, Tyagi S, Marras SAE. Multiplex real-time PCR assays that measure the abundance of extremely rare mutations associated with cancer. *PLoS One*. 2016;11: e0156546.
5. Allegra CJ, Rumble RB, Hamilton SR, Mangu PB, Roach N, Hantel A, et al. Extended RAS gene mutation testing in metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy: American society of clinical oncology provisional clinical opinion. *J Clin Oncol*. 2016;34:179–85.
6. Das V, Kalita J, Pal M. Predictive and prognostic biomarkers in colorectal cancer: A systematic review of recent advances and challenges. *Biomed Pharmacother*. 2017;87:8–19.
7. Khodakov D, Wang C, Zhang DY. Diagnostics based on nucleic acid sequence variant profiling: PCR, hybridization, and NGS approaches. *Adv Drug Deliv Rev*. 2016;105:3–19.
8. Kim S, Misra A. SNP Genotyping: Technologies and Biomedical Applications. *Annu Rev Biomed Eng*. 2007;9:289–320.
9. Tsiatis AC, Norris-Kirby A, Rich RG, Hafez MJ, Gocke CD, Eshleman JR, et al. Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: Diagnostic and clinical implications. *J Mol Diagnostics*. 2010;12:425–32.
10. Milbury CA, Li J, Makrigiorgos GM. PCR-based methods for the enrichment of minority alleles and mutations. *Clin Chem*. 2009;55:632–40.
11. Lin CC, Huang WL, Wei F, Su WC, Wong DT. Emerging platforms using liquid biopsy to detect EGFR mutations in lung cancer. *Expert Rev Mol Diagn*. 2015;15:1427–40.
12. Strohmeier O, Laßmann S, Riedel B, Mark D, Roth G, Werner M, et al. Multiplex genotyping of KRAS point mutations in tumor cell DNA by allele-specific real-time PCR on a centrifugal microfluidic disk segment. *Microchim Acta*. 2014;181:1681–88.
13. Markou A, Tzanikou E, Ladas I, Makrigiorgos GM, Lianidou E. Nuclease-Assisted Minor Allele Enrichment Using Overlapping Probes-Assisted Amplification-Refractory Mutation System: An Approach for the Improvement of Amplification-Refractory Mutation System-Polymerase Chain Reaction Specificity in Liquid Biopsies. *Anal Chem*. 2019;91:13105–11.
14. Zhang H, Song J, Ren H, Xu Z, Wang X. Detection of Low-Abundance KRAS Mutations in Colorectal Cancer Using Microfluidic Capillary Electrophoresis-Based Restriction Fragment Length Polymorphism Method with Optimized Assay Conditions. *PLoS One*. 2013;8:54510.
15. Hindson CM, Chevillet JR, Briggs HA, Gallichotte EN, Ruf IK, Hindson BJ, et al. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods*. 2013;10:1003–05.
16. Li J, Wang L, Mamon H, Kulke MH, Berbeco R, Makrigiorgos GM. Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat Med*. 2008;14:579–84.
17. Castellanos-Rizaldos E, Milbury CA, Guha M, Makrigiorgos G.M. COLD-PCR enriches low-level variant DNA sequences and increases the sensitivity of genetic testing. *Molecular Diagnostics for Melanoma: Methods and Protocols, Methods in Molecular Biology*. 2014. p. 623–39.
18. Briones C, Moreno M. Applications of peptide nucleic acids (PNAs) and locked nucleic acids (LNAs) in biosensor development. *Anal Bioanal Chem*. 2012;402:3071–89.
19. Huang JF, Zeng DZ, Duan GJ, Shi Y, Deng GH, Xia H, et al. Single-tubed wild-type blocking quantitative PCR detection assay for the sensitive detection of codon 12 and 13 KRAS mutations. *PLoS One*. 2015;10:1–23.

20. Kim HR, Lee SY, Hyun DS, Lee MK, Lee HK, Choi CM, et al. Detection of EGFR mutations in circulating free DNA by PNA-mediated PCR clamping. *J Exp Clin Cancer Res.* 2013;32:50.
21. Zhang S, Chen Z, Huang C, Ding C, Li C, Chen J, et al. Ultrasensitive and quantitative detection of: EGFR mutations in plasma samples from patients with non-small-cell lung cancer using a dual PNA clamping-mediated LNA-PNA PCR clamp. *Analyst.* 2019;144:1718–24.
22. Tolnai Z, Harkai Á, Szeitner Z, Scholz ÉN, Percze K, Gyurkovics A, et al. A simple modification increases specificity and efficiency of asymmetric PCR. *Anal Chim Acta.* 2019;1047:225–30 .
23. Jia Y, Sanchez JA, Wangh LJ. Kinetic hairpin oligonucleotide blockers for selective amplification of rare mutations. *Sci Rep.* 2014;4:1–8.
24. Vashist SK, Luppá PB, Yeo LY, Ozcan A, Luong JHT. Emerging Technologies for Next-Generation Point-of-Care Testing. *Trends Biotechnol.* 2015;33:692–705.
25. Peterson AW. The effect of surface probe density on DNA hybridization. *Nucleic Acids Res.* 2001;29:5163–68.
26. Falcomatà C, Schneider G, Saur D. Personalizing KRAS-mutant allele-specific therapies. *Cancer Discov.* 2020;10:23–5.
27. McEvoy AC, Wood BA, Ardakani NM, Pereira MR, Pearce R, Cowell L, et al. Droplet Digital PCR for Mutation Detection in Formalin-Fixed, Paraffin-Embedded Melanoma Tissues: A Comparison with Sanger Sequencing and Pyrosequencing. *J Mol Diagnostics.* 2018;20:240–52.
28. Milbury CA, Li J, Makrigiorgos GM. Ice-COLD-PCR enables rapid amplification and robust enrichment for low-abundance unknown DNA mutations. *Nucleic Acids Res.* 2011;39:e2-e2.
29. Xiang Z, Wan R, Zou B, Qi X, Huang Q, Kumar S, et al. Highly sensitive and specific real-time PCR by employing serial invasive reaction as a sequence identifier for quantifying EGFR mutation abundance in cfDNA. *Anal. Bioanal Chem.* 2018;410:6751-59.
30. Huang T, Zhuge J, Zhang WW. Sensitive detection of BRAF V600E mutation by amplification refractory mutation system (ARMS)-PCR. *Biomark Res.* 2013;1:1-6.
31. Matsuda K. PCR-Based Detection Methods for Single-Nucleotide Polymorphism or Mutation: Real-Time PCR and Its Substantial Contribution Toward Technological Refinement. *Adv Clin Chem.* 2017;80:45–72.
32. Lázaro A, Yamanaka ES, Maquieira Á, Tortajada-Genaro LA. Allele-specific ligation and recombinase polymerase amplification for the detection of single nucleotide polymorphisms. *Sensors Actuators B Chem.* 2019;298:126877.

### 3.1.8 Supplementary information

#### Oligonucleotide selection

The EAB-qPCR method introduces an oligonucleotide in solution to specifically block amplification of wild-type DNA, so that a mutated DNA could be selectively amplified. The aim is the allele suppression mechanism based on the formation of blocker-template hybrid inhibits the DNA polymerization, enabling the duplication of targeted sequences.

Primers were chosen for a selective amplification of the target region with a high amplification yield following the standard design algorithms for qPCR methods based on thermodynamic data (e.g. GC percentage, length, melting temperature, absence of secondary structures). The other requirements relate to the blocker.

- a. The blocker must strongly hybridize to the native template (wide variation in free energy,  $\Delta G$ ). To minimize the undesired inhibition of mutant variants, mutations must be located in a central position given a greater destabilization of mismatched complexes (low  $\Delta G$ ).
- b. The blocker/template hybrid must be more stable than the primer/template hybrid to establish the intermediate step of the thermal cycle.
- c. The blocker should partially overlap the forward primer. This clamp strategy induces greater competition at the binding site by destabilizing the formation of primer/blocker/template complexes.
- d. The 3'-end must be functionalized to avoid blocker elongation by polymerase activity during the thermal cycling.

The candidate blocking sequences was examined by computational tools, and also experimentally. By applying the specific design requirements, the number of potential oligonucleotides exceeded 100. Nevertheless, *in silico* candidate sequences were classified into a few groups by considering similarities in thermodynamic properties. A representative set of these candidates was chosen (Table SI.1) and the EAB-qPCR method was assayed.

**Table SI.1.1.** Tested blockers for assay optimization of EAB-qPCR method. Bold letters indicate target single-nucleotide mutations. Target: *KRAS* gene (codon 12-13).

Blocker	Sequence (5'-3')	L (nt)	% GC	T <sub>m</sub> (°C)	O <sup>a</sup> (nt)	T <sub>f</sub> <sup>b</sup> (°C)	ΔG <sup>c</sup> (Kcal/mol)	ΔCt <sup>d</sup>
A	GGAGCT <b>GGTGGC</b> GTAG	16	62	55.9	1	37.4	22.0	<1
B	TGGAGCT <b>GGTGGC</b> GTAGG	18	67	60.8	2	37.4	25.8	9
C	TGGAGCT <b>GGTGGC</b> GTAGG CAAG	22	64	67.9	2	44.0	32.5	3
D	TTGGAGCT <b>GGTGGC</b> GTAG GCAAGAGCCGG	25	60	70.7	3	44.0	46.0	<2
E	TAGTTGGAGCT <b>GGTGGC</b> G TAGGCAAGAGTGC	31	58	76.2	5	41.5	46.1	<2
F	TGTGGTAGTTGGAGCT <b>GG</b> <b>TGGC</b> GTAGGCAAGAGTGC C	37	59	80.9	11	50.9	56.7	<2

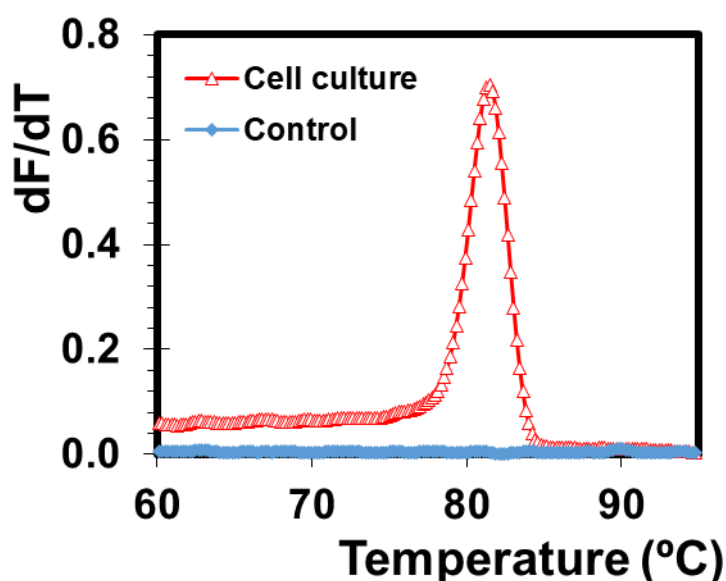
<sup>a</sup> Number of nucleotides of the blocker that overlap with the forward primer (clamp strategy).

<sup>b</sup> Folding temperature, melting temperature for secondary structures. In all cases, T<sub>f</sub> were lower than working temperatures.

<sup>c</sup> Free-energy of template hybrid for [Na<sup>+</sup>] = 1 M at 25 °C, pH = 7.

<sup>d</sup> Discrimination factor between wild-type and mutants.

The crucial blocker inertness as a primer for DNA polymerase was examined by preparing mixtures containing only the reverse primer and blocker. Negative responses (comparable to blanks) were recorded (t-test: p-value > 0.05), if the blocker included 3'-end functionalization (2',3'-dideoxycytidine). Thus, the absence of nonspecific blocker elongation was corroborated because the terminal pyrimidine-like nucleotide avoided further polymerization. The EAB-qPCR method yielded positive responses for the reaction mixtures containing forward primer, reverse primer and blocker (t-test: p-value < 0.05). Amplification specificity was confirmed because the melting curve recorded a single peak at 81.5 °C (Fig. SI.1).



**Figure SI.1.** Melting curve plotted as first derivative of the fluorescence as a function of temperature. T<sub>m</sub> peak: 81.5 °C. Sample: Cell line at 10<sup>5</sup> copies.

The blocker length for the EAB-qPCR method was studied, evaluating the measured  $C_q$  values (Table SI.1). For long oligonucleotides, the high  $C_q$  values revealed that the binding of the blocker was excessively strong, even for the mutant strands. Conversely, short oligonucleotides scarcely modified the competition against the primer for the template strands, and a mild blocking effect measured through the discrimination factor was recorded ( $\Delta C_q < 1$ ). The intermediate oligonucleotides generated enough stable blocker/template hybrids by inhibiting primer recognition for only the wild-type variant ( $\Delta C_q > 3$ ). From these experimental data, a design criterion based on thermodynamic parameters is suggested. Indeed, the estimated variation of free energy was about 0.6 Kcal/mol, the equivalent to a 2.5 °C increment in the melting temperature of the blocker in relation to the overlapping primer. For genotyping of *KRAS* gene, the selected blocker was an oligonucleotide with a melting temperature of 60.8 °C (Table SI.2). In short, the EAB-qPCR method requires a blocker oligonucleotide with stronger hybrids for the wild-type template than the mutant template ( $\Delta G_{\text{Blocker,wild-type}} > \Delta G_{\text{Blocker,mutant}}$ ).

**Table SI.1.2.** Sequences of oligonucleotides used for the EAB-qPCR method. Target: *KRAS* gene (codon 12-13). NCBI Gene 3845.

Function	Sequence (5'-3') <sup>a</sup>	Length (nt)	%GC	Tm (°C)	$\Delta G$ (Kcal/mol) <sup>b</sup>
Forward primer	CTGAATATAAACTTGTGGTAGTTG	24	33	58.3	26.4
Reverse primer	CTCTATTGTTGGATCATATTCGT	23	35	57.6	26.0
Blocker	TGGAGCT <u>GGTGGC</u> GTAGG	18	67	60.8	25.8

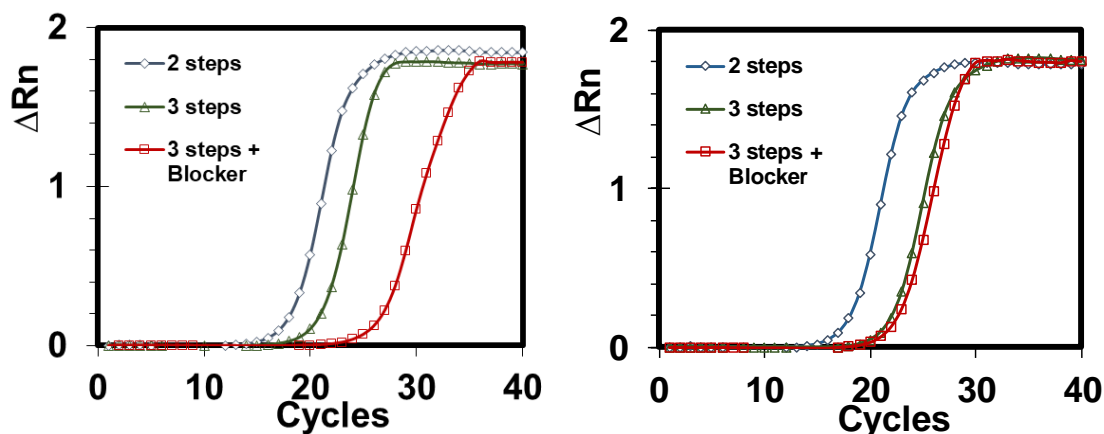
<sup>a</sup> Amplicon length: 107 pb

<sup>b</sup> Free-energy of template hybrid for  $[Na^+] = 1 \text{ M}$  at 25 °C, pH = 7

### Comparison to conventional qPCR

Figure SI.2 shows the amplification curves depending on the reaction conditions. They include conventional qPCR based on a thermal cycling of two steps (denaturation and annealing-extension), qPCR based on a thermal cycling of three steps (denaturation, annealing and extension) and EAB-qPCR method based on a thermal cycling of three steps and the presence of blocker. The amplification curves showed that our method resulted in delayed detection (about 2.5-3.5 cycles). The interpretation was associated with a partial re-hybridization of the template strands due to the reaction solution cooling. The results demonstrated that the correct discrimination between mutant and wild-type variants is achieved by EAB-qPCR method.





**Figure SI.2.** Real time curves depending on the thermal cycling and the presence of blocker. (Left) Wild-type variant. (Right) Mutant variant c.38G>A (p.G13D).

### Comparison to blocked qPCR

Conventional blocked qPCR is a method based on thermal cycling of two steps (denaturation and annealing-extension) and the presence of blocker. EAB-qPCR method is an enhanced variant for promoting the wild-type template/blocker hybrid against the template/primer hybrid. For that, the reaction conditions are modified, including thermal cycling, primer stoichiometry and blocker features. Table SI.3 compares the genotyping performances of both methods, measuring the allele suppression during DNA amplification. Figure SI.3 shows the method sensitivity, in terms of serial dilutions of the mutant template and mutant percentage able to be detected in presence of wild-type variant.

**Table SI.3.** Threshold cycle ( $C_t$ ), percentage of blocking (% B) and discrimination factor ( $\Delta C_t$ ) for conventional blocked qPCR and EAB-qPCR. Sample: Cell lines at  $10^7$  copies. Mutant: c.38G>A (p.G13D).

		Conventional qPCR (symmetric format)	Conventional qPCR (asymmetric format)	EAB-qPCR
Wild-type	$C_t$	20.90	23.41	31.19
	% B	43.7	90.7	99.7
Mutant	$C_t$	20.33	20.28	23.31
	% B	15.2	12.1	14.0
	$\Delta C_t$	<b>0.6</b>	<b>3.1</b>	<b>7.9</b>

Calculations:

$$\text{Amplification equation: } N = N_0 \cdot E^{\text{cycle}}$$

$N$ : number of copies after c-cycles

$N_0$ : initial number of template copies

$E$ : Amplification efficiency (theoretical value = 2)

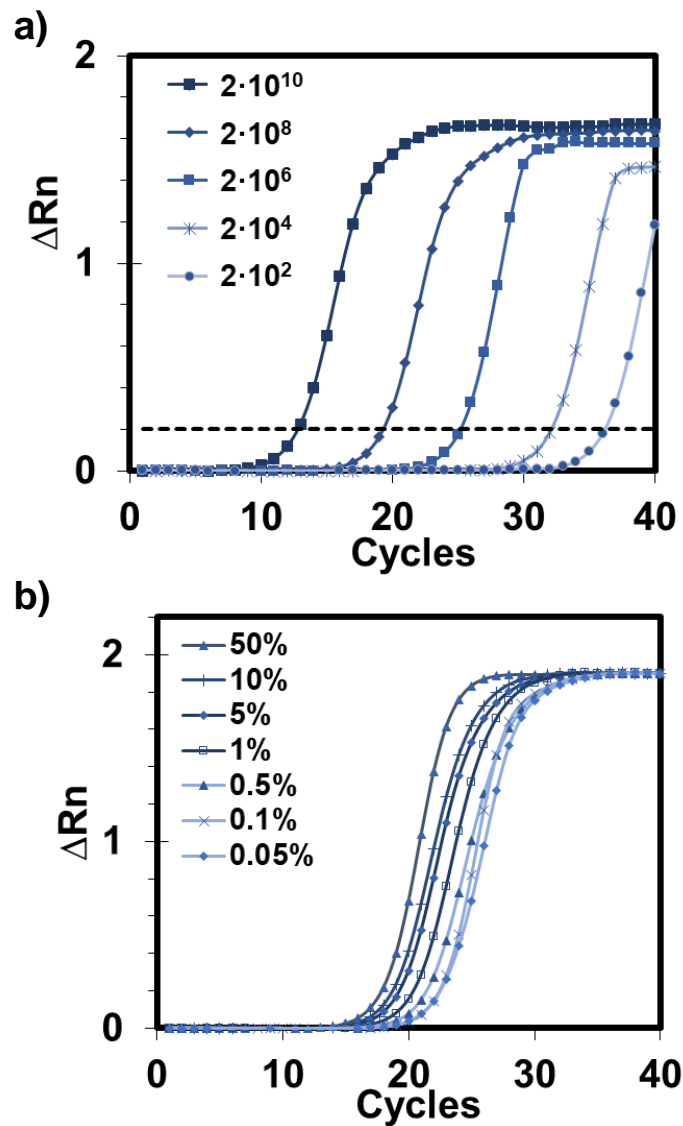
From the measured data at the threshold cycle ( $C_t$ ), the equation can be rearranged and the initial number of template copies ( $N_{0,measured}$ ) can be calculated:

$$C_t = constant - \frac{1}{E} \cdot \log(N_{0,measured})$$

The blocking percentage is calculated as:

$$\%B = \frac{N_{0,measured}}{N_{0,added}} \cdot 100$$

In EAB-qPCR, the percentage of blocking was calculated correcting the delay for the inclusion of an additional stage during thermal cycling.



**Figure SI.3.** Conventional qPCR method. a) Mutant dilutions. b) Mixtures of wild-type:mutant. Mutant variant c.38G>A (p.G13D).

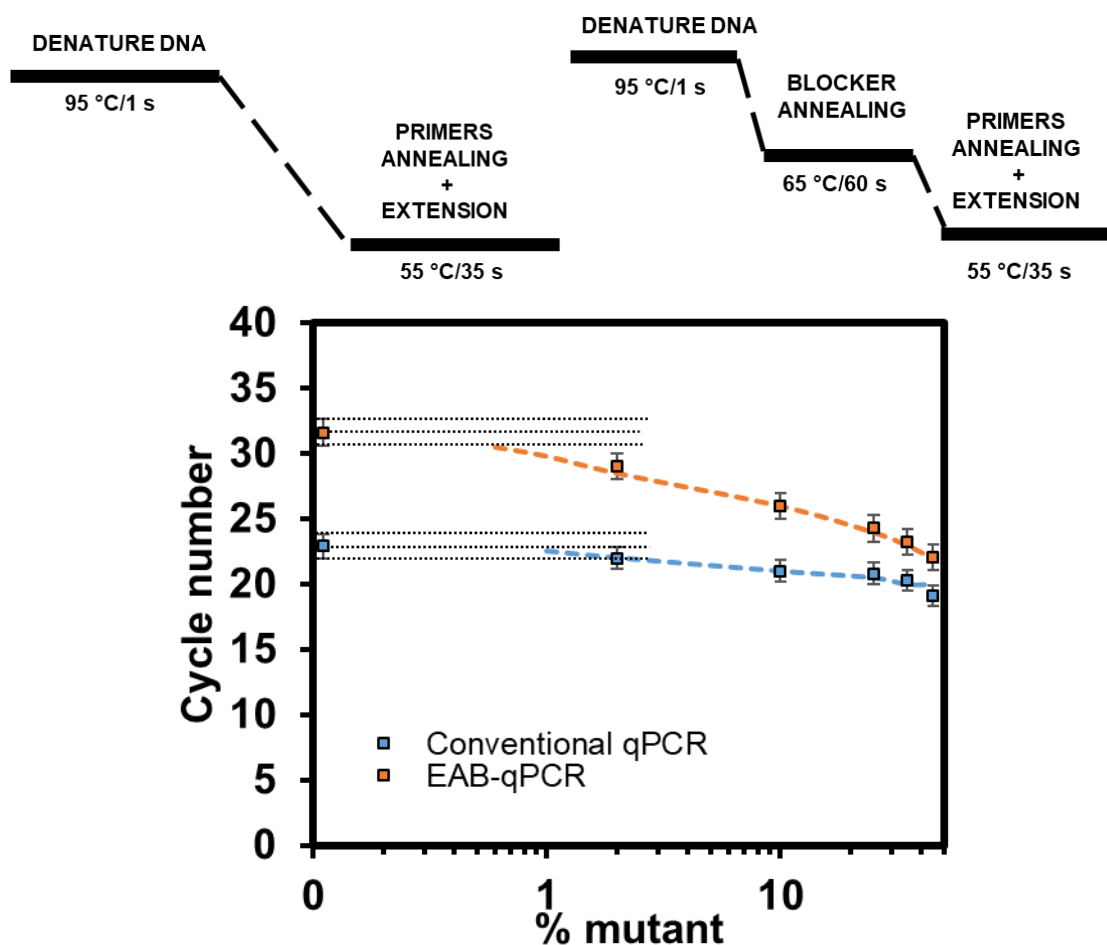
### Mutant detection from clinical samples

The capability of EAB-qPCR as a diagnostic tool in metastatic colon cancer was examined and compared to two reference methods (conventional blocked qPCR and NGS).

**Conventional blocked qPCR.** Thermal cycling was 2 min at 50 °C, 10 min at 95 °C, followed by 40 cycles of amplification of 1 s at 95 °C (denaturation) and 35 s at 55 °C (primer annealing and extension).

**Next Generation Sequencing.** The method was developed using the OncoPrint Solid Tumor Assay panel (Thermo Fisher Scientific, USA), which determines the mutational status of 22 oncogenes in cancer, including the *KRAS* oncogene. The employed technology was the Ion Personal Genome Machine® (PGM) System (Thermo Fisher Scientific, USA) and its associated bioinformatics tools (Torrent Server, Ion Reporter Server and IGV). Each result was obtained by averaging 307 448 readings with a mean coverage of 2926x. Thus, each nucleotide was read 2926 times on average, and 96.94% of the regions of interest obtained a coverage that equaled or exceeded 20% of the mean coverage (uniformity). The readings aligned with the sequences of interest were 93.32%.

The content of the mutation from biopsied tissue samples, reported by sequencing method, correlated with the measured  $C_q$  values (Figure SI.4). From these data, the estimated detection limit of EAB-qPCR (0.5 %) was lower than conventional blocked qPCR and agreed with the measured values from serial dilution experiments. Also, interpolating on the regression line, it would be possible to quantify the content of the mutation quite accurately. Since antiEGFR treatment is administered when the mutation is below 1%, patients with approximately  $C_q \approx 31$  obtained by EAB-qPCR, should necessarily receive the immunotherapy.



**Figure SI.4.** (Top) Thermal cycling for conventional qPCR (left) and EAB-qPCR (right). (Bottom) Estimation of detection limits from patient samples (biopsied tissues). Black dashed lines indicate the cycle number ( $C_t$ ) for the amplification of wild-type variant. The intersection between background values and the experimental curve is an estimation of the minimal mutant percentage able to produce a detectable modification on amplification curves.

Table SI.4. shows the genotyping results for the detection of *KRAS* mutation (codons 12-13) for the studied methods. A total agreement was reported in the double-blind study, validating that EAB-qPCR allows to discriminate between the wild-type allele and the mutant variants (c.35G>A, c.34G>T and c.35G>C). Compared to NGS method, EAB-qPCR does not identify which of the mutations is present. But, most of clinical scenarios only needs a technique to know the presence or absence of mutant alleles to select the proper treatment. Then, our method fulfills the requirements to support a wide range of personalized medicine decisions by a cheaper molecular diagnostic tool.

Table SI.1.4. KRAS mutation analysis of 20 clinical samples by studied methods

Patient sample	EAB-qPCR	qPCR	NGS	DNA code	Protein code
	Genotype	Genotype	Genotype		
1	WT	WT	WT		
2	WT	N.A.	WT		
3	WT	WT	WT		
4	WT	N.A.	WT		
5	WT	WT	WT		
6	WT	WT	WT		
7	WT	WT	WT		
8	WT	WT	WT		
9	WT	WT	WT		
10	WT	WT	WT		
11	WT	WT	WT		
12	M	M	M	c.35G>A	p.G12D
13	M	M	M	c.35G>A	p.G12D
14	M	M	M	c.35G>A	p.G12D
15	M	M	M	c.35G>A	p.G12D
16	M	M	M	c.34G>T	p.G12C
17	M	M	M	c.34G>T	p.G12C
18	M	M	M	c.35G>T	p.G12V
19	M	M	M	c.35G>T	p.G12V
20	M	M	M	c.35G>C	p.G12A

WT: wild-type; M: mutant; N.A.: not assigned





**CHAPTER 2:**

**Allele-specific ligation and  
recombinase polymerase  
amplification for the detection of  
single nucleotide polymorphisms**

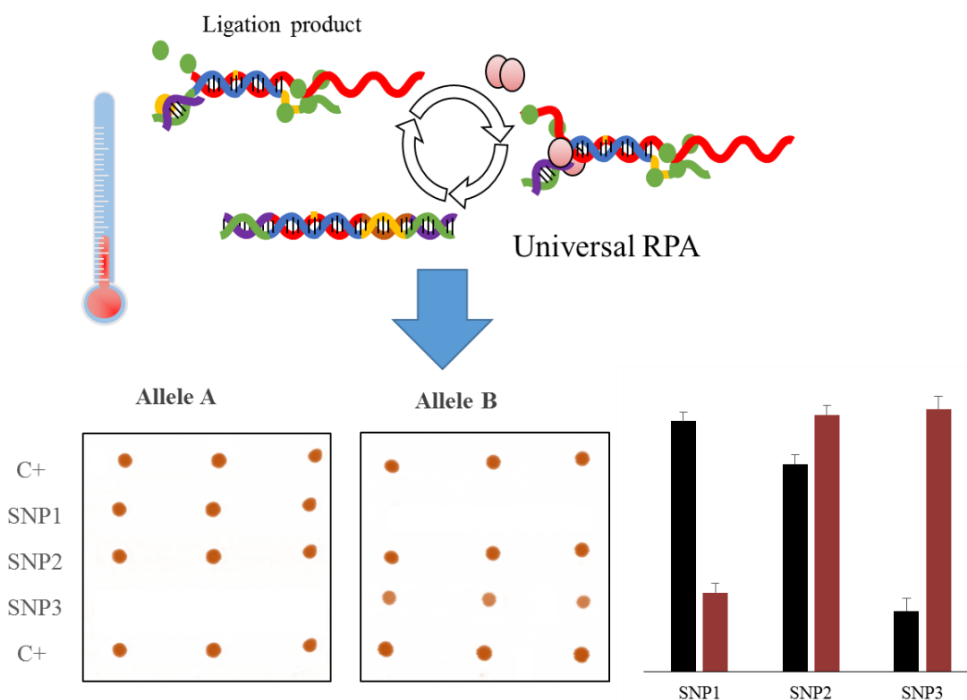




### 3.2 CHAPTER 2: ALLELE-SPECIFIC LIGATION AND RECOMBINASE POLYMERASE AMPLIFICATION FOR THE DETECTION OF SINGLE NUCLEOTIDE POLYMORPHISMS

In this chapter, a different strategy is addressed to overcome the technical requirements and complexity associated with current SNV detection technologies. The solution is based on alternative reactions to the standard PCR. The novelty is to combine a ligation process, as a base sequence discrimination strategy, and an isothermal reaction, such as recombinase polymerase amplification (RPA) to increase the copy number. In particular, the novel method avoids using the thermocycler and reduce energy consumption enabling the POC detection of multiple SNV.

Until now, a limitation of isothermal techniques is their low multiplex capacity. Our new RPA method solves it by using universal primers and barcoded probes that simultaneously amplified allele-specific products obtained due to the high selectivity of the ligases. Subsequently, the generated RPA products are hybridized with allele-specific probes. To demonstrate the principle, we chose a high-performance detection platform: compact disc technology. Concretely, probes are anchored to a Blu-ray disc and recognition products are optically detected by immunostaining using a modified Blu-ray disc drive. The purpose is to evaluate the analytical performance of this technology as DNA assay support and detector to obtain a high-throughput, low-cost POC genotyping system. As a proof-of-concept, the assay is applied for simultaneous recognition in patient samples of three SNV related to response to treatment with anticoagulants such as warfarin (*CYP2C9\*2*, *CYP2C9\*3* and *VKROC1*).



**Summary figure.** Overview chapter 2



# Allele-specific ligation and recombinase polymerase amplification for the detection of single nucleotide polymorphisms

Ana Lázaro <sup>a</sup>, Eric Seiti Yamanaka <sup>a</sup>, Ángel Maquieira <sup>a, b, c</sup> ✉, Luis A. Tortajada-Genaro <sup>a, b, c</sup> ✉

## 3.2.1. Abstract

A novel multiplex detection of single nucleotide polymorphisms (SNPs), with point-of-care testing as its aim, is reported for supporting pharmacogenetic-based decisions. The strategy relies on allele-specific ligation to discriminate base sequence variations at the SNP site and the extension of generated products by isothermal amplification and recombinase polymerase amplification (RPA). Having demonstrated the assay principle, the variables for the adequate integration of the ligation-amplification process were studied and compared to a conventional PCR approach. One key result was the development of RPA in a universal format using short-length primers, which enabled detection based on selective hybridisation on a barcode-DNA chip and a low-cost optical sensor. As proof of concept, we successfully discriminated genetic variants related to cardiovascular diseases and the adequate prescription of oral anticoagulant antagonists of vitamin K (genes *CYP2C9* and *VKROC1*).

**Keywords:** universal RPA; ligation; SNP; optical sensor; hybridization chip; personalised healthcare.

## 3.2.2 Introduction

The implementation of personalised healthcare needs better accessibility to genetic information given current technological and economical barriers [1]. Although the use of sequencing techniques increases every year, this approach is expensive or unaffordable in many clinical scenarios [2]. Thus simple, accurate and cost-effective solutions are also required as alternatives. The excellent performance of ligase-mediated detection makes these techniques one of the preferred alternatives for routine

genotyping applications [3,4]. In order to achieve the required sensitivity, ligation has been coupled to an amplification step prior to a quantification technique, such as capillary electrophoresis [5], chemiluminescence [6], bead-based colorimetric detection [7], chip-based fluorescence [8], real-time fluorescence [9] and chip-based reflection [10]. However, a pending challenge is a faster response time and an easier operation system to be integrated into and miniaturised on a compact platform.

Isothermal amplification approaches, which replicate nucleic acids at a constant temperature, are a powerful alternative to conventional PCR, and they open up new ways to achieve the required sensitivity in point-of-care devices [11]. In the last few decades, particular rolling circle amplification (RCA) features have been extensively exploited for ligase-based genotyping [12,13]. Recently, loop-mediated isothermal amplification (LAMP) has been integrated into a ligation method for the end-point detection of microRNAs [14].

This study addresses the recombinase polymerase amplification (RPA) of allele-specific ligation products. This isothermal technique combines enzymes to facilitate the binding of primers to the DNA target and the stabilisation of reaction intermediates to avoid heating double-stranded nucleic acid for template separation [15]. With the adequate integration, the potential advantages are a simplification of the assay platform (materials, dimensions and bonding technique), and the demanded equipment [16,17]. In fact, this technique is also compatible with fast-response diagnostic and equipment-free approaches [18].

Our research aims to face the complex challenge of the simultaneous amplification of several products generated in an enzyme-mediated ligation, despite the limited multiplexing capability of RPA [19]. The research hypothesis states that generic primers might enable an increment in copy numbers for all the formed ligation products (universal RPA). Regarding detection, RPA has been combined with conventional techniques [19] until advanced hybridisation methods on several platforms, such as microtiter plates [20], microdevices [21] and arrays [22]. In the present study, a novel approach based on a barcode chip and a portable optical sensing device is developed. The expected advantages are high-throughput capability and easy selection of hybridisation conditions, a flexible working range, and better assay selectivity and sensitivity [23,24]. In fact a similar approach has recently been successfully applied for ligation-PCR products [10].

As the RPA-based method should require a shorter incubation time (30-90 min), a lower constant temperature (35-40°C) and a simpler protocol (i.e. single incubation step), an integrated assay would be more amenable for point-of-care applications. As demonstrated, the proposed methodology was applied to the genotyping of single

nucleotide polymorphisms (SNPs) in the pharmacogenetics field. These methods are required for the discrimination of adverse drug reactions given their significant impact on public health in terms of patient status, death rates and healthcare costs [25].

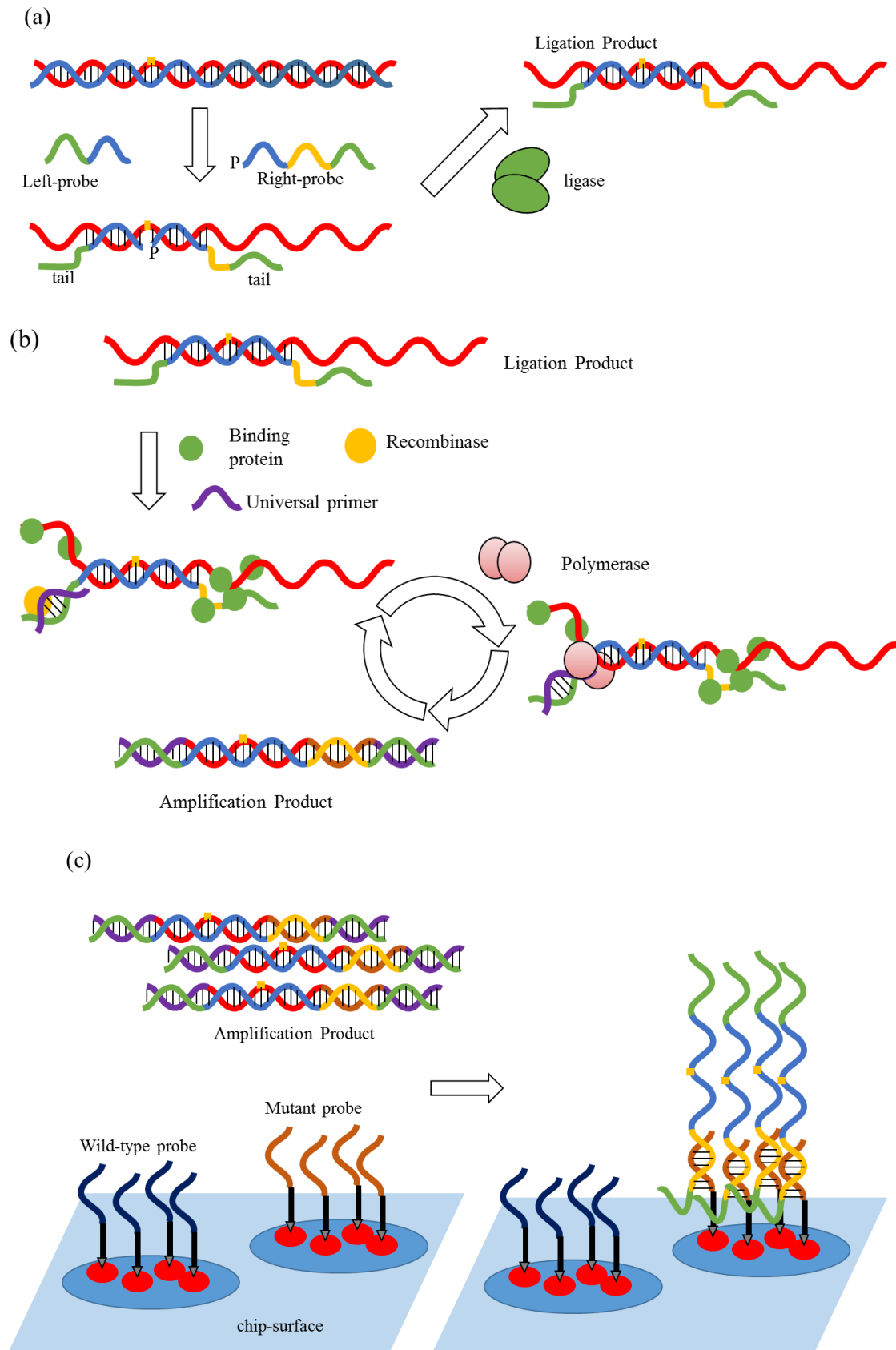
### **3.2.3 Materials and methods**

**Principle of SNP genotyping.** The method combines a ligation and an isothermal amplification in a universal format (Fig. 1a and 1b). The first stage is based on using ligase and two specific oligonucleotides per genomic variant. In the presence of the target nucleotide, the left probe with a 5'-tail (forward universal primer) ligates to the right probe with a 3'-tail (reverse universal primer). In a second stage, these specific tails enable isothermal amplification. Indeed the mechanism is based on the action of binding proteins and recombinase to yield a cyclic process of primer annealing and enzymatic extension. The universal design of the primers leads to the amplification of all the products simultaneously with a single primer pair.

One main advantage of this method is its multiplexing capability. The incorporation of a barcode into the ligation tail enables detection based on the selective hybridisation of each ligation amplification. In a third stage, each allele-specific product hybridises to the complementary probe attached to the chip surface (Fig. 1c). The resulting image pattern can be related to a precise genetic profile.

**Ligation.** Two reaction solutions contained the correspondent ligation probes at 50 nM (allele wild-type or mutant mixture) and 30 ng genomic DNA in Tris-EDTA buffer (Tris-base 10 mM, EDTA 1 mM, pH 8) was prepared. After probe annealing (5 min, 98°C and 30 min, 65°C), the ligase (Salsa Ligase-65, MRC-Holland, The Netherlands) was added. The solutions were incubated for the ligation process (54°C, 15 min) and enzyme deactivation (98°C, 5 min). In each assay batch, the ligation control analysed the genomic DNA samples together. For optimisation purposes, another ligase (ampligase, Lucigen-Epicentre, USA) was also studied.

**Isothermal amplification.** The reagents used for universal RPA were TwistAmp Basic RPA kit (TwistDx, UK). Two reaction mixtures (12.5 µL) were prepared with rehydrated buffer, 14 mM of magnesium acetate, 400 nM of the associated upstream primer (ACTTCGTCAGTAACGGAC or GAGTCGAGGTCATATCGT), 400 nM of the downstream primer (GACTCACTATAGGCAGAC), 10 µM of digoxigenin dUTP and 1.25 µL of the ligation product. The solutions were heated at 37°C for 40 min in an oven. In each assay batch, the amplification controls were analysed together with the ligation products.



**Figure 1.** Scheme of the multiplexed discrimination assay for single-base changes (polymorphisms or mutations): a) allele-specific ligation. b) Universal RPA. c) Hybridisation in an array format.

**Detection for single assays.** The RPA products from the discrimination assay of a single polymorphism were visualised by agarose gel electrophoresis. After clean-up by silica-gel membrane adsorption (PCR purification kit, Jena Bioscience, Germany), electrophoretic separation was done in 3% agarose gel, at 110 V, fluorescent dye (Realsafe Nucleic acid Staining Solution 2x, Real Lab., Spain). Products were detected after both the addition of the fluorescent dye and the measurement of the response in a microplate reader (Wallac, model Victor 1420 multilabel counter, Finland).

**Detection for multiplexed assays.** The amplification products of several polymorphisms were simultaneously detected by a hybridisation assay based on Blu-Ray technology (chip and reader). The layout of the barcode probes enabled the simultaneous analysis of 36 samples per disc (Supplementary Information). The ligation-RPA product (6  $\mu$ L) was mixed with hybridisation solution (21  $\mu$ L), composed of SSC buffer 3 $\times$  (sodium citrate 45 mM, NaCl 450 mM, pH 7), 20% formamide, and 2.5 $\times$  Denhardt's reagent. A positive hybridisation control (labelled amplification product of *ACTB* gene) was added (3  $\mu$ L). The solution was denatured at 92°C for 10 min and transferred to the array surface. Discs were incubated at 37°C for 45 min in a conventional oven, and gently washed for 1 min first with SSC 0.1 $\times$  and then with SSC 0.01 $\times$ .

Chip colorimetric staining was based on incubation with conjugated antibodies (anti-digoxigenin) and the addition of colorimetric substrate (Supplementary Information). Finally, the disc was placed into the Blu-Ray drive and scanned by the pickup laser (405 nm). The reflected light was collected and digitised to generate monochromatic images (tagged image file format with resolution of 16 bit). The optical intensity signals of each spot were quantified using in-home software.

**Sample analysis.** The assay performances were evaluated by applying the method for the genotyping of single-nucleotide changes from human samples. The biosensing method was applied to the SNPs associated with widely prescribed drugs in primary care, such as heart diseases [26]. The target variants were rs1057910, rs1799853 and rs9923231 located in genes *CYP2C9* and *VKORC1*. Primers and probes were designed according to the thermodynamic parameters associated with the perfect-match and mismatched duplexes (Supplementary Information). Several quality controls were carried out to ensure that each step and the entire assay were correctly performed. Subjects (n=30) were recruited for the present study according to ethical guidelines. The DNA extracts from buccal smear samples were diluted to 4 ng/ $\mu$ L (1,300 copies) and analysed as described in the previous sections (multiplexed format).

For assignment purposes, a genotype decision rule was constructed based on a spot signal-to-noise ratio for both allele-specific products. The automated algorithm enabled signal transformation and sample classification. Firstly, responses were converted into polar coordinates ( $\theta_{\text{normalized}}$ ,  $r$ ). Then patients were grouped into the correspondent population based on their relative position.

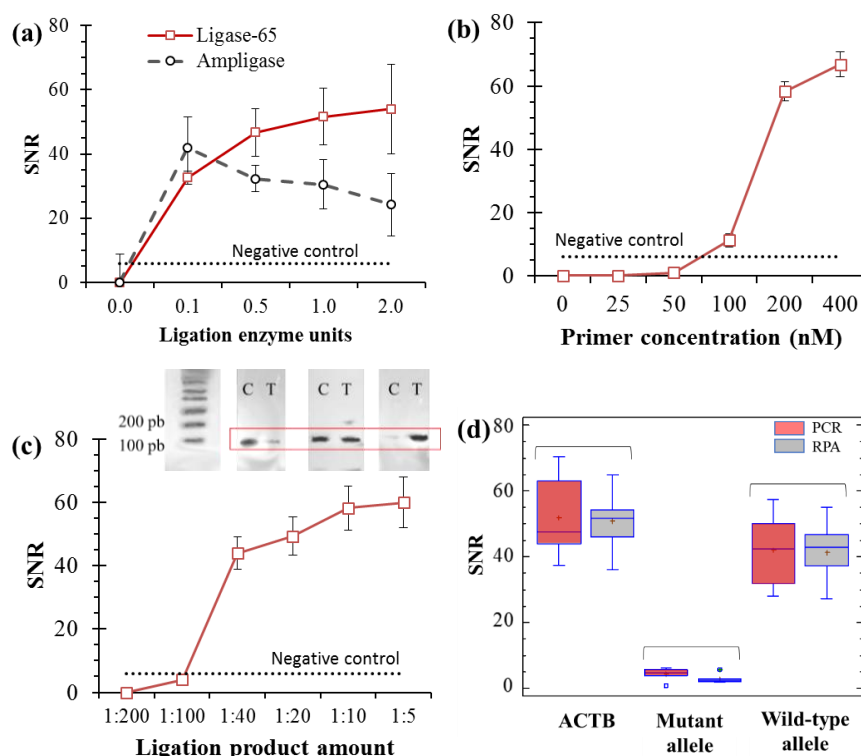
### 3.2.4 Results and discussion

#### Set-up of the ligation-universal amplification method

The main research challenge to discriminate single-base changes in several targeted genes is the combination of selective ligation and RPA in a universal format. The *in-silico* method was employed to select probes by considering that the ligation oligonucleotides must incorporate primer sequences as a tail for their latter isothermal amplification. The thermodynamic-based design provided sequences without homology to other sequences in the human genome or stable secondary structures. Initial experiments were performed using artificial ligation products to confirm the correct selection of oligonucleotides (Supplementary Information). One relevant result was about RPA primer length compared to conventional approaches that recommend values between 30 and 35 [15]. Shorter sequences (18 nucleotides) provided sensitive and precise amplification yields ( $> 10^7$ ) with no loss of selectivity. These performances agreed with those previously reported, where short nucleotides can be used for certain biosensing applications [27]. In short, the isothermal amplification of ligation products, which included the selected tails, was feasible.

The following experiments focused on ligation by studying two approaches that rely on the selective activity of ligase-65 and ampligase, respectively. The first enzyme is used in the technique called multiplex ligation-dependent probe amplification (MLPA) [5]. The second enzyme has been employed in different ligation-based techniques [9]. The objective was to establish the conditions for which two probes were bound adjacently on a target sequence, ligated and amplified. Negative responses (signal-to-noise ratios below 3) were obtained in all the phials with incomplete reaction mixtures or non-complementary templates. Positive assays were reported for the reaction, which included perfect-match probes to the DNA template and demonstrated that both ligation approaches were compatible to the RPA mechanism-based amplification (Fig. 2a). Nevertheless, the ligase-65 reactions (enzyme 1U, 60°C, 15 min) provided higher and more reproducible signals than the ampligase reactions in the presence of the complementary target DNA.

Regarding the RPA conditions, the saturated signals for the perfect-match duplexes were achieved after a 40-minute incubation with universal primers and ligation products. This amplification kinetics agreed with previous studies based on conventional RPA [15]. Figure 2b shows that the primer concentration for the amplification based on a universal format was comparable to previous studies (400 nM). It is worth emphasising the low responses of the negative controls because one important drawback of ligation-based methods is false-positives (i.e. absence of template DNA) [4]. Another relevant result was the amount of ligation solution required for correct isothermal amplification (Fig. 2c). The best results were achieved for a dilution factor of 1:5 because excessive dilution reduces the final number of copies, and higher ligation solution values drastically modified the RPA conditions, probably due to buffer capacity and the inhibiting environments of the ligation components. A parallel study was performed to compare the amplification yields in a thermal cycling regime (PCR) and the isothermal mode (RPA). Statistics analyses confirmed that both methods provided comparable mean signals (t-test,  $p > 0.05$ ) and comparable standard deviations (F-test,  $p > 0.05$ ) (Fig. 2d). Compared to ligation-PCR methods, ligation-RPA requires less thermal variations and less precise heating/cooling technologies. These features could favour an easier miniaturisation and automation (i.e. centrifugal-based microfluidic system) [11,17,28].



**Figure 2.** Integration of ligation and universal RPA. a) Ligation enzyme concentration. b) Amplification primer concentration. c) Ligation product amount. d) Comparison of PCR (red) and RPA (blue) for different reaction mixtures. SNR: signal-to-noise ratio. DNA template: 1000 copies (wild-type). Reference genes: *ZFX* and *ACTB*. Tested gene: *VKORC1*. Expected product length: 103 pb.

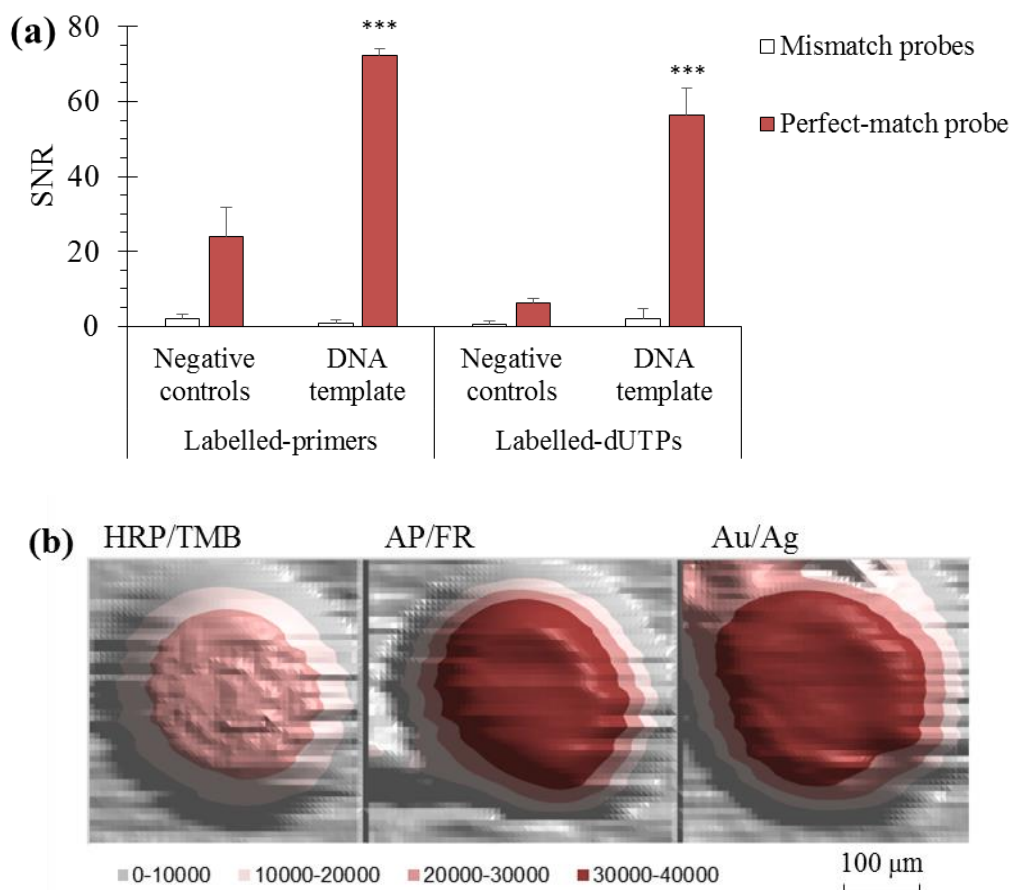


### Set-up of optical detection

The detection of a single polymorphism or mutations was easily achieved by conventional approaches, such as gel electrophoresis and fluorescent reading. The detection of a band or a fluorescent response indicated the presence or the absence of nucleotide change. However, the simultaneous detection of several targets required a sensing assay with multiplexing capability. Among the different techniques currently available for DNA diagnostics, the detection of ligated-amplified products was approached using a solid-phase hybridisation assay for its excellent performance [10,20].

The proposed method was based on inserting a barcode into the ligation oligonucleotide. Firstly, the experiments proved that its presence did not modify the ligation or amplification yields (t-test,  $p > 0.05$ ). Under the selected conditions, the combination of both the isothermal enzymatic processes (ligation-amplification) was feasible, and opened up an innovative path to increase assay sensitivity based on universal hybridisation probes. Secondly, detection in an array format based on Blu-ray technology (BD) was studied for its potential for point-of-care applications [29]. The main challenge was to select suitable conditions, as described in Supplementary Information. By considering the assay restrictions, different labelled-reagents and immunostaining methods were compared (Fig. 3). In all the approaches, the platform preserved the general optical/mechanical properties to enable the correct reading of the generated spots, with sensitivity in the order of fmol. However, the staining based on labelled-nucleotides/alkaline phosphatase-antibody/fast red substrate displayed better analytical performance than other combinations.

The BD-based method was successfully compared with the DNA assay performed on planar plastic chips and the later imaging using a conventional scanner. Nevertheless, BD technology displayed appealing features compared to other analytical platforms [4,30]. Firstly, the high-hydrophobic nature of the disc surface is interesting due to the RPA mixture composition (low background signals). Secondly, a high-resolution reader is suitable for high-density arrays of several barcode probes. Thirdly, chip/reader features (optical, mechanical, etc.) match an accurate, versatile and affordable detection system for point-of-care scenarios [11]. In summary, the developed method is an interesting approach as portability, simplicity and low cost in both acquisition and maintenance are required in diagnostic systems that use nucleic acids as biomarkers [25].



**Figure 3.** Optical detection of RPA products (array format). a) Effect of digoxigenin-labelling method during amplification by comparing 5'-functionalised universal forward primers and labelled 2'-deoxyuridine 5'-triphosphate (dUTPs). b) Response curve of an array spot depending on the staining method. DNA template: 1,000 copies. Sample: genomic DNA mutant in the *VKORC1* gene. \*\*\*: p-value <0.001. SNR: signal-to-noise ratio. HRP/TMB: Horseradish peroxidase system. AP/FR: Alkaline phosphatase system. Au/Ag: Nanoparticle system.

### Multiplex evaluation

The simultaneous discrimination of several polymorphisms was studied. The first task was to select compatible ligation probes based on a similar estimated stability of hybrids (template-ligation probe and RPA product-barcode) and comparable alteration due to the presence of a single-nucleotide mismatch (Supplementary Information). The determined features estimated a feasible discriminatory analysis of three targets simultaneously given the low cross-reactivity probability.

The following activities focused on the experimental confirmation. The optimised method was tested using different formats from an assay run for the genotyping of a single SNP (1-plex) to an assay for three polymorphisms (triplex). Table 1 shows the recorded spot signals in each case, indicating that no molecular recognition was produced for the mismatched templates (null target-to-target cross-reactivity). The recognition profiles corresponded precisely with the expected perfect-match complexes (probe-template), independently of assay multiplexing. The results demonstrated that all

the sets of ligation oligonucleotides were correctly ligated and amplified in the same solution to yield unequivocal hybridisation. Therefore, our method involves a higher multiplexing capability approach compared to other approaches, such as performing an oligonucleotide ligation assay after the isothermal amplification of a specific region [30].

**Table 1.** Spot mean SNRs depending on the ligation mixture (single, duplex or triplex) and the forward primer used in the amplification reaction (wild-type or mutant). Patient 1: heterozygous in rs1799853, wild-type in rs1057910, and wild-type in rs9923231. Patient 2: wild-type in rs1799853, wild-type in rs1057910, and heterozygous in rs9923231.

		Array probe <sup>(b)</sup>							
		Wild-type RPA			Mutant RPA				
	Mix <sup>(a)</sup>	rs1057910	rs1799853	rs9923231	rs1057910	rs1799853	rs9923231		
Patient 1	single	1	<u>46±10</u>	6±3	2±3	<u>46±3</u>	2±3	4±2	
		2	4±2	<u>34±7</u>	3±2	5±2	6±5	2±1	
		3	4±2	6±6	<u>61±4</u>	5±9	6±3	7±1	
	duplex	1-2	<u>67±8</u>	<u>64±5</u>	7±3	<u>45±2</u>	5±3	6±5	
		1-3	<u>74±9</u>	6±6	<u>73±10</u>	<u>65±7</u>	8±2	7±3	
		2-3	7±3	<u>58±7</u>	<u>80±5</u>	6±3	6±4	6±3	
	triplex	1-2-3	<u>57±6</u>	<u>30±4</u>	<u>52±4</u>	<u>55±5</u>	10±3	5±2	
	Patient 2	single	1	<u>26±5</u>	4±2	3±1	9±5	6±3	2±1
			2	3±1	<u>29±1</u>	2±1	3±2	3±2	3±2
3			1±2	4±2	<u>57±3</u>	2±2	2±3	<u>27±8</u>	
duplex		1-2	<u>42±12</u>	<u>25±5</u>	3±2	5±5	4±2	3±2	
		1-3	<u>60±13</u>	3±2	<u>44±10</u>	5±5	3±2	<u>66±6</u>	
		2-3	4±1	<u>39±5</u>	<u>67±7</u>	1±2	4±1	<u>48±3</u>	
triplex		1-2-3	<u>49±8</u>	<u>35±3</u>	<u>56±1</u>	7±4	3±3	<u>37±2</u>	

<sup>(a)</sup> Ligation mixtures, 1: rs1799853, 2: rs1057910, 3: rs9923231

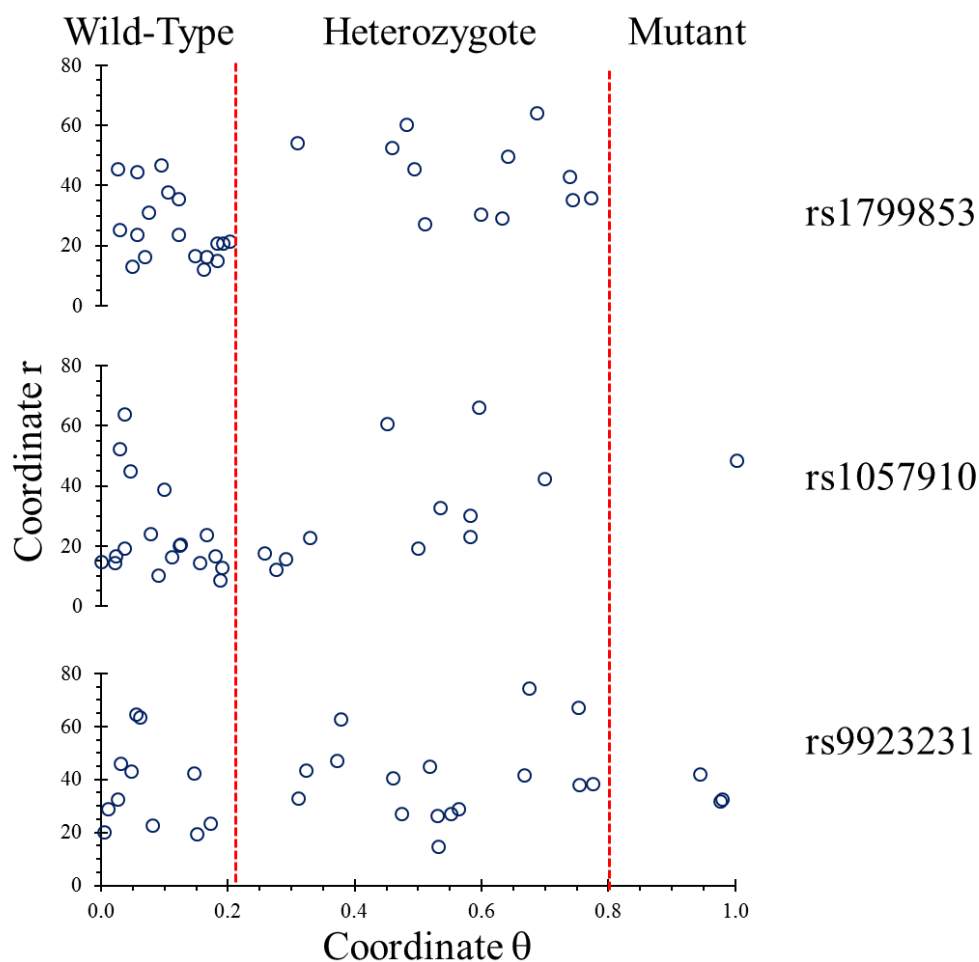
<sup>(b)</sup> Underlined values indicate a significant difference from the negative control response (t-test, p-value > 0.05).

### Demonstration of accurate SNP scoring

Method performances were assessed for the classification of patients suffering cardiovascular diseases. In a pilot study, 30 different individuals were genotyped by identifying the specific variant for each targeted polymorphism according to the hybridisation patterns (Supplementary Information). A heterozygote patient led to positive amplifications for both RPA reactions and consequent hybridisation in both probes. A homozygous patient produced single positive amplification and consequent hybridisation in the specific probe. Nevertheless, an automated algorithm was applied to reduce patient classification uncertainty due to residual signals. The assignment was based on its position on the discrimination map (Fig. 4). The output was that a nucleotide change was detected in 12 (40%), 12 (40%) and 19 (63%) cases for rs1799853 (*CYP2C9* gene), rs1057910 (*CYP2C9* gene) and rs9923231 (*VKORC1* gene), respectively.

Therefore, the uncertainty of a drug treatment's individual efficacy could be minimised based on the specific genetic profile. A revision of anticoagulant drug doses is recommendable for mutant individuals [26].

Assay reproducibility was determined from the replicated samples and was expressed as a relative standard deviation. The values were 6% for the intra-assay and 19% for the inter-assay. Assay accuracy was validated by the agreement of the genotypes assigned by the independent sequencing of patient samples (Supplementary Information). Although further research is needed, the investigated assay showed certain advantages compared to sequencing methods. It provided key genetic information for personalised diagnostic, prognostic and therapy selection tools more simply and cheaply. The developed approach can be potentially extrapolated to other screening genomic applications because the assay principles are based on universal formats. The biosensing assay is compatible to the technologies used for decentralised scenarios, such as a doctor's office, remote locations, emergency needs and ubiquitous low-resource health systems [2,3,11,19,29].



**Figure 4.** Discrimination map plotted from the response signals of patients (wild-type and mutant reactions) for each polymorphism.

### 3.2.5 Conclusions

The developed method exploits the advantages of ligation-mediated discrimination as an effective tool to detect single base-pair variations. Nevertheless, its novelty lies in the way that the required detection sensitivity is achieved: an isothermal and universal method. Multiplexing capabilities and integration potentiality (i.e., combination to portable instruments) are higher than other PCR-mediated methods.

Regarding the health impact, the assay allows a screening DNA diagnostic and opens up new avenues of personalised medicine from tailored therapies to direct-to-patient approaches. Given its excellent versatility, the method can be useful for the simultaneous detection of a large number of single base-pair polymorphisms or mutations. The recommendation for effective development is a conscientious design of oligonucleotides to guarantee the specific ligation, amplification and hybridisation of targeted variations.

Finally, the integrated platform fulfils the demanded requirements for next-generation genetic testing devices. The resulting system is robust, simple, sensitive and cost-effective, and can be easily converted into a biosensing device for onsite detection with limited laboratory infrastructure.

### Acknowledgements

The authors acknowledge the financial support received from the Generalitat Valenciana (GRISOLIA/2014/024 PhD Grant and GVA-FPI-2017 PhD Grant) and the Spanish Ministry of Economy and Competitiveness (MINECO Project CTQ2016-75749-R).

### 3.2.6 References

- [1] M. V. Relling, W.E. Evans, Pharmacogenomics in the clinic, *Nature*. 526 (2015) 343–350. doi:10.1038/nature15817.
- [2] S.A. Scott, Clinical pharmacogenomics: opportunities and challenges at point of care, *Clin. Pharmacol. Ther.* 93 (2013) 33–35. doi:10.1038/clpt.2012.196.
- [3] W. Shen, Y. Tian, T. Ran, Z. Gao, Genotyping and quantification techniques for single-nucleotide polymorphisms, *TrAC Trends Anal. Chem.* 69 (2015) 1–13. doi:10.1016/j.trac.2015.03.008.
- [4] A.A. Gibriel, O. Adel, Advances in ligase chain reaction and ligation-based amplifications for genotyping assays: detection and applications, *Mutat. Res. Rev. Mutat. Res.* 773 (2017) 66–90. doi:10.1016/j.mrrev.2017.05.001.
- [5] J.P. Schouten, C.J. Mcelgunn, R. Waaijer, D. Zwiijnenburg, F. Diepvens, G. Pals, Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification, *Nucleic Acids Res.* 30 (2002) e57. www.mrc-holland.com. (accessed February 17, 2019).
- [6] H.-Q. Wang, W.-Y. Liu, Z. Wu, L.-J. Tang, X.-M. Xu, R.-Q. Yu, J.-H. Jiang, Homogeneous label-free genotyping of single nucleotide polymorphism using ligation-mediated strand displacement amplification with DNAzyme-based chemiluminescence detection, *Anal. Chem.* 83 (2011) 1883–1889. doi:10.1021/ac200138v.
- [7] X. Chen, A. Ying, Z. Gao, Highly sensitive and selective colorimetric genotyping of single-nucleotide polymorphisms based on enzyme-amplified ligation on magnetic beads, *Biosens. Bioelectron.* 36 (2012) 89–94. doi:10.1016/j.bios.2012.03.045.

- [8] J. Ritari, J. Hultman, R. Fingerroos, J. Tarkkanen, J. Pullat, Detection of human papillomaviruses by polymerase chain reaction and ligation reaction on universal microarray, *PLoS One*. 7 (2012) e34211. doi:10.1371/journal.pone.0034211.
- [9] Y. Sun, X. Lu, F. Su, L. Wang, C. Liu, X. Duan, Z. Li, Real-time fluorescence ligase chain reaction for sensitive detection of single nucleotide polymorphism based on fluorescence resonance energy transfer, *Biosens. Bioelectron.* 74 (2015) 705–710. doi:10.1016/j.bios.2015.07.028.
- [10] L.A. Tortajada-Genaro, R. Niñoles, S. Mena, Á. Maquieira, Digital versatile discs as platforms for multiplexed genotyping based on selective ligation and universal microarray detection, *Analyst*. 144 (2019) 707–715. doi:10.1039/C8AN01830H.
- [11] M.C. Giuffrida, G. Spoto, Integration of isothermal amplification methods in microfluidic devices: recent advances, *Biosens. Bioelectron.* 90 (2017) 174–186. doi:10.1016/j.bios.2016.11.045.
- [12] J. Pickering, A. Bamford, V. Godbole, J. Briggs, G. Scozzafava, P. Roe, C. Wheeler, F. Ghouze, S. Cuss, Integration of DNA ligation and rolling circle amplification for the homogeneous, end-point detection of single nucleotide polymorphisms, *Nucleic Acids Res.* 30 (2002) e60. [https://watermark.silverchair.com/gnf060.pdf?token=AQECAHi208BE49Ooan9kKhW\\_Ercy7Dm3ZL\\_9Cf3qfKAc485ysgAAAj4wggI6BgkqhkiG9w0BBwaggIrlMIICJwIBADCCAIAGCSqGS1b3DQEHATAeBgIghkqgBZQMEAS4wEQQMT1RpJ07Xt7qzxXHsAgEQgIIB8SGf33Vy1MFWRe2\\_YsQfNinuH5SJaV6\\_Amh8e6KmGgmuJEGF](https://watermark.silverchair.com/gnf060.pdf?token=AQECAHi208BE49Ooan9kKhW_Ercy7Dm3ZL_9Cf3qfKAc485ysgAAAj4wggI6BgkqhkiG9w0BBwaggIrlMIICJwIBADCCAIAGCSqGS1b3DQEHATAeBgIghkqgBZQMEAS4wEQQMT1RpJ07Xt7qzxXHsAgEQgIIB8SGf33Vy1MFWRe2_YsQfNinuH5SJaV6_Amh8e6KmGgmuJEGF) (accessed February 17, 2019).
- [13] H.Y. Heo, S. Chung, Y.T. Kim, D.H. Kim, T.S. Seo, A valveless rotary microfluidic device for multiplex point mutation identification based on ligation-rolling circle amplification, *Biosens. Bioelectron.* 78 (2016) 140–146. doi:10.1016/j.bios.2015.11.039.
- [14] W. Du, M. Lv, J. Li, R. Yu, J. Jiang, A ligation-based loop-mediated isothermal amplification (ligation-LAMP) strategy for highly selective microRNA detection, *Chem. Commun.* 52 (2016) 12721–12724. doi:10.1039/c6cc06160e.
- [15] O. Piepenburg, C.H. Williams, D.L. Stemple, N.A. Armes, DNA detection using recombination proteins, *PLoS Biol.* 4 (2006) e204. doi:10.1371/journal.pbio.0040204.
- [16] M.C. Giuffrida, G. Spoto, Integration of isothermal amplification methods in microfluidic devices: Recent advances, *Biosens. Bioelectron.* 90 (2017) 174–186. doi:10.1016/j.bios.2016.11.045.
- [17] Y. Zhao, F. Chen, Q. Li, L. Wang, C. Fan, Isothermal Amplification of Nucleic Acids, *Chem. Rev.* 115 (2015) 12491–12545. doi:10.1021/acs.chemrev.5b00428.
- [18] Z.A. Crannell, B. Rohrman, R. Richards-Kortum, Equipment-free incubation of recombinase polymerase amplification reactions using body heat, *PLoS One*. 9 (2014) 1–7. doi:10.1371/journal.pone.0112146.
- [19] I.M. Lobato, C.K. O’Sullivan, Recombinase polymerase amplification: basics, applications and recent advances, *Trends Anal. Chem.* 98 (2018) 19–35. doi:10.1016/j.trac.2017.10.015.
- [20] S. Santiago-Felipe, L.A. Tortajada-Genaro, R. Puchades, A. Maquieira, Recombinase polymerase and enzyme-linked immunosorbent assay as a DNA amplification-detection strategy for food analysis, *Anal. Chim. Acta.* 811 (2014) 81–87. doi:10.1016/j.aca.2013.12.017.
- [21] G. Choi, J.H. Jung, B.H. Park, S.J. Oh, J.H. Seo, J.S. Choi, D.H. Kim, T.S. Seo, A centrifugal direct recombinase polymerase amplification (direct-RPA) microdevice for multiplex and real-time identification of food poisoning bacteria, *Lab Chip.* 16 (2016) 2309–2316. doi:10.1039/c6lc00329j.
- [22] S. Martorell, S. Palanca, Á. Maquieira, L.A. Tortajada-Genaro, Blocked recombinase polymerase amplification for mutation analysis of PIK3CA gene, *Anal. Biochem.* 544 (2018) 49–56. doi:10.1016/j.ab.2017.12.013.
- [23] C.X. Li, Q. Pan, Y.G. Guo, Y. Li, H.F. Gao, D. Zhang, H. Hu, W.L. Xing, K. Mitchelson, K. Xia, P. Dai, J. Cheng, Construction of a multiplex allele-specific PCR-based universal array (ASPUA) and its application to hearing loss screening, *Hum. Mutat.* 29 (2008) 306–314. doi:10.1002/humu.20622.
- [24] Y. Guo, J. Cheng, P. Wang, J. Guo, X. Ding, Q. Dong, Y. Jiang, Development of multiplex reverse transcription-ligase detection reaction-polymerase chain reaction (MRLP) mediated universal DNA microarray for diagnostic platform, *Biosens. Bioelectron.* 26 (2011) 3719–3724. doi:10.1016/j.bios.2011.02.027.
- [25] A. Alfirevic, M. Pirmohamed, Genomics of adverse drug reactions, *Trends Pharmacol. Sci.* 38 (2017) 100–109. doi:10.1016/j.tips.2016.11.003.
- [26] A.S. Tseng, R.D. Patel, H.E. Quist, A. Kekic, J.T. Maddux, C.B. Grilli, F.E. Shamoun, Clinical review of the pharmacogenomics of direct oral anticoagulants, *Cardiovasc. Drugs Ther.*

32 (2018) 121–126. doi:10.1007/s10557-018-6774-1.

[27] S. Santiago-Felipe, L.A. Tortajada-Genaro, R. Puchades, Á. Maquieira, Parallel solid-phase isothermal amplification and detection of multiple DNA targets in microliter-sized wells of a digital versatile disc, *Microchim. Acta.* 183 (2016) 1195–1202. doi:10.1007/s00604-016-1745-3.

[28] L.A. Tortajada-Genaro, S. Santiago-Felipe, M. Amasia, A. Russom, Á. Maquieira, Isothermal solid-phase recombinase polymerase amplification on microfluidic digital versatile discs (DVDs), *RSC Adv.* 5 (2015) 29987–29995. doi:10.1039/c5ra02778k.

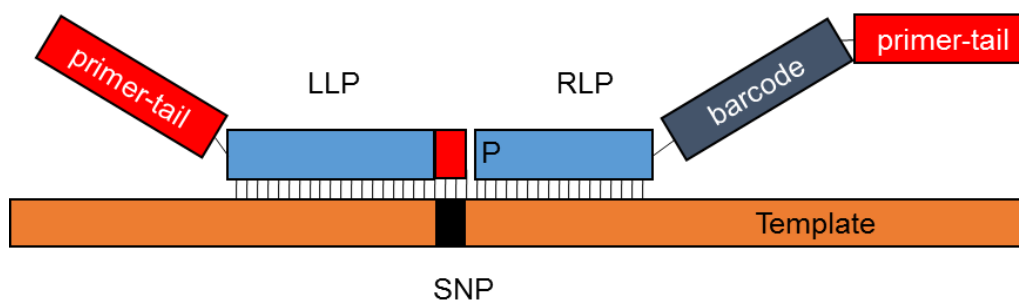
[29] E.E. Te Hwu, A. Boisen, Hacking CD/DVD/Blu-ray for biosensing, *ACS Sensors.* 3 (2018) 1222–1232. doi:10.1021/acssensors.8b00340.

[30] M.E. Natoli, B.A. Rohrman, C. De Santiago, G.U. van Zyl, R.R. Richards-Kortum, Paper-based detection of HIV-1 drug resistance using isothermal amplification and an oligonucleotide ligation assay, *Anal. Biochem.* 544 (2018) 64–71. doi:10.1016/j.ab.2017.12.008.

### 3.2.7 Supplementary information

#### Assay principle. Design recommendations

The assay combines a highly selective ligation process and a common amplification of the generated products. The key element of the developed method is the inclusion of coding tails in the ligation probes for RPA-based amplification (Figure SI.1). In case of multiplexed reactions, the detection can be performed by a selective hybridization. A barcode is used to address products to the specific probes attached to the array substrate.



**Figure SI.1.** Structure of ligation probes. LLP: left ligation probe. RLP: right ligation probe

The design specifications are:

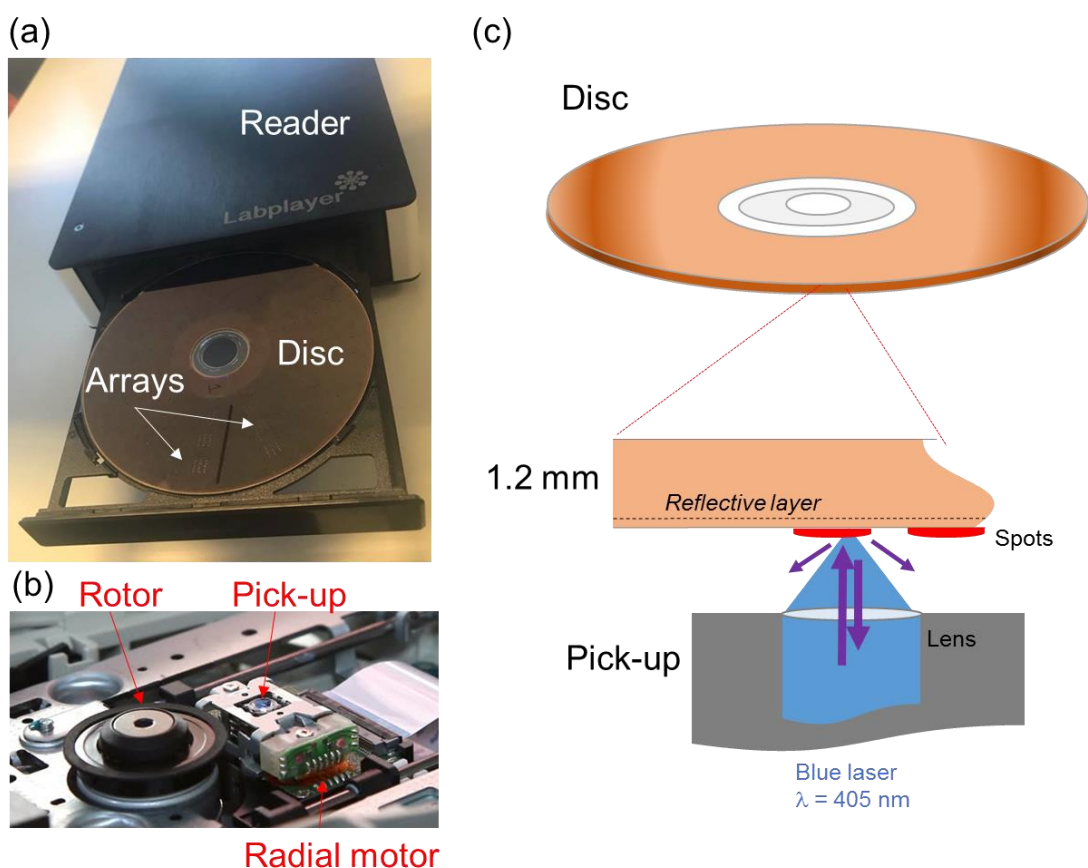
- The left and right ligation probes must be selected based on the position respect to the polymorphism location (upstream or downstream) in a single strand.
- Two left ligation probes are required per each target hotspot. The 3'-end is the interrogating nucleotide (wild-type or mutant).
- The 5'-tail of left ligation probe corresponds to the sequence for universal isothermal amplification.
- The right ligation probe (RLP) contains a 5'-phosphate group. The 3'-tail is the sequence for universal isothermal amplification.
- All these sequences must show no homology to any human genome regions, thus avoiding false signals due to mismatch interactions.
- The formation of stable secondary structures must be minimized.

- Cross-reaction between ligation probes must be avoided.
- Perfect-match hybrids must be stable enough for ligation, amplification and detection.

In case of multiplexed assays, the ligation probe must include a barcode for each target polymorphism or mutation. The recommended position is inserted between the specific region and the amplification tail.

### Assay principle. Simultaneous detection of several targets

The multiplexed detection of ligation-RPA products was performed using solid-phase hybridization assay based on Blu-Ray technology (**Figure SI.2**). The reading was based on light-reflection principle for products anchored in an array format. This system took advantages from the striking characteristics to use this consumer electronic device for chemical and biochemical sensing compared to the conventional fluorescence scanners [1]. High quality mass-produced discs have unique properties for enabling effective immobilization of DNA probes.



**Figure SI.2.** Sensing of hybridization assay in array format based on Blu-ray technology: a) Image of reader and disc with 36 arrays (5×3 spots per array). b) Image of main reader components for disc rotation and laser scanning c) Scheme of optical sensing. The presence of a spot modifies the light intensity, yielding a reduction of reflected intensity, related to the target concentration.



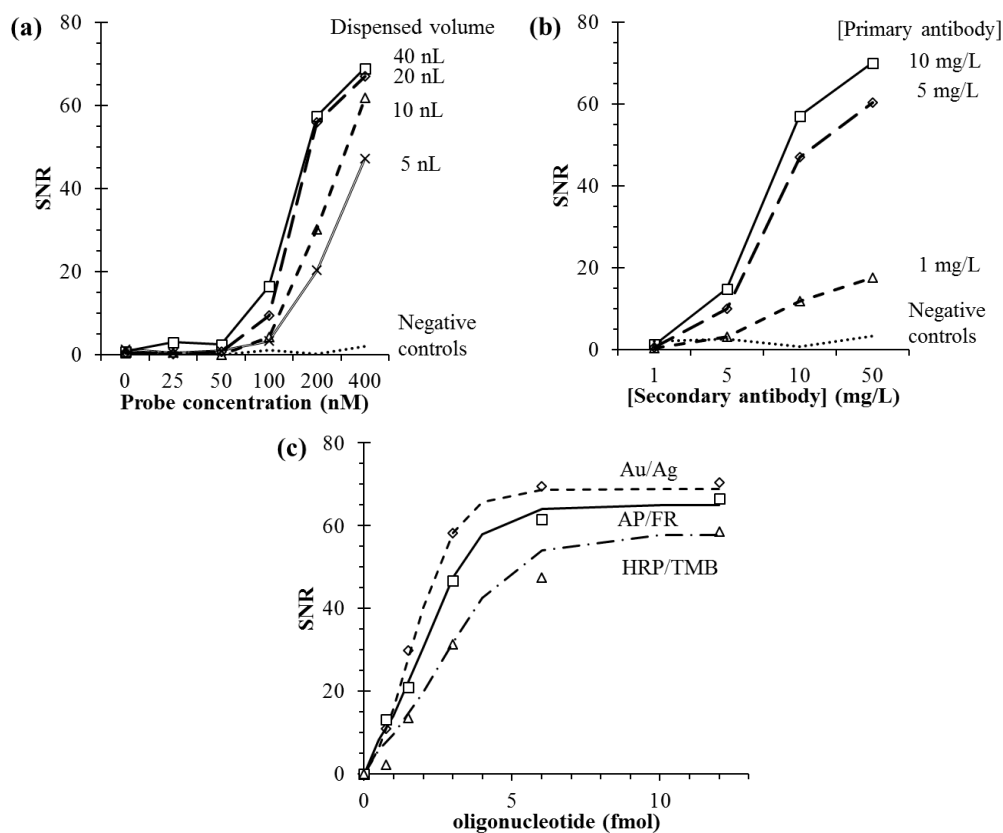
In this research, the immobilization of probes on BD was performed via physisorption, mainly through hydrophobic interactions. Briefly, each biotinylated probe in printing buffer (10 mg/L streptavidin, 50 mM carbonate buffer, pH 9.6) was arrayed on Blu-ray discs. The equipment was a non-contact printer (AD 1500 BioDot Inc., CA, USA) at relative humidity of 90% and room temperature. In this strategy, the immobilization yields were low (0.5–1 fmol/mm<sup>2</sup>). On the contrary, simplicity and avoiding the blocking step were confirmed as the main advantages. Optimal immobilization conditions of biotin-labelled oligonucleotides/streptavidin were experimentally selected on the basis on spot signal intensities. Serial dilutions of probes (up to 400 nM) were arrayed on the disc varying the dispensed volume (5–40 nL), yielding high signal-to-noise responses (**Figure SI.3a**). The spot diameters (400±10 µm) were suitable considering the optical features of the reader, since the laser wavelength ( $\lambda = 405$  nm) and numerical aperture of the objective lens (NA = 0.85) led to a focal spot of 250 nm-diameter on the disc surface. Thus, the optical detector showed a great precision to focus on small spots, allowing for retrieving more digital information (measurements per spot) in a fast-response unequivocal reading.

On the other hand, prior to hybridization on the microarray, the ligated-amplified products were denatured to prevent double stranded molecules that could reduce efficient hybridization onto the microarray. Compared to the initial studies based on PCR assays performed in planar plastic chips, a non-ionic surfactant (Tween 20 0.5%) was also added to the hybridization buffer in order to reduce the background signal. On the other hand, the barcode array led to few restrictive buffer composition (formamide 20 %) and few aggressive washing conditions. Under the selected conditions, RPA products were specifically recognized in short incubation (45 min).

For a reading based on reflection principle, a colorimetric staining of DNA complexes formed on the Blu-Ray disc surface was selected. The strategy consisted of labelling the products during RPA with digoxigenin, a popular hapten and common functionalization of nucleotides. Latter, an immunostaining approach based on anti-digoxigenin rabbit antibody (primary) and conjugated anti-rabbit antibody (secondary) was examined (**Figure SI.3b**). The resulting conditions were similar to the studies performed in other optical discs [2]. Nevertheless, the percentage of detergent required in Blu-ray approach for obtaining low backgrounds was ten-folds higher (0.5 %) than the needed for other compact discs. Three conjugation systems were compared: (a) horseradish peroxidase combined to colorimetric substrate TMB, (b) alkaline phosphatase combined to colorimetric substrate Fast-red and (c) gold nanoparticle combined to silver developer (**Figure SI.3c**).

The Blu Ray drive used in this study was from LG Electronics Inc. (Englewood Cliffs, NJ, USA), and connected to the laptop through a USB2.0 universal serial bus interface. The generated insoluble products for perfect-match probes modified the reflected beam detected by the pickup photo-diode. The attenuated analogic signals were directly acquired from the photodiode of the BD drive and related to the amount of the target analyte. Higher responses were achieved using gold nanoparticle-labelled antibody/silver salt staining based on the light scattering of silver deposition. However, enzymatic conjugates yielded insoluble deposits with higher reproducibility and lower background signal (light absorption and scattering).

The selected reagents were monoclonal rabbit antidigoxigenin-antibody (Invitrogen, CA, USA) at 5 mg/L and goat antirabbit-antibody labelled with alkaline phosphatase (Abcam, UK) at 10 mg/L in phosphate buffered saline solution containing 0.5 % Tween-20 (PBST). The colorimetric substrate was 1 mL of Fast Red/naphtol (4-chloro-2-methylbenzenediazonium/3-hydroxy-2-naphthoic acid 2,4-dimethylanilide phosphate) (Sigma, MI, USA). The protocol was incubation with conjugated-antibodies (anti-digoxigenin) at room temperature for 20 min, washing with PBST and water, and incubation with colorimetric substrate (room temperature, 10 min, dark).



**Figure SI.3.** Set-up of array detection based on signal-to-noise ratio (SNR): a) Probe concentration and dispensed volume; b) Antibody concentrations (primary and secondary). c) staining method. HRP/TMB: Horseradish peroxidase system. AP/FR: Alkaline phosphatase system. Au/Ag: Nanoparticle system.

The extrapolation of the proposed detection methodology to other analytical platforms or biosensing systems based on DNA hybridization was studied. The same assay was also performed on conventional plastic chips (polycarbonate Makrolon, 25×75 mm, 12 arrays per chip, 3 chip replicates). Later, a conventional scanner (Perfection 1640SU Office, Epson) imaged the generated spots. The results were comparable in terms of spot morphology (circle-shape), spot homogeneity (standard deviation < 20%) and response-noise ratios (SNR  $\approx$  60).

### **Proof-of-concept application. Pharmacogenomic information**

The discrimination capabilities of single nucleotide changes were studied focusing on a relevant pharmacogenomics application in a high-impact disease. Identifying individuals who are at risk for poor therapeutic responses may be useful for amending treatment decisions to optimize benefit and minimize harm [3].

The most widely prescribed anticoagulant drugs worldwide are warfarin, acenocoumarol and phenprocoumon. They are used to prevent thromboembolic diseases in patients with deep vein thrombosis, atrial fibrillation, recurrent stroke or heart valve prosthesis. These coumarin-type drugs show similar mechanisms of action because they inhibit the protein called vitamin K epoxide reductase complex, subunit 1 (VKORC1). However, inappropriate dosing has been associated to a substantial risk of both major and minor hemorrhage.

Many studies have demonstrated that some gene genotypes help to predict the coumarin-type drug dose. Accounting for genetic variation, age, height, body weight, interacting drugs, and indication for warfarin therapy explained the variability in warfarin dose.

The variants rs1799853, rs1057910 and rs9923231 have been associated to dosage of coumarin-type drugs with high clinical annotation levels of evidence (**Table SI.1**). In fact, Clinical Pharmacogenetics Implementation Consortium (CPIC) and other medical societies have published dosing guidelines taking into consideration patient genotype for those variants.

**Table SI.1.** Variants associated to cardiovascular diseases. Source. PharmGKB database.

Variant	Gene	Change	Level	Molecule	Type
rs1799853	<i>CYP2C9*2</i>	C>T	Level 1A	Warfarin	Dosage
rs1057910	<i>CYP2C9*3</i>	A>G	Level 1A	Warfarin	Dosage/toxicity
rs9923231	<i>VKORC1</i>	C>T	Level 1A	Warfarin	Dosage
			Level 1B	Acenocoumarol Phenprocoumon	Dosage
rs7294	<i>VKORC1</i>	C>T	Level 1B	Warfarin	Dosage
rs9934438	<i>VKORC1</i>	G>A	Level 1B	Warfarin	Dosage
rs2108622	<i>CYP4F2</i>	C>T	Level 1B	Warfarin	Dosage

*CYP2C9 gene: cytochrome P450, family 2, subfamily C, polypeptide 9*

### Proof-of-concept application. Selection of oligonucleotides

The design goal was the discrimination of three polymorphisms: rs1799853 (*CYP2C9\*2*, g.8633C>T), rs1057910 (*CYP2C9\*3*, g.47639A>C) and rs9923231 (*VKORC1*, g.3588C>T). A design workflow was employed for the selection of candidate sequences [1]. The algorithm includes several calculations for evaluating the thermodynamical properties of candidate oligonucleotides. For instance, the melting temperature of possible secondary structures (e.g. hairpins, self-annealing) and unspecific duplexes were computed. Furthermore, the algorithm selected the barcode sequences that are compatible with the designed sets.

Due to RPA mechanism, the length of tail was especially studied. Conventional RPA-based methods employ a longer primer compared to PCR for the formation of recombinase-template-primed complexes [4]. The recommended length is usually between 30 and 35 nucleotides. However, the inclusion of these long sequences in the ligation probes produced the formation of stable secondary structures (i.e. hairpin). In previous investigations, we have concluded that shorter sequences can be used for RPA-based methods, controlling the design and conditions [5]. Therefore, the selection of shorter tails (18 nucleotides) yielded ligation probes with low stability for secondary structures under reaction conditions. The estimated variation of free-energy was  $-2.5 \pm 1.5$  kcal/mol.

Table SI.2.2 shows the oligonucleotides used for the ligation, RPA-based amplification and detection assays. All oligonucleotides were purchased from Eurofins (Luxembourg).

Single experiments (1 gene per assay) were performed for confirming the correct design. Synthetic ligation products were prepared, containing the candidate forward primer and reverse primer sequences at the extremes 5' and 3', respectively. Then, these products (1333 copies) were amplified (RPA reaction) using the complementary primers

(18 nucleotides). Positive band of agarose electrophoresis were detected for candidate oligonucleotides, the length of amplification products being 99 bp, 100 bp and 103 bp, respectively. Therefore, these assays concluded that the amplification of ligation products using an isothermal technique was feasible, even selecting short RPA tails.

**Table SI.2.** List of tested oligonucleotides.

Gen	Function	Sequence (5'-3')	Length (nt)	%GC	Tm (°C)
<b>Ligation</b>					
<i>CYP2C9*2</i>	LLP-A	ACTTCGTCAGTAACGGAC- CGGGCTTCCTCTTGAACACA	18 20	50 55	53.8 60.5
<i>rs1799853</i>	LLP-B	GAGTCGAGGTCATATCGT- CGGGCTTCCTCTTGAACACG	18 20	50 60	53.8 62.5
		[P]-GTCCTCAATGCTCCTCTTCCC- CGTTCTAGCCTAACCGCCTTGA-	21 22	57 55	63.2 64.2
	RLP	GTCTGCCTATAGTGAGTC	18	50	53.8
	<i>CYP2C9*3</i>	LLP-A	ACTTCGTCAGTAACGGAC- GCTGGTGGGGAGAAGGTCAAT	18 21	50 57
<i>rs1057910</i>	LLP-B	GAGTCGAGGTCATATCGT- GCTGGTGGGGAGAAGGTCAAG	18 21	50 62	53.8 65.3
		[P]-GTATCTCTGGACCTCGTGAC- GCGATTCATAGACCCGTTTCCG-	21 22	57 55	63.2 64.2
	RLP	GTCTGCCTATAGTGAGTC	18	50	53.8
	<i>VKORC1</i>	LLP-A	ACTTCGTCAGTAACGGAC- AGACCTGAAAAACAACCATTGGCCA	18 25	50 44
<i>rs9923231</i>	LLP-B	GAGTCGAGGTCATATCGT- AGACCTGAAAAACAACCATTGGCCG	18 25	50 48	53.8 65.8
		[P]-GGTGCGGTGGCTCACGCCTA- ATAGCAGGTTACGGTCCGACGA-	20 22	70 55	66.6 64.2
	RLP	GTCTGCCTATAGTGAGTC	18	50	53.8
	<b>Amplification</b>				
	U-FP-A	ACTTCGTCAGTAACGGAC	18	50	53.8
	U-FP-B	GAGTCGAGGTCATATCGT	18	50	53.8
	U-RP	GACTCACTATAGGCAGAC	18	50	53.8
<b>Hybridization</b>					
<i>CYP2C9*2</i>	u-probe	[Btntg]TTTTTTTTTTT- TCAAGGCGGTTAGGCTAGAACG	22	55	64.2
<i>CYP2C9*3</i>	u-probe	[Btntg]TTTTTTTTTTT- CGGAAACGGGTCTATGAATCGC	22	55	64.2
<i>VKORC1</i>	u-probe	[Btntg]TTTTTTTTTTT- TCGTCCGACCGTAACCTGCTAT	22	55	64.2

LLP: left ligation probe; RLP: right ligation probe; U-FP: universal forward primer; U-RP: universal reverse primer; p: probe; [P]: 5'-phosphate group; [Btntg]: biotin TEG; Tm: melting temperature.

For a multiplex reaction, the candidate oligonucleotides do not have to show regions of homology. The partial complementarities were estimated calculating the

variation of free energy related to the cross-hybridization of ligation probes (**Table SI.3**). Perfect-match complexes were more stable than nonspecific complexes, the variation of 16.8-28.0 kcal/mol being for inter-targets. The genotyping capability was confirmed because the estimated variation was 0.2-2.1 kcal/mol for single nucleotide mismatch. Regarding to RPA primers (18 nucleotides), the estimated free energies were -22.5 kcal/mol (forward) and -20.5 kcal/mol (reverse). Finally, the stability variations between perfect-match and mismatch barcode probes were 23.9-25.9 kcal/mol. In conclusion, the estimated values were suitable for developing a selective multiplexed assay.

**Table SI.3.** Cross-hybridization calculations of ligation probes, expressed as variation of free energy for template-probe complexes ( $\Delta G$ , kcal/mol). WT: wild-type, M: mutant.

Template		Ligation probes					
		CYP2C9*2		CYP2C9*3		VKORC1	
		WT	M	WT	M	WT	M
CYP2C9*2	WT	<u>-26.4</u>	-25.7	-5.8	-4.7	-9.0	-5.1
	M	-26.2	<u>-27.8</u>	-8.2	-5.7	-8.2	-4.7
CYP2C9*3	WT	-6.9	-11.0	<u>-28.0</u>	-27.5	-7.8	-10.0
	M	-6.9	-11.0	-26.8	<u>-28.9</u>	-7.8	-10.7
VKORC1	WT	-8.6	-7.7	-9.7	-9.9	<u>-31.6</u>	-31.1
	M	-8.6	-7.7	-10.8	-9.7	-31.3	<u>-33.1</u>

### Proof-of-concept application. Preparation of controls

**Ligation.** Human genomic DNAs was amplified using a specific forward primers (AGTAAACAAAGCCCCTGC) and reverse primer (TTTGGGGCGATGAGTGT) for *ZFX* gene. The reaction mixture contained Taq polymerase at 0.04 U/ $\mu$ L, MgCl<sub>2</sub> at 5 mM and dNTPs at 200  $\mu$ M in amplification buffer (Biotools). The cycling conditions were 35 cycles of denaturation at 90 °C for 30 s, primer annealing at 60 °C for 30 s and elongation at 72 °C for 30 s. Then, a dilution of this product (10<sup>6</sup>) was added in the ligation control vial instead of genomic DNA. Ligation oligonucleotides were ACTTCGTCAGTAACGGAC-AGTAAACAAAGCCCCTGCATGAGAA, GAGTCGAGGTCATATCGT-AGTAAACAAA GCCCCTGCATGAGAG and [P]-AAATGCTTCCCACACTCATCGCCCCAAA-CAAACGTA CCAAGCCCGCGTCG-GTCTGCCTATAGTGAGTC. Therefore, a positive result indicates the ligation was correctly produced, while a negative indicates a fail during the ligation.

Four negative ligation assays were also prepared. Ligation controls 1 and 2 were incomplete reaction mixtures, lacking left or right ligation probe, respectively. In controls 3 and 4, no template or non-human DNA (bacterial genomic DNA) was employed. Later, the products obtained were amplified using universal RPA mixture.

**Amplification.** Six amplification controls were generated for each studied variant in targeted genes (**Table SI.4**). Genomic DNAs (previously genotyped as homozygotes) were amplified using the corresponded allele-specific forward primer (including universal amplification tail) and the reverse primer (including both hybridization and universal amplification tails). Since the mutant homozygotes are non-registered variants for *CYP2C9* gene in Hapmap database (population frequency 0 %), synthetic DNA solutions were prepared from single strand DNA (mutant variant). The amplification conditions were as described above. The results were homozygote products with a length of 111 pb for rs1799853, 139 pb for rs1057910, and 129 pb for rs9923231. Then, dilutions of these products (factor 10<sup>6</sup>) were added in the control vials instead of ligation products (six vials). The amplification-hybridization protocol was applied as vials for patient samples. Therefore, a positive result indicates that the assay was correctly produced and a selective genotyping, while a false-negative or false-positive indicates a fail during the assay.

**Detection.** The immobilization control was a doubled-labelled probe (digoxigenin-T10-GTCATGGGCCTCGTGTCGGAAA ACC-biotin TEG). The hybridization control was composed by an amplification product of *ACTB* gene and a specific probe (biotin TEG-T10-CAACCGCGAGAAGATGACCCAGATCA). This product was generated following the protocol described above, the primers being GCACCACACC TTCTACAATGAG and digoxigenin-GGCCACCAGAAGAGGTAGC. Therefore, a positive result indicates that the immobilization-hybridization and array staining were correctly produced.

**Table SI.4.** List of oligonucleotides for RPA controls. Genotypes were determined by Sanger sequencing

Gen	Function	Sequence (5'-3')	Length (nt)	%GC	Tm (°C)
<b>Universal RPA controls</b>					
<i>CYP2C9*2</i>	WT template	Homozygote genomic DNA (WT-genotyped)			
	M template	AATTTTGGGATGGGGAAGAGGAGCATTGAG GACTGTGTTCAAGAGGAAGCCCG	53	51	84.3
	FP-WT	GAGTCGAGGTCATATCGT- CGGGCTTCCTCTTGAACACG	38	55	80.0
	FP-M	ACTTCGTCAGTAACGGAC- CGGGCTTCCTCTTGAACACA GACTCACTATAGGCAGAC-	38	53	78.9
<i>CYP2C9*3</i>	RP	TCAAGGCGGTTAGGCTAGAACG- AATTTTGGGATGGGGAAGAG	60	50	85.7
	WT template	Homozygote genomic DNA (WT-genotyped)	-	-	-
	M template	ATGCAAGACAGGAGCCACATGCCCTACACA GATGCTGTGGTGCACGAGGTCCAGAGATAC CTTGACCTTCTCCCCACCAGC	81	57	92.1
	FP-WT	ACTTCGTCAGTAACGGAC- GCTGGTGGGGAGAAGGTCAAT	39	54	80.0
<i>VKORC1</i>	FP-M	GAGTCGAGGTCATATCGT- GCTGGTGGGGAGAAGGTCAAG GACTCACTATAGGCAGAC-	39	56	81.0
	RP	CGGAAACGGGTCTATGAATCGC- ATGCAAGACAGGAGCCACAT	60	52	86.4
	WT template	Homozygote genomic DNA (WT-genotyped)	-	-	-
	M template	Homozygote genomic DNA (M-genotyped)	-	-	-
<i>VKORC1</i>	FP-WT	ACTTCGTCAGTAACGGAC- AGACCTGAAAAACAACCATTGGCCA	43	47	78.9
	FP-M	GAGTCGAGGTCATATCGT- AGACCTGAAAAACAACCATTGGCCG GACTCACTATAGGCAGAC-	43	49	79.9
	RP	TCGTCCGACCGTAACCTGCTAT- CTCAAGTGATCCACCCACCT	60	53	87.1

WT: wild-type; M: mutant; FP: forward primer; RP: reverse primer; Tm: melting temperature.



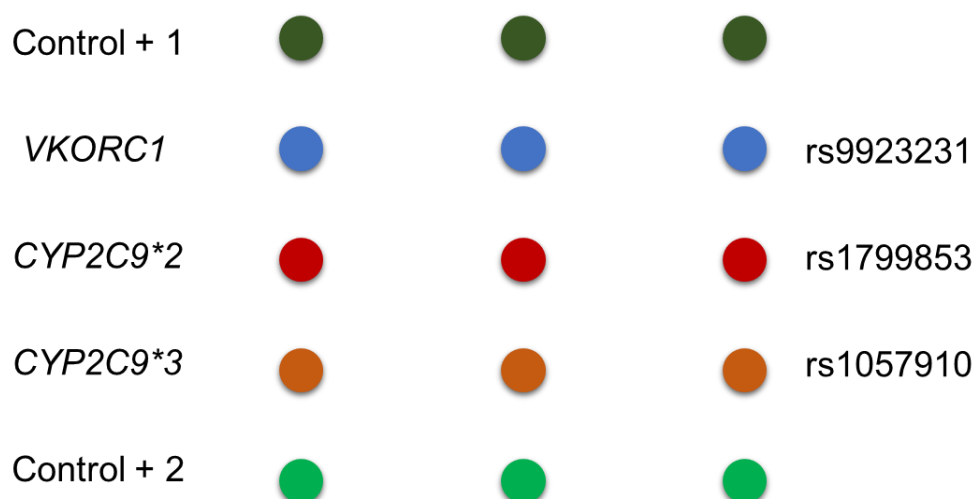
### Proof-of-concept application. Comparison of patient samples

By using the developed methodology, the aim was to obtain the genetic profiles of patients in order to achieve the best drug response and highest safety margins. Then, we evaluated if the method can support the accurate assignation to the specific population group.

DNA extracts were obtained from buccal smear samples, using a Purelink Genomic DNA mini kit (Thermo Fisher Scientific, USA). The purified products were eluted with Tris-HCl buffer (Tris 10 mM, pH 8.6) and had their genomic DNA content quantified with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA).

Then, the developed method was applied for genotyping of targeted polymorphisms. Regarding the detection, the probes of ligation-amplification products were immobilized (**Figure SI.4**). The array layout consisted of 5 dots per row (3 replicates), each with a 1-mm track pitch corresponding to targeted genes and controls. Quality control involved checking the corresponding steps (ligation, amplification and hybridization). Then, the analytical results of a patient sample was declared valid if assay controls provided the expected optical signals (negative or positive). Examples of patient results are shown in **Figure SI.5**. The limitation observed was the need of establishing a discrimination factor due to the residual responses. Including this calculation in the data processing step, an accurate discrimination of wild-type and mutant variants was achieved in automated mode.

The results were compared to those obtained by the reference method based on Sanger sequencing (**Figure SI.6** and **Table SI.5**). The assignation to population groups is shown in **Table SI.6**, confirming that our system correctly typed the polymorphisms.



**Figure SI.4.** Array layout. Positive controls: immobilization (1) and hybridization (2). Target polymorphisms: 3. Replicates: 3



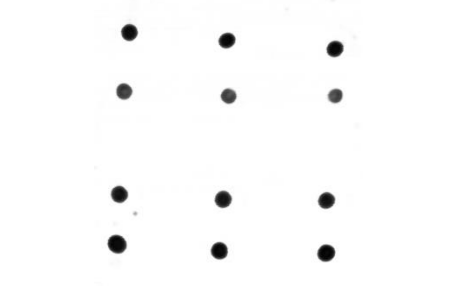

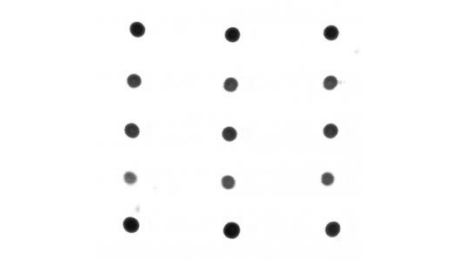

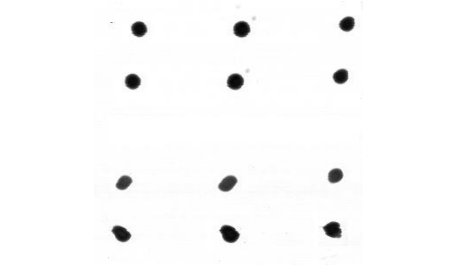
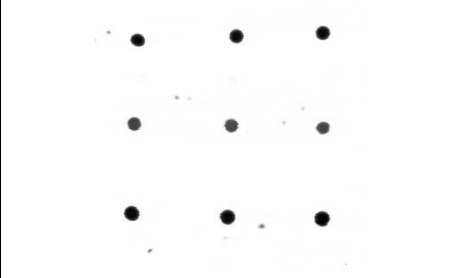

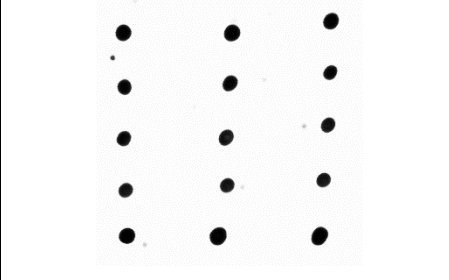
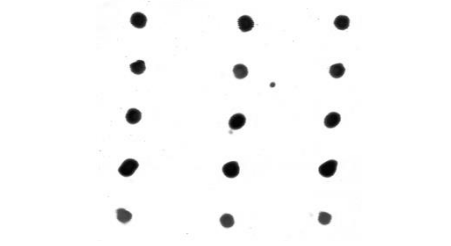

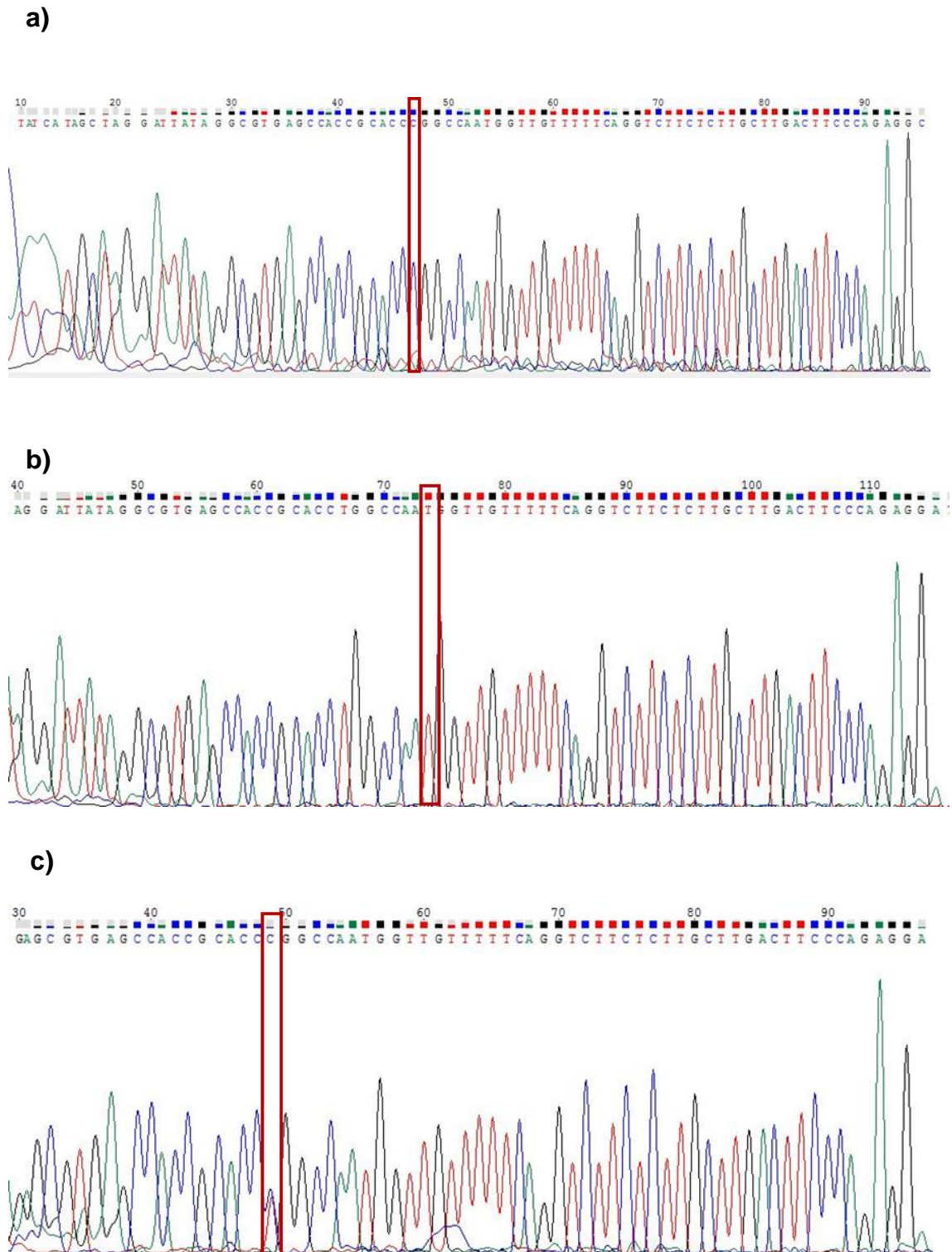
Patient	Mix-A	Mix-B
1		
2		
3		
4		
5		
6		

Figure SI.5. Examples of array images from patient samples.



**Figure SI.6.** Electropherogram obtained in Sanger sequencing for the VKORC1 gene, indicating the SNP position rs9923231. a) Individual 1: homozygous native genotype. b) Individual 2: homozygous mutant genotype. c) Individual 3: heterozygous genotype.

**Table SI.5.** Results for example assays

			Patient					
Mix	Probe		1	2	3	4	5	6
DEVELOPED	C+ 1		+	+	+	+	+	+
METHOD	A	VKORC1 T	-	+	+	+	+	+
		CYP2C9*2 T	-	-	+	-	-	+
	CYP2C9*3 A	+	+	+	+	+	+	
	C+ 2	+	+	+	+	+	+	
	B	C+ 1		+	+	+	+	+
VKORC1 C		+	+	+	-	+	+	
CYP2C9*2 C		+	+	+	+	+	+	
CYP2C9*3 C		-	-	-	-	+	+	
C+ 2		+	+	+	+	+	+	
REFERENCE	VKORC1		CC	CT	CT	TT	CT	CT
METHOD	CYP2C9*2		CC	CC	CT	CC	CC	CT
	CYP2C9*3		AA	AA	AA	AA	AC	AC

WT: Wild-type ligation probe; M-p: Mutant ligation probe

**Table SI.6.** Assignment to genetic populations

Variant	Wild-type	Heterozygote	Mutant
CYP2C9*2	CC	CT	TT
	18 (60%)	12 (40%)	0 (0 %)
CYP2C9*3	AA	AC	CC
	18 (60 %)	11 (37%)	1 (3%)
VKORC1	CC	CT	TT
	11 (37%)	16 (53%)	3 (10%)

## References

- [1] Armandis-Chover, T., Morais, S., González-Martínez, M. Á., Puchades, R., & Maquieira, Á. (2014). High density MicroArrays on Blu-ray discs for massive screening. *Biosensors and Bioelectronics*, 51, 109-114.
- [2] Tortajada-Genaro, L. A., Puchades, R., & Maquieira, Á. (2017). Primer design for SNP genotyping based on allele-specific amplification—Application to organ transplantation pharmacogenomics. *Journal of pharmaceutical and biomedical analysis*, 136, 14-21.
- [3] Joseph, P. G., Pare, G., Ross, S., Roberts, R., & Anand, S. S. (2014). Pharmacogenetics in cardiovascular disease: the challenge of moving from promise to realization: concepts discussed at the Canadian Network and Centre for Trials Internationally Network Conference (CANNeCTIN), June 2009. *Clinical cardiology*, 37(1), 48-56.
- [4] Piepenburg, O., Williams, C. H., Stemple, D. L., & Armes, N. A. (2006). DNA detection using recombination proteins. *PLoS biology*, 4(7), e204.
- [5] Santiago-Felipe, S., Tortajada-Genaro, L. A., Morais, S., Puchades, R., & Maquieira, Á. (2014). One-pot isothermal DNA amplification–Hybridisation and detection by a disc-based method. *Sensors and Actuators B: Chemical*, 204, 273-281.



## **CHAPTER 3:**

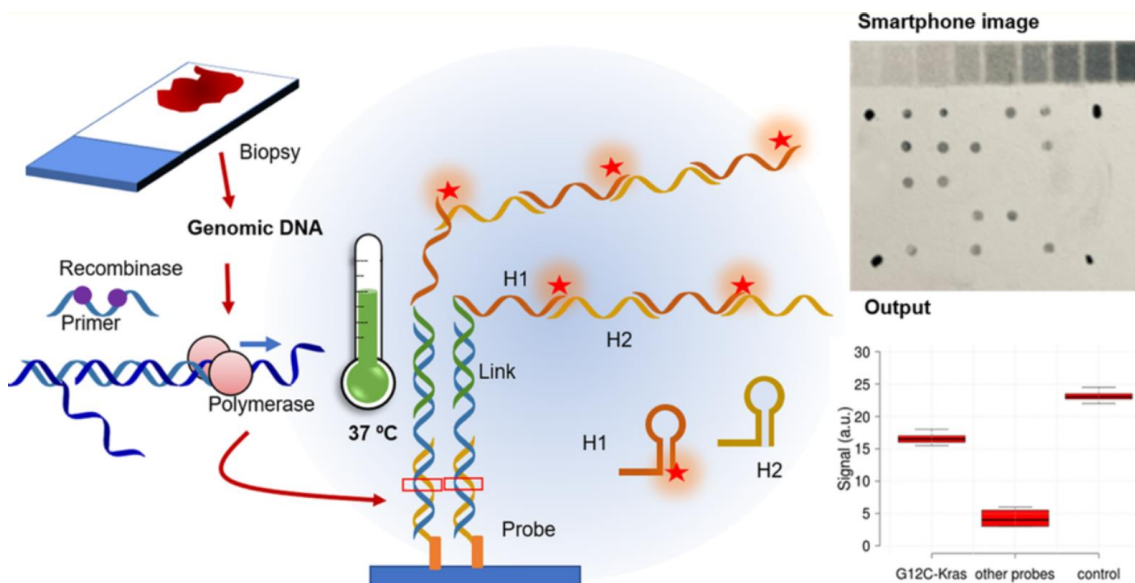
**Discrimination of single nucleotide variants based on allele-specific hybridization chain reaction and smartphone detection**



### 3.3 CHAPTER 3: DISCRIMINATION OF SINGLE NUCLEOTIDE VARIANTS BASED ON ALLELE-SPECIFIC HYBRIDIZATION CHAIN REACTION AND SMARTPHONE DETECTION

This chapter involves an advance regarding the approaches described in previous studies. The research aims to improve the genotyping performance of isothermal amplification methods for application in low-resource settings, minimizing the use of enzymes. To this end, the capabilities of hybridization chain reaction (HCR), a free-enzyme DNA recognition process, are explored to discriminate the mutational status of oncogenes. The main challenge is to apply HCR in an array format to genomic DNA from human clinical samples and identify the SNV present. To overcome the drawbacks of published papers regarding the limited detection limits, we first propose a DNA treatment step. A rapid RPA could allow the generated enough small-sized products to trigger allele-specific HCR (AS-HCR) on-chip to overcome the structural difficulty of gDNA.

After demonstrating the assay principle and determining the experimental conditions, the HCR products are detected by simple colorimetric immunostaining. Another approached challenge is using the smartphone as an analytical reader to achieve a low-cost, ubiquitous, and portable method. This solution simplifies those described in previous chapters of the current thesis. As a proof of concept, the method's capabilities are evaluated by multiplex screening for SNV in the *KRAS* gene (codon 12-13) and *NRAS* gene (codon 61) in human cell cultures and cancer biopsy tissue samples. The nucleotide changes determine matched the results of next-generation sequencing.



Summary figure. Overview chapter 3

## Discrimination of Single-Nucleotide Variants Based on an Allele-Specific Hybridization Chain Reaction and Smartphone Detection

Ana Lázaro, Ángel Maquieira, and Luis A. Tortajada-Genaro\*

 Cite This: *ACS Sens.* 2022, 7, 758–765 Read Online

### 3.3.1 Abstract

Massive DNA testing requires novel technologies to support a sustainable health system. In recent years, DNA superstructures have emerged as alternative probes and transducers. We herein report a multiplexed and highly sensitive approach based on allele-specific hybridization chain reaction (AS-HCR) in the array format to detect single nucleotide variants. Fast isothermal amplification was developed before activating the HCR process on a chip to work with genomic DNA. The assay principle was demonstrated, the variables for integrating the AS-HCR process and smartphone-based detection were also studied. The results were compared to a conventional PCR-based test. The developed multiplex method enabled higher selectivity against single-base mismatch sequences at concentrations as low as  $10^3$  copies with a limit of detection of 0.7 % of the mutant DNA percentage and good reproducibility (relative error: 5 % for intra-assay and 17 % for inter-assay). As proof of concept, the AS-HCR method was applied to clinical samples, including human cell cultures and biopsied tissues of cancer patients. Accurate identification of single nucleotide mutations in *KRAS* and *NRAS* genes was validated, considering those obtained from the reference sequencing method. To conclude, AS-HCR is a rapid, simple, accurate, cost-effective isothermal method that detects clinically relevant genetic variants and has a high potential for point-of-care demands.

**Keywords:** DNA biosensing; single-nucleotide mutation; allele-specific probe; hybridization chain reaction; cancer biomarker genes

### 3.3.2 Introduction

Certain changes in a specific position in the genome sequence, called single nucleotide variations (SNV), are closely associated with genetic diseases and cancer.<sup>1</sup> Therefore, disease-related SNVs serve as biomarkers for efficient clinical diagnosis and prognosis based on the genomic profile of primary tumors.<sup>2</sup> However, the genotyping of



SNVs is difficult due to close molecular similarity and their presence at trace levels.<sup>3</sup> A specific copy number increment is required to detect and avoid inference from high-abundant variants.<sup>4</sup> The most widely used amplification methods are thermocycling techniques, including polymerase reaction chain (PCR)<sup>5</sup> and ligase chain reaction.<sup>6</sup> Despite their good performance, the requirement of sophisticated equipment considerably limits applications for point-of-care testing (POC).<sup>7</sup> Important signs of progress have been achieved in the last years, developing biosensing methods<sup>8-10</sup> and microfluidics-integrated sensors<sup>11</sup> with excellent single-base specificity.

Isothermal techniques are appealing alternative tools that work with simple assays and avoid heating/cooling cycles. The first category is enzyme-based methods, such as loop-mediated amplification (LAMP), helicase-dependent amplification (HDA), and recombinase polymerase amplification (RPA).<sup>12</sup> For instance, their direct integration into consumer electronic devices has allowed the diagnosis of diseases in resource-limited settings.<sup>13,14</sup> Another innovative approach is to use isothermal reactions as a previous step in recognition/detection assays. Here the goal is to achieve the required sensitivity based on the generation of small-sized products that improve access to probes<sup>15</sup> or the biorecognition process, such as nucleic acid detection systems supported on clustered regularly interspaced short palindromic repeats (CRISPR/Cas).<sup>16</sup>

The second group of isothermal techniques are enzyme-free methods, such as toehold-mediated strand displacement amplification<sup>17</sup> and hybridization chain reaction (HCR).<sup>18</sup> In a typical HCR approach, recognizing a target initiates the cross-opening of two DNA hairpins to yield a DNA superstructure of nicked amplified double helices capable of flexibly detecting.<sup>19</sup> Due to its simplicity and excellent efficiency, HCRs have been combined with different molecular reporters and readout detection approaches, enabling the sensitive detection of certain nucleic acids.<sup>20,21</sup> Nevertheless, HCR-based methods remain in the laboratory stage, and their practical applications are still challenging.<sup>22</sup> To date, there have been two critical limitations: (i) given the structural complexity of genomic DNA (gDNA), only short DNA sequences from clinical samples are amplified by HCR, such as microRNAs,<sup>23</sup> short gene sequences,<sup>24</sup> and specific circulating tumor DNA;<sup>25</sup> (ii) none of the reported studies identify the change in a single-nucleotide in gDNA due to the hybridization complexity associated with secondary structures of nucleic acids.<sup>26</sup>

The discrimination of multiple SNVs in gDNA is herein presented. The assay principle consisted of an allele-specific hybridization chain reaction in the array format (AS-HCR), including a previous short RPA. Thus, the double isothermal amplification process is sensitive enough even for a few target copies. By focusing on supporting personalized medicine in low-resource settings, a simple, reliable colorimetric detection

system for POC is also developed. A smartphone was selected for its widespread presence, portability, and capacity to transmit data at a user-friendly interface as a biomedical reader.<sup>27</sup> The study also included the analysis of method requirements and its versatility to identify other important SNV in critical genes for developing diseases.

### **3.3.3 Materials and methods**

**Reagents.** The fabrication of chips is described in the Supplementary Information. The hybridization buffer was sodium citrate 45 mM, NaCl 450 mM, 25% formamide, and 2.5% Denhardt's reagent, adjusted to pH 7. The positive control of allele-specific hybridization (digoxigenin-labeled amplification product of the  $\beta$ -actin gene) was added (1 nM). Washing buffers were dilutions of saline citrate buffer (1:10 and 1:100).

The studied application was the discrimination of SNVs in oncogenes. A set of oligonucleotides (primers, probe, and link) was complementary to a specific region close to each target SNV (Figure SI.1-SI.2). In this study, the targets were the point mutations located at *KRAS* (codons 12 and 13) and *NRAS* (codon 61) genes. The design criteria and the list of the used oligonucleotides, supplied by Eurofins (Luxembourg), are shown in Table SI.1 and SI.2.

**Patients.** Human SK-N-AS cells (brain) with wild-type *KRAS* and HCT116 cells with mutant *KRAS* G13D (colon cancer) were purchased from the American Type Culture Collection (ATCC, USA) and used for method optimization purposes (passage number lower than 4).

The pathologically confirmed metastatic colorectal cancer DNA samples (n=36) were obtained from formalin-fixed paraffin-embedded (FFPE) biopsy tissues of patients from the Oncologic Service of the Hospital Clínico Universitario La Fe (Valencia, Spain). Informed consent was provided by each patient. Genomic DNA was extracted as described in the Supplementary Information.

**RPA.** Fast isothermal amplification was performed using the TwistAmp Basic RPA kit (TwistDx, UK). Reaction mixtures (12.5  $\mu$ L) were prepared with rehydrated buffer, 14 mM of magnesium acetate, 200 nM of each forward and reverse primer, 4 ng of gDNA, and the enzyme pellet. Solutions were incubated at 37 °C for 10 min in an oven (Beschickung- Loading Modell 100-800, Germany).

**AS-HCR process.** RPA products (6  $\mu$ L) were mixed with the hybridization buffer (20  $\mu$ L) and the oligonucleotide solution (4  $\mu$ L), with final concentrations of 100 nM of links and 500 nM of H1 and H2. Solutions were denatured (92 °C, 10 min) for opening the double-strand RPA product and transferred to the chip of 12-assay zones. The chip was incubated in an oven at 37 °C for 30 min and rinsed once with washing buffer for 1 min.

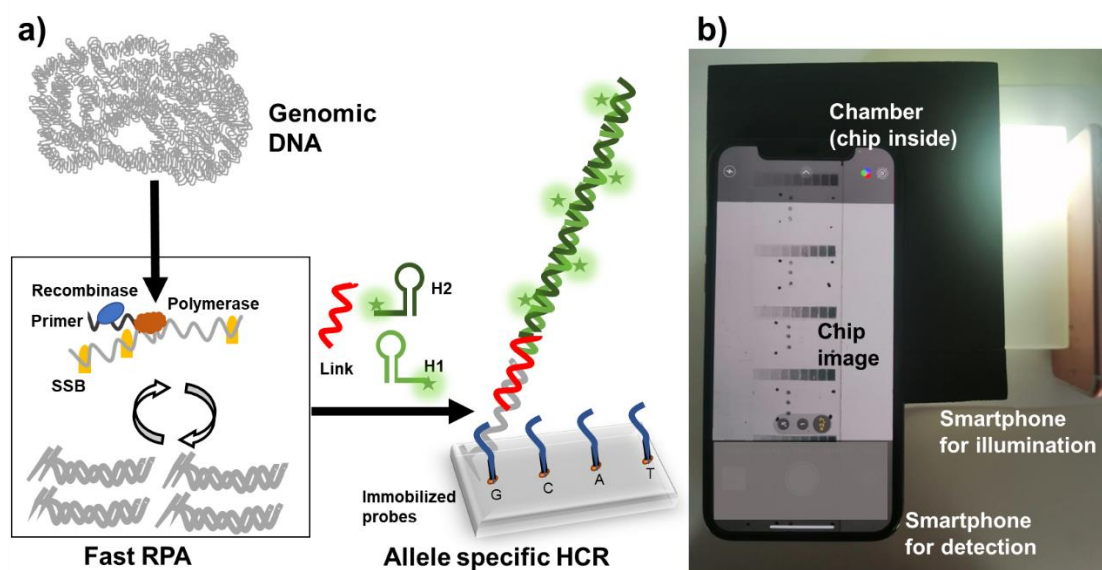
**Detection.** In order to detect the complex formed by the HCR reaction, the chip was stained as described previously.<sup>24</sup> The imaging system was an iPhone 11 Pro's triple camera, which offers x2.7 optical zoom, 12 MP resolution, and f1.8 aperture in a specific reading assembly (Figure 1b). After sending the images to a PC, the Image J free software was used to convert images into a tagged image file format on a 16-bit grayscale and analyze them. The analytical signals were considered the variation of the spot intensity and the chip background. Signal-to-noise ratios (SNR) were obtained by dividing the mean spot signals between the associated noise values, calculated as the standard deviation from 15 blank measurements. The limit of detection limits was inferred from the experimental concentration corresponding to the SNR equal to 3.

**Reference methods.** PCR combining with allele-specific hybridization was used to compare the analytical performance of AS-HCR using a smartphone and a chip scanner. Ion Torrent PGM technology (ThermoFisher Scientific, USA) was applied to validate the somatic mutations detection of patient samples. Both methods are described in the Supplementary Information.

### 3.3.4 Results and discussion

#### **Principle of the genotyping of SNVs based on AS-HCR**

The schematic of genotyping by AS-HCR is presented in Figure 1a. The strategy is based on two isothermal recognition events between specific oligonucleotides (primers, probes, and reporters) and a target sequence. First, the specific human genome region is amplified by a fast-isothermal technique, recombinase polymerase amplification (RPA), to generate short-length products. Subsequently, selective sequential hybridizations are activated on the chip surface. The target DNA sequence (G, C, A or T variant) hybridizes with the corresponding immobilized allele-specific probe. Then the 5'-end of the product hybridizes to the link sequence. At the same time, the link acts as the universal initiator to trigger the cascade self-assembly of the two partially complementary DNA hairpins (H1 and H2). The labeling of hairpins yields a hybridization pattern by discriminating SNVs. The approach is compatible with several transducer methods, such as optical or electrochemical. The presented study developed colorimetric chip staining was based on antigen/antibody recognition and metallographic development. Thus, chip imaging can be obtained by a smartphone.



**Figure 1.** Scheme of genotyping by AS-HCR with colorimetric detection. a) Mechanism of the AS-HCR method for identifying single nucleotide variants from genomic DNA. b) Assembly for smartphone detection.

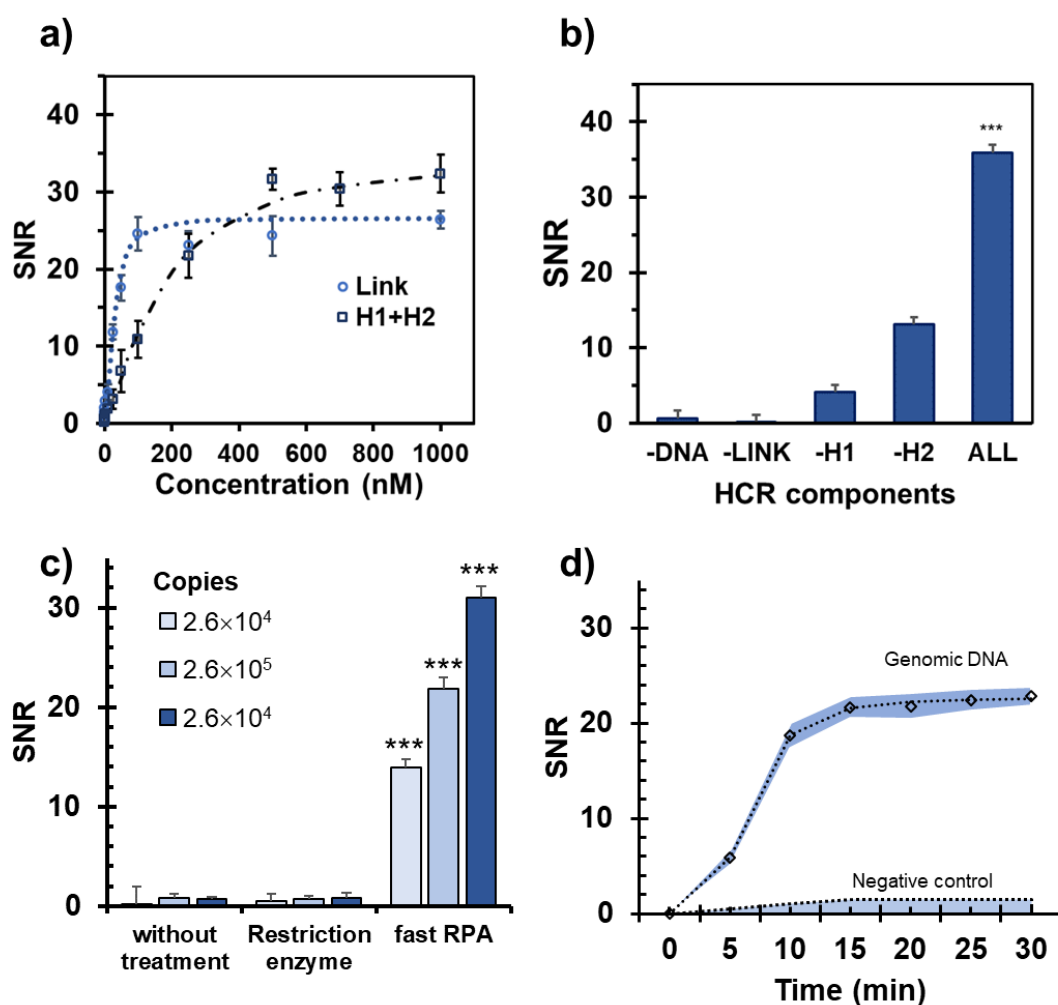
### HCR method setup for short templates and human genome

Solid-phase hybridization assays in an array format and later optical detection were studied on plastic chips. Conventional HCR reaction requirements were examined using a 72-nucleotide sequence as a model template. The study of the oligonucleotide design (Tables SI.1-SI.3, Figures SI.1-SI.2) and experimental variables (Figures SI.3-SI.4) yielded high-intensity signals. The most relevant variable in the amplification performance was the sequences and concentrations of link and H1/H2 oligonucleotides (Figure 2a). The selectivity of assays was confirmed because Only when all the components were present, a long polymer structure of dsDNA was formed and easily detected (SNR>30) (Figure 2b). These observations agreed with the experimental conditions described in previous studies.<sup>28</sup> Applied to short templates, HCR proved to be a powerful enzyme-free amplification technique, and the stoichiometric ratio between the target molecule and markers substantially increased (8.8-fold). However, the method failed to directly use the gDNA (6.4 billion base pairs) extracted from human cells as a template. From gDNA, steric hindrances and strand template stability prevented the formation of the link/template complex, and HCR could not be activated regardless of the initial concentration.

In order to improve its applicability in actual clinical practice, gDNA treatment was studied before the HCR method (Figure 2c). The first option was to use restriction enzymes for genome fragmentation. Although DNA digestion was achieved, there were

no positive signals. We hypothesized that the presence of steric hindrances still affected recognition processes.<sup>29</sup>

The second option was to incorporate an isothermal pre-amplification step. RPA was selected for its high yield, low working temperature, and high reaction rate. Although the two techniques (RPA-HCR) have not been combined to date, the strategy provided the expected results. The formation of a few copies was enough to produce a perfect-match hybrid with the immobilized probe, to be recognized for the linker and, consequently, to trigger the HCR reaction (Figure 2d).



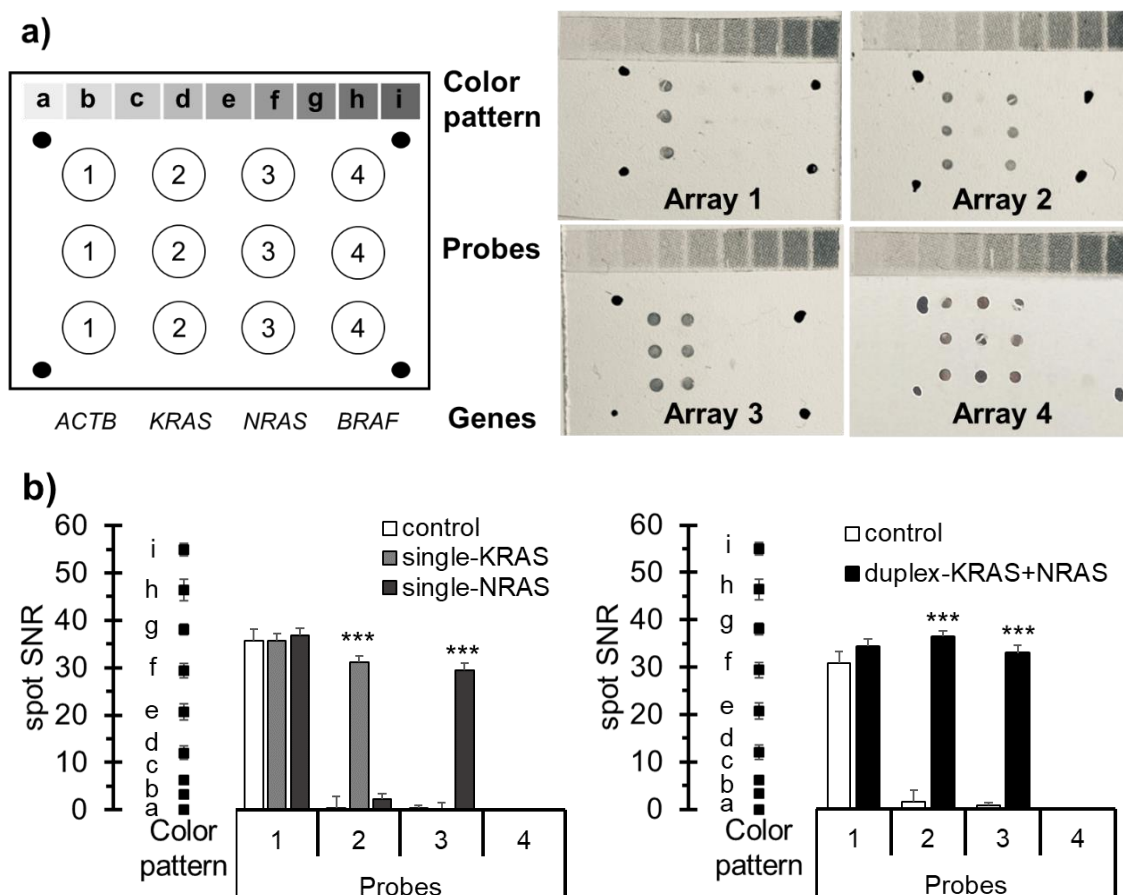
**Figure 2.** HCR method setup. a) Concentration of link ( $[H1] = [H2] = 500$  nM) and H1 plus H2 ( $[Link] = 500$  nM). b) Essential components for HCR (the – sign indicates the component that was not added for each reaction). c) Comparison of the detected signals depending on sample treatment. d) Effect of RPA time. Target: *KRAS* gene (codons 12-13). \*\*\* Student's t-test p-values < 0.05. Sample: human DNA.

We observed that one crucial requisite for successful detection was a correct primer design to yield short amplification products. The estimation involved a competitive hybridization template/link requiring products lower than 80 bp. Another relevant result was that a short amplification step (10 min) sufficed to register detectable signals. Hence the combination of two highly selective efficient amplification processes (RPA-HCR) led to fast and straightforward isothermal approach, i.e., compatible with POC applications.

### **Smartphone-based detection of HCR products**

The detection of the previously described products in an array format was achieved with conventional laboratory instruments such as a chip scanner. The next step was to adjust the method to perform colorimetric smartphone detection for POC applications. To that end, of the different techniques currently available for labeling DNA chips, the immunoassay/nanoparticle system was chosen for its simple protocol.<sup>30</sup> The setup of smartphone-based detection is described as Supplementary Information (Figures SI.5-SI.6). Array quality was confirmed by studying the morphology of spots, including cross-section profiles, intensity, and diameter homogeneity (Figures SI.7). Under the selected conditions, each array spot (diameter 250  $\mu\text{m}$ ) corresponded to 314 pixels. Positive tests depicted up to SNR = 30 in 16-bit grayscale units, while negative ones fell within the chip background range.

A correct pattern was obtained in the single assay format, as positive responses were registered only for complementary probes of the target RPA products (Figure SI.8). The same experiments were successfully performed using a mixture of products from two target regions (Figure 3). Qualitative and quantitative responses supported that the assay was feasible for different targets during single and duplex detection. Also, inter-assay robustness was assured by doing replicates (ANOVA, p-value < 0.05). As a result of excellent performance, this study demonstrates the colorimetric smartphone detection of HCR products for the first time. It opens up a new path for developing fast, portable and easy-to-handle platforms, particularly useful in decentralized scenarios and low-resource health systems.



**Figure 3.** Smartphone detection of RPA-HCR products for the *KRAS* and *NRAS* targets. a) Qualitative results: Layout (left) and resulting images (right) taken by the smartphone using control (array 1), *KRAS* product (array 2), *NRAS* product (array 3), and a mixture of both products (1:1) (array 4). b) Quantitative results: Spots signals for single assays (left) and duplex (right). SNR: signal-to-noise ratio. \*\*\* Student's t-test p-values < 0.05. The color pattern ranges from a (lower intensity) to i (higher intensity).

### Allele-specific HCR setup

Previous HCR-based methods can detect DNA sequences, but they cannot identify any change of the involved nucleotide.<sup>31</sup> The challenge of genotyping SNV was approached by considering the integration of AS-HCR and allele-specific probes immobilized on chips in a microarray format.

Specific probes were designed to maximize the selective recognition process for a single nucleotide by considering thermodynamic calculations (Table SI.1). The results indicated that differences in free energy variation must be significant enough for a discriminatory interaction, with an approximate threshold of 4 kcal/mol, equivalent to a rise in the melting temperature of 6 °C. Moreover, the relative position of probes and the linker was optimized because competition will disrupt the recognition/amplification process if both overlap the same template sequence. The estimation was a separation

of nine nucleotides between probe (3'-end of the target) and the linker (5'-end of the target).

The reaction conditions were revised to maximize hybridization efficiency. The selection criteria were the best discrimination capability when comparing the detection signals from the perfect-match probe and the mismatches addressing the same locus. Regarding the hybridization buffer, the biorecognition process of SNVs was improved by controlling the ionic strength and concentration of formamide (Figure 4ab). As expected, reagents stabilized single-stranded DNA by promoting only the entirely complementary sequences to hybridize.<sup>30</sup> The highest selectivity with no loss of analytical signal was obtained for 25 % formamide and 500 mM citrate buffer (pH = 7). The best responses were obtained incubating at 37 °C for 30 min.

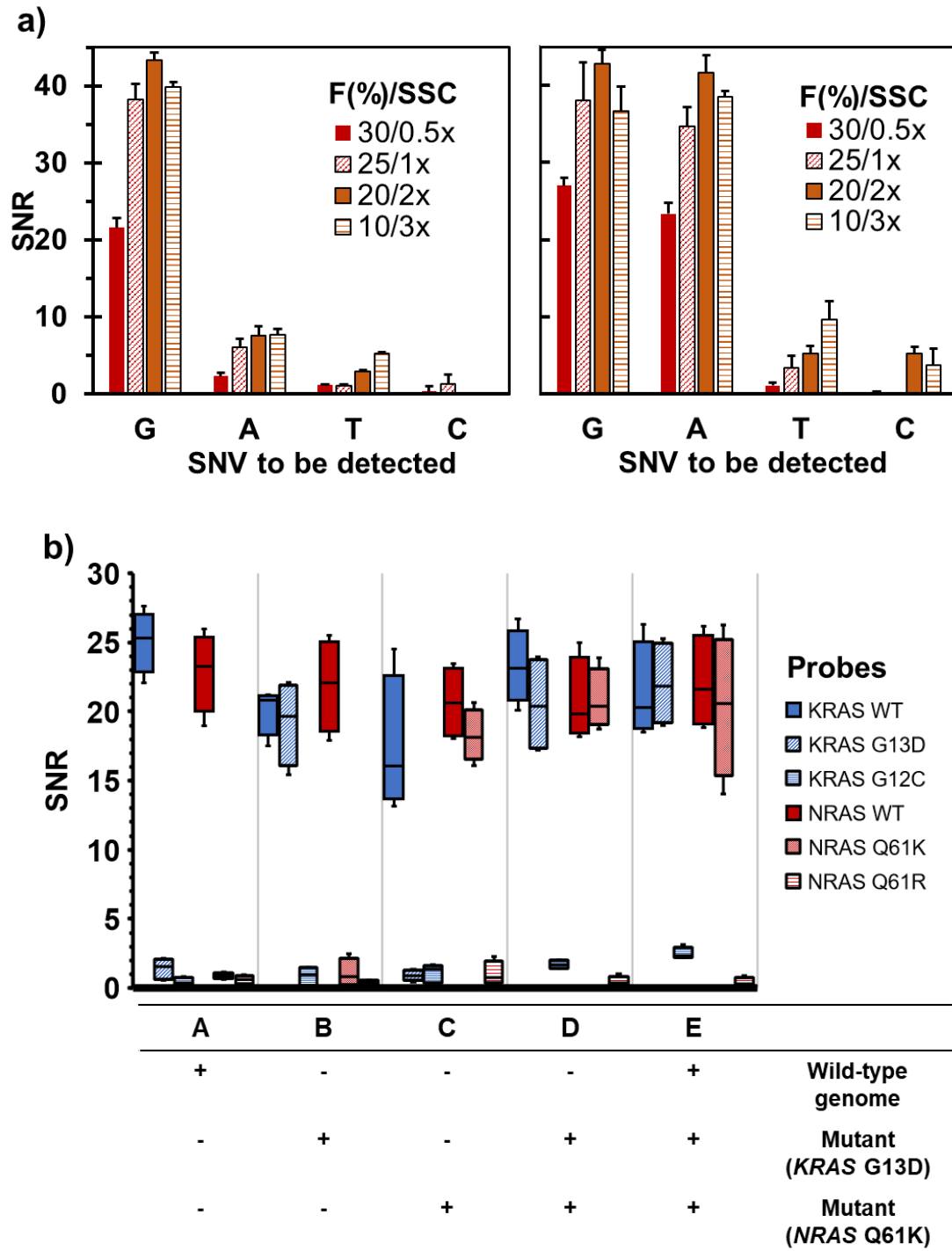
The correct identification of the present DNA variant was achieved because the spot SNR was 23-42 for the perfect-match probes and 0-9 for the mismatch-probes (single-nucleotide change). Consequently, the signal-generating assay was highly specific. For instance, the wild-type cells for the *KRAS* gene only provided positive responses for the correct probe (G probe) with minimal nonspecific intensity in the remaining probe, while the mutant cells (c.38G>A) hybridized to two probes (G and A probes) with similar intensities. These values agree with a gDNA from a human cell culture being heterozygote for this locus.

Therefore, the results demonstrated selective polymer formation by the perfect-match interactions among the corresponding specific probe, the template, the link, and H1/H2. Hence, the accurate discrimination of only one mismatch in one same locus is possible by the AS-HCR method.

### **Multiplexed AS-HCR**

The next challenge was the simultaneous discrimination of several polymorphisms in a single assay. Multiplexed detection required compatible RPA primers, HCR links, and allele-specific probes to be selected for all the target SNVs. As proof of concept, duplex-fast RPA for two human genes was optimized and performed from culture cells (SK-N-AS and HCT116). The formed products were hybridized to the probes immobilized in an array format to perform the AS-HCR method (Figure 4c). The recognition profiles corresponded precisely with the expected perfect-match complexes (probe-template) regardless of assay multiplexing. The results depicted that this methodology is a promising tool to identify the SNVs from gDNA accurately. Thus, RPA combined with AS-HCR in the microarray format can perform the multiplex discrimination of different loci and simultaneously determine many samples by simplifying the analysis and reducing times compared to other detection techniques.



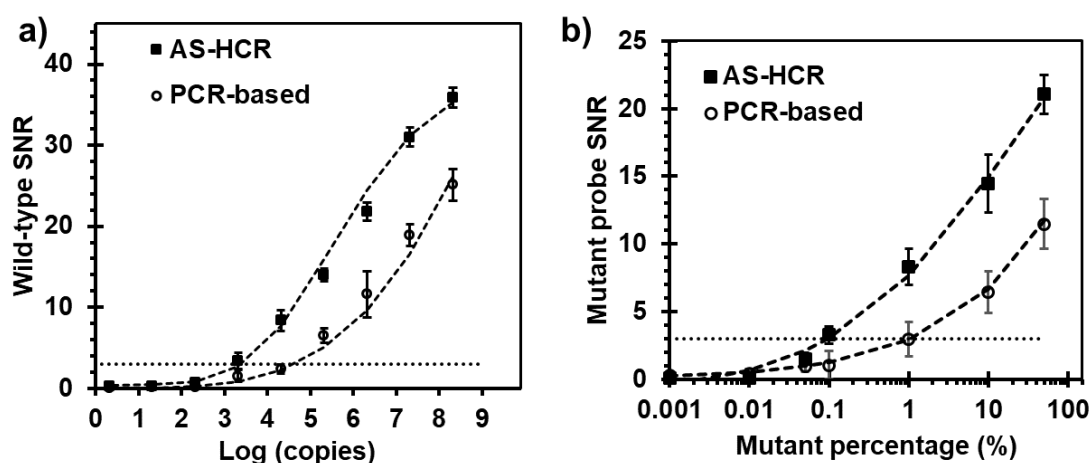


**Figure 4.** Multiplexing capability. a) SNV discrimination ability of wild-type cell line (left) and mutant cell line (c.38G>A KRAS) (right). F(%): formamide percentage and SSC: sodium saline citrate buffer. b) Multiplexing ability of AS-HCR using different mixtures of genomic DNA from cell lines (NGS genotyped samples). SNR: signal-to-noise ratio. As the mutant cell lines were heterozygous, wild-type probes also showed a positive signal.

### Comparison to the PCR-based array and other DNA-biosensors

A conventional allele-specific method based on PCR and hybridization chip was chosen to evaluate the biosensing performances of AS-HCR (Table SI.4). Sensitivity was obtained from serial dilutions of the human genomic template and gDNA mixtures. As shown in Figure 5, an exponential correlation was observed between the spot response and template copies. However, the copies required for the novel method were 1.7-fold lower than the PCR approach. The detection limits, expressed in mutant abundance, were 1.1 % for the PCR-based method and 0.7 % for AS-HCR, i.e., almost 5-fold lower. Each copy of initiators could trigger the HCR event, which resulted in the linkage of many oligonucleotides and showed very high potential in signal amplification for DNA detection purposes.

Reproducibility was determined from the replicated assays and expressed as a relative standard deviation. Values were 5 % for intra-assay and 17 % for inter-assay and, thus, the high consistency among the parallel results confirmed the robustness. Therefore, AS-HCR allowed the specific DNA variants to be identified despite the high wild-type/mutant cell ratio.



**Figure 5.** Comparison of the assay sensitivity of AS-HCR and PCR-based method. a) Effect of DNA copy number. b) Effect of the mutant template percentage, where the total amount remained constant. SNR: signal-to-noise ratio. Mutant: *KRAS* c.34G>T.

Concerning the operational features, the first difference lay in the analysis time, which was 1 h for the novel method (amplification: 10 min, chip assay: 50 min) and 2 h for the conventional method (amplification: 70 min, chip assay: 50 min). Second, AS-HCR was isothermal, low-cost, and instrument-free. Third, this method was less sensitive to the inhibiting factors of the PCR amplification often found in complex matrices.<sup>32</sup> Finally, some authors have pointed that HCR approaches can effectively reduce false-positive results and cross-contamination from amplicons or carry-over pollution, which

frequently occur in PCR<sup>20</sup>. The reason is that the repetitive copy of the template is not controlled.

Although further research is needed, our method depicted some unique advantages over sensing processes in traditional DNA assays, approaching the requirements for an ASSURED test. The use of cheap chips, reagents, and a smartphone as an array reader supports DNA testing in low-resource laboratories (*Affordable*). Most conventional biosensors rely on the hybridization between target molecules and signal probes at a 1:1 stoichiometric ratio, which restricts detection sensitivity.<sup>33</sup> In this novel approach, a single target DNA molecule produces a cascade of hybridizations to form a long concatemers structure and, thus, generates a one-to-multiple amplification effect (*Sensitive*). AS-HCR is an isothermal alternative that does not involve a thermal cycler (*Equipment-free*). Moreover, the specific recognition process (*Selective and Robust*) is compatible with several transduction principles depending on the hairpin marker. If the two hairpins are labeled with digoxigenin, an easy immunoassay can be performed, and the resulting arrays can be read by the colorimetric mode or naked-eye detection (*User-friendly*). Finally, it can be used for POC applications thanks to its simplicity, portability, and low cost in both acquisition and maintenance terms (*Deliverable to end-users*).

### **Genotyping patient samples**

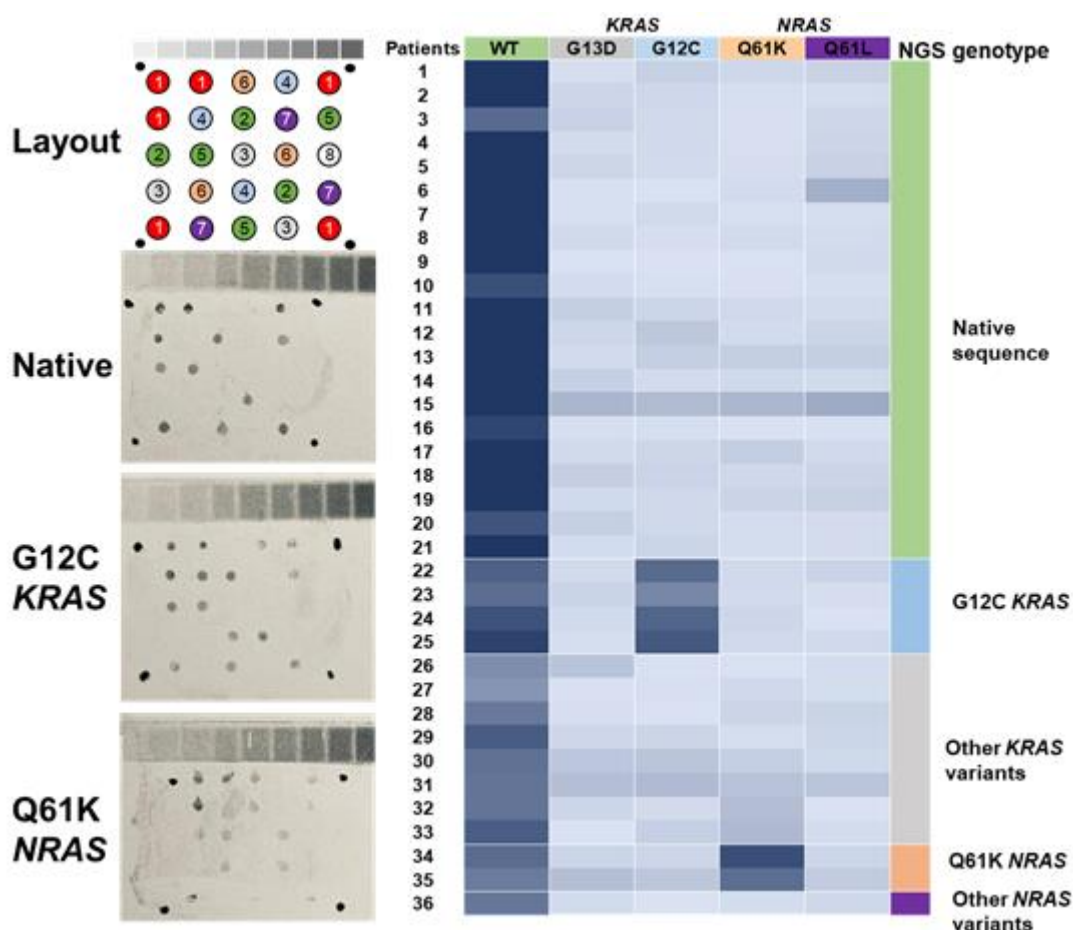
The biosensing ability to distinguish SNVs in real samples, related to solid cancer screening, was tested. As proof of concept, the proposed method was applied to identify the specific mutations in the *KRAS* and *NRAS* genes from biopsied tumors tissues, which were formalin-fixed and paraffin-embedded (FFPE). The validation set was obtained from patients with metastatic colon cancer (n=36). The analysis was conducted as a blind test (Figure 6a, Figure SI.8).

In all cases, the high sensitivity gave a positive response (SNR>3) for the probes corresponding to the wild-type genotype. However, in some patients a small amount of and/or poor-quality DNA due to the nature of clinical samples and their conservation mode. According to the hybridization patterns, positive signals were associated with the p.G12C mutation (c.34G>T) in the *KRAS* gene and the p.Q61K mutation (c.181C>A) in the *NRAS* gene. Four (11.11%) and two patients (5.56 %) were respectively identified. Negligible signals were obtained for the other probes.

Assay accuracy was validated by the independent sequencing of patient samples (Figure 6b, Table SI.5). Complete agreement was obtained, which demonstrates the capability of AS-HCR to identify specific mutations independently of the mutation type

and its position. In addition, the small proportion of mutated cells in tumor tissue did not limit the assay's success.

Although further research is needed, the developed approach is adequate for clinical applications. The AS-HCR method provides key genetic information that can be used to apply personalized medicine to patients with metastatic colon cancer. Considering the state-of-art (Table SI.6 and SI.7), our approach can be classified as a high-moderate sensitive method for mutational analyses (0.1–1%) useful for solid tissues. This method is much simpler and cheaper than instrumental methods and can be applied to other diseases that require genomic screening, focused on the genotyping of nucleotide variations.<sup>6,33,34</sup> These findings facilitate the differential diagnosis and determination of disease prognosis, as well as identification of the drug resistance of tumors and, hence, therapy selection.<sup>(17)</sup><sup>35,36</sup> Table SI.8 shows more examples of the possible applications of the novel technology based on the discrimination of nucleotide variations.



**Figure 6.** Identification of specific mutations from cancer patients using the AS-HCR method. (Left) Spotting layout and images obtained by a smartphone. (right) Heating map of probe responses, classified as the population groups identified by AS-HCR and Ion Torrent sequencing technology. Probes: 1. positive control, 2. WT-KRAS, 3. G13D-KRAS, 4. G12C-KRAS, 5. WT-NRAS, 6. Q61K-NRAS, 7. Q61R-NRAS, 8. negative control.

### **3.3.5 Conclusions**

Currently, reliable, fast, economical, and sensitive methods to discriminate SNVs for clinical applications and other scientific fields need to be developed. The developed HCR method is a powerful, cost-effective alternative for this purpose. In this study, we made several relevant research advances: (i) the treatment of the gDNA with RPA allowed an HCR approach to be applied to complex clinical samples; (ii) the integration of alleles specific to hybridization chain reactions provides the ultrasensitive and selective genotyping of several SNVs; (iii) as the link acts as the trigger for the HCR reaction, the multiplexing capability is high because the amplification is based on a universal pair of oligonucleotides (H1 and H2). Thus, it is possible to extend this technology to other targets by designing the correct probe and link; (iv) consumer electronic devices, such as smartphones, simplify the implementation of molecular methods outside centralized laboratories; (v) the approach shows a high-throughput analysis, portability, and availability. Given its advantages, this novel method implies promising DNA testing opportunities by supporting personalized medicine. The massive genotyping of specific biomarkers (target SNVs) will improve the diagnosis, prognosis, and assignment of the appropriate treatment to each individual patient.

#### **Acknowledgements**

The authors acknowledge the financial support received from EU FEDER, the Spanish Ministry of Economy and Competitiveness (PID2019-110713RB-I00), and the Generalitat Valenciana (PROMETEO/2020/094 and GVA-FPI-2017 Ph.D. grant).

#### **Ethical approval**

Research involving human subjects complied with all the relevant national regulations and institution policies, followed the Declaration of Helsinki tenets (as revised in 2013), and was approved by the authors' Institutional Review Board (Ethics Committee LAFE2015/096). Informed consent was obtained from all the individuals included in this study.

#### **Author Contributions**

The manuscript was written with contributions by all the authors. All authors approved to the final version of the manuscript.

#### **Notes**

The authors declare no competing financial interest.

### 3.3.6 References

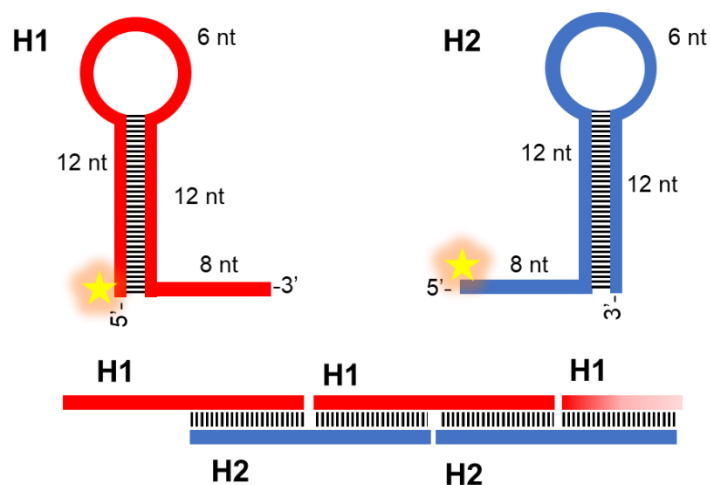
- (1) Zhang, P.; Xia, J. H.; Zhu, J.; Gao, P.; Tian, Y. J.; Du, M.; Guo, Y. C.; Suleman, S.; Zhang, Q.; Kohli, M.; Tillmans, L. S.; Thibodeau, S. N.; French, A. J.; Cerhan, J. R.; Wang, L. D.; Wei, G. H.; Wang, L. High-Throughput Screening of Prostate Cancer Risk Loci by Single Nucleotide Polymorphisms Sequencing. *Nat. Commun.* **2018**, *9*, 1-12.
- (2) Falcomatà, C.; Schneider, G.; Saur, D. Personalizing KRAS-Mutant Allele-Specific Therapies. *Cancer Discov.* **2020**, *10*, 23-25.
- (3) Abi, A.; Safavi, A. Targeted Detection of Single-Nucleotide Variations: Progress and Promise. *ACS Sens.* **2019**, *4* (4), 792-807.
- (4) Lee, S. H.; Park, S. min; Kim, B. N.; Kwon, O. S.; Rho, W. Y.; Jun, B. H. Emerging Ultrafast Nucleic Acid Amplification Technologies for Next-Generation Molecular Diagnostics. *Biosens. Bioelectron.* **2019**, *141*, 111448.
- (5) Matsuda, K. PCR-Based Detection Methods for Single-Nucleotide Polymorphism or Mutation: Real-Time PCR and Its Substantial Contribution Toward Technological Refinement. *Adv. Clin. Chem.* **2017**, *80*, 45–72.
- (6) Gibriel, A. A.; Adel, O. Advances in Ligase Chain Reaction and Ligation-Based Amplifications for Genotyping Assays: Detection and Applications. *Mutat. Res., Rev. Mutat. Res.* **2017**, *773*, 66–90.
- (7) Shen, W.; Tian, Y.; Ran, T.; Gao, Z. Genotyping and Quantification Techniques for Single-Nucleotide Polymorphisms. *Trends Anal. Chem.* **2015**, *69*, 1–13.
- (8) Das, J.; Ivanov, I.; Montermini, L.; Rak, J.; Sargent, E. H.; Kelley, S. O. An Electrochemical Clamp Assay for Direct, Rapid Analysis of Circulating Nucleic Acids in Serum. *Nat. Chem.* **2015**, *7*, 569-575.
- (9) Chang, K.; Deng, S.; Chen, M. Novel biosensing methodologies for improving the detection of single nucleotide polymorphism. *Biosens. Bioelectron.* **2015**, *66*, 297-307.
- (10) Koo, K. M.; Trau, M. Direct Enhanced Detection of Multiple Circulating Tumor DNA Variants in Unprocessed Plasma by Magnetic-Assisted Bioelectrocatalytic Cycling. *ACS Sens.* **2020**, *5* (10), 3217-3225.
- (11) Hajjalyani, M.; Hosseinzadeh, L.; Wu, J. J. Microfluidics-Integrated Sensors toward Rapid Detection of Single Nucleotide Variations. *ACS Omega.* **2021**, *6* (38), 24297-24303.
- (12) Zhao, Y.; Chen, F.; Li, Q.; Wang, L.; Fan, C. Isothermal Amplification of Nucleic Acids. *Chem. Rev.* **2015**, *115* (22), 12491-12545.
- (13) Tortajada-Genaro, L. A.; Yamanaka, E. S.; Maquieira, Á. Consumer Electronics Devices for DNA Genotyping Based on Loop-Mediated Isothermal Amplification and Array Hybridisation. *Talanta.* **2019**, *198*, 424–431.
- (14) Ross, G. M. S.; Salentijn, G. I.; Nielen, M. W. F. A Critical Comparison between Flow-through and Lateral Flow Immunoassay Formats for Visual and Smartphone-Based Multiplex Allergen Detection. *Biosens. Bioelectron.* **2019**, *9* (4), 143.
- (15) Lobato, I. M.; O’Sullivan, C. K. Recombinase Polymerase Amplification: Basics, Applications and Recent Advances. *Trends Anal. Chem.* **2018**, *98*, 19–35.
- (16) Gootenberg, J. S.; Abudayyeh, O. O.; Kellner, M. J.; Joung, J.; Collins, J. J.; Zhang, F. Multiplexed and Portable Nucleic Acid Detection Platform with Cas13, Cas12a, and Csm6. *Science.* **2018**, *360*, 439-444.
- (17) Tang, W.; Zhong, W.; Tan, Y.; Wang, G. A.; Li, Feng; Liu, Y. DNA Strand Displacement Reaction: A Powerful Tool for Discriminating Single Nucleotide Variants. *DNA Nanotechnol.* **2020**, 377-406.
- (18) Yang, D.; Tang, Y.; Miao, P. Hybridization Chain Reaction Directed DNA Superstructures Assembly for Biosensing Applications. *Trends Anal. Chem.* **2017**, *94*, 1–13.
- (19) Augspurger, E. E.; Rana, M.; Yigit, M. V. Chemical and Biological Sensing Using Hybridization Chain Reaction. *ACS Sens.* **2018**, *3* (5), 878–902.
- (20) Bi, S.; Yue, S.; Zhang, S. Hybridization Chain Reaction: A Versatile Molecular Tool for Biosensing, Bioimaging, and Biomedicine. *Chem. Soc. Rev.* **2017**, *46* (14), 4281–4298.
- (21) Qiu, X.; Wang, P.; Cao, Z. Hybridization Chain Reaction Modulated DNA-Hosted Silver Nanoclusters for Fluorescent Identification of Single Nucleotide Polymorphisms in the Let-7 MiRNA Family. *Biosens. Bioelectron.* **2014**, *60*, 351–357.
- (22) Shan Ang, Y.; Lanry Yung, L.-Y. Rational Design of Hybridization Chain Reaction Monomers for Robust Signal Amplification. *Chem. Commun.* **2016**, *52* (22), 4219-4222.
- (23) Miao, P.; Tang, Y.; Yin, J. MicroRNA Detection Based on Analyte Triggered Nanoparticle Localization on a Tetrahedral DNA Modified Electrode Followed by Hybridization Chain

- Reaction Dual Amplification. *Chem. Commun.* **2015**, 51 (86), 15629–15632.
- (24) Wang, X.; Ge, L.; Yu, Y.; Dong, S.; Li, F. Highly Sensitive Electrogenerated Chemiluminescence Biosensor Based on Hybridization Chain Reaction and Amplification of Gold Nanoparticles for DNA Detection. *Sens. Actuators, B.* **2015**, 220, 942–948.
- (25) Li, R.; Zou, L.; Luo, Y.; Zhang, M.; Ling, L. Ultrasensitive Colorimetric Detection of Circulating Tumor DNA Using Hybridization Chain Reaction and the Pivot of Triplex DNA. *Sci. Rep.* **2017**, 7(1), 1-10.
- (26) Wu, J.; Lv, J.; Zheng, X.; Wu, Z. S. Hybridization chain reaction and its applications in biosensing. *Talanta.* **2021**, 122637.
- (27) Roda, A.; Michelini, E.; Zangheri, M.; Di Fusco, M.; Calabria, D.; Simoni, P. Smartphone-Based Biosensors: A Critical Review and Perspectives. *Trends Anal. Chem.* **2016**, 79, 317–325.
- (28) Lázaro, A.; Yamanaka, E. S.; Maquieira, Á.; Tortajada-Genaro, L. A. Allele-Specific Ligation and Recombinase Polymerase Amplification for the Detection of Single Nucleotide Polymorphisms. *Sens. Actuators, B.* **2019**, 126877.
- (29) Ying, N.; Sun, T.; Chen, Z.; Song, G.; Qi, B.; Bu, S.; Sun, X.; Wan, J.; Li, Z. Colorimetric Detection of MicroRNA Based Hybridization Chain Reaction for Signal Amplification and Enzyme for Visualization. *Anal. Biochem.* **2017**, 528, 7-12.
- (30) Figg, C. A.; Winegar, P. H.; Hayes, O. G.; Mirkin, C. A. Controlling the DNA Hybridization Chain Reaction. *J. Am. Chem. Soc.* **2020**, 142 (19), 8596-8601.
- (31) Arnandis-Chover, T.; Morais, S.; Tortajada-Genaro, L. A.; Puchades, R.; Maquieira, A.; Berganza, J.; Olabarria, G. Detection of Food-Borne Pathogens with DNA Arrays on Disk. *Talanta.* **2012**, 101, 405–412.
- (32) Bi, S.; Yue, S.; Zhang, S. Hybridization Chain Reaction: A Versatile Molecular Tool for Biosensing, Bioimaging, and Biomedicine. *Chem. Soc. Rev.* **2017**, 46 (14), 4281–4298.
- (33) Yang, H.; Gao, Y.; Wang, S.; Qin, Y.; Xu, L.; Jin, D.; Yang, F.; Zhang, G. J. In Situ Hybridization Chain Reaction Mediated Ultrasensitive Enzyme-Free and Conjugation-Free Electrochemical Genosensor for BRCA-1 Gene in Complex Matrices. *Biosens. Bioelectron.* **2016**, 80, 450–455.
- (34) Van Krieken, J. H. J.; Rouleau, E.; Ligtenberg, M. J.; Normanno, N.; Patterson, S. D.; Jung, A. RAS testing in metastatic colorectal cancer: advances in Europe. *Virchows Archiv.* **2016**, 468 (4), 383-396.
- (35) Khodakov, D.; Wang, C.; Zhang, D. Y. Diagnostics Based on Nucleic Acid Sequence Variant Profiling: PCR, Hybridization, and NGS Approaches. *Adv. Drug Deliv. Rev.* **2016**, 105, 3–19.
- (36) Stolze, B.; Reinhart, S.; Bullinger, L.; Fröhling, S.; Scholl, C. (). Comparative analysis of KRAS codon 12, 13, 18, 61 and 117 mutations using human MCF10A isogenic cell lines. *Sci. Rep.*, **2015**, 5(1), 1-9.

### 3.3.7 Supplementary information

#### Design of oligonucleotides

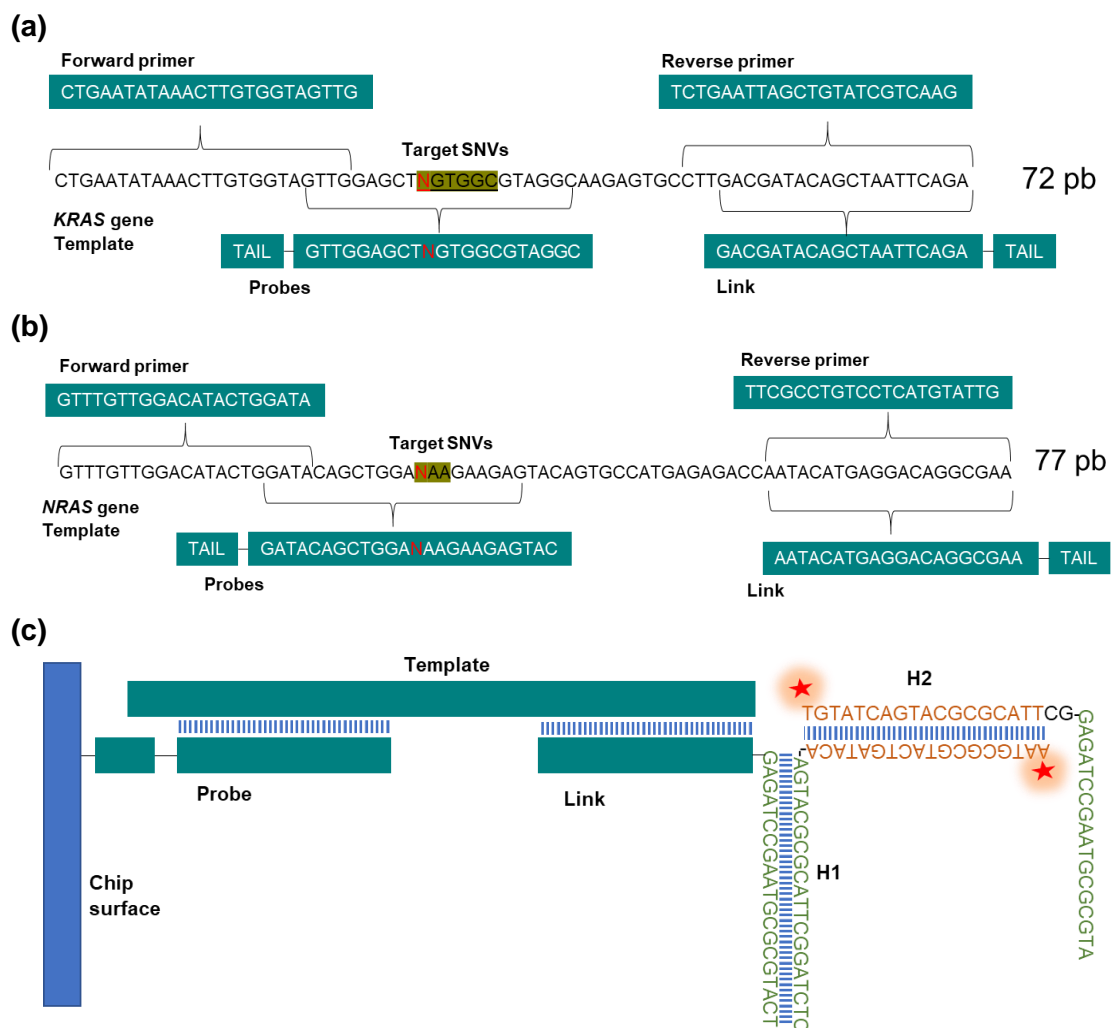
Selection of the H1 and H2 oligos. A correct design was crucial for effective amplification because any slight change in hairpins could lead to an error in the HCR process. The thermodynamic analysis of the candidate oligonucleotides (different lengths and %GC) reported critical information about the sequence of H1, H2, and their secondary structure. Hairpin sets had the short loop protected by the long stem to store potential energy. Reducing toehold and loop size promotes hairpin metastability, while increasing it activates hairpin polymerization.<sup>1</sup> The typical structure utilized in HCR systems is a 6-base loop and 18 bp stem, and both H1 and H2 oligos are symmetric.<sup>2,3</sup> In our approach, a shorter stem structure and a pair of sticky ends were chosen to reduce any unexpected polymerization. The structure of the final design of H1 and H2 (38 nucleotides) had 12 base pair stems, 6 base pair loops, and 8 base complementary tails (Figure SI.1). Both molecules were functionalized with a reporter group in the optical detection mode. The immunostaining option was chosen for the colorimetric approach, and hairpins were modified at their 5' end with digoxigenin as the reporter molecule.



**Figure SI.1.** Structure of the H1 and H2 oligonucleotides.



Target-specific oligonucleotides. Oligonucleotide sets were designed to analyze the most important hotspot mutations in Kirsten rat sarcoma-2 oncogene (*KRAS*, codons 12 and 13) and Neuroblastoma Ras viral oncogene (*NRAS*, codon 61). Therefore, oligonucleotides were complementary to a specific region close to the target SNVs (Figure SI.2). Primer sequences defined the region extremes, probe sequences included the target nucleotide at a central position, and a link sequence was designed at one region extreme. The specific sequences were chosen based on the selection criteria described in previous papers for single nucleotide polymorphisms<sup>4,5</sup> and single-point mutations.<sup>6</sup>



**Figure SI.2.** Design of the oligonucleotides for the AS-HCR method: (a) *KRAS* gene (codons 12-13). (b) *NRAS* gene (codon 61). (c) Universal enzyme-free amplification based on HCR.

In order to achieve multiplexing capability, a specific link was designed as an intermediate to trigger a universal HCR reaction by considering the requirement previously described for single assays.<sup>7</sup> For this purpose, the link was designed with a specific secondary structure. It was formed by base-pair stems, loops, and two tails at the 5'-end and the 3'-end. Therefore, the base-pair hybridization of the template/link and link/H1 hybrids was more stable than the hairpin structure of H1 (Table SI.1). The list of the chosen oligonucleotides is shown in Table SI.2.

The estimated free energy variation for the complexes between the specific probes to its fully complementary target sequence varied from -22.0 to -26.6 kcal/mol. The values calculated for the mismatched complexes ranged between -17.4 and -22.7 kcal/mol. The estimated differences were meaningful enough for a discriminatory assay, and about 4.8 kcal/mol was the equivalent to an increment in the melting temperature of 6.6 °C.

**Table SI.1.** Thermodynamic analysis of the AS-HCR oligonucleotides. (a) Chip hybridization (b) Link/H1/H2 system Software: Dinamelt software for secondary structures.<sup>8</sup>

(a)	Oligo	Sequence (5'-3')	$\Delta G$	mean $\Delta G$	$\Delta G$
			Perfect-match template	Mismatch templates	Difference
<i>KRAS</i> gene	Probe 1	GTTGGAGCTGGTGGCGTAGGC	-25.4	-22.7	2.7
	Probe 2	AGTTGGAGCTG <u>I</u> TGGCGTAGGC	-25.5	-20.8	4.7
	Probe 3	AGTTGGAGCTG <u>A</u> TGGCGTAGGC	-25.3	-21.0	4.3
	Probe 4	AGTTGGAGCTG <u>C</u> TGGCGTAGGC	-26.6	-20.2	6.4
	Probe 5	GTTGGAGCTGGT <u>G</u> ACGTAGGC	-24.0	-18.2	5.8
<i>NRAS</i> gene	Probe 1	ATACAGCTGGACAAGAAGAGTACA	-23.0	-18.1	4.9
	Probe 2	GATACAGCTGGAA <u>A</u> AGAAGAGTAC	-22.0	-17.4	4.6
	Probe 3	TACAGCTGGAC <u>G</u> AGAAGAGTAC	-22.7	-17.7	5.1

$\Delta G$  variation of free energy (kcal/mol) considering  $[Na^+] = 0.15$  M; The underlined bases indicate the target SNVs.

(b)	Oligo	Sequence (5'-3')	Folding nt / $\Delta G$	Hybridization nt / $\Delta G$		
				T-DNA	Link	H1
<i>KRAS</i> gene	Link	GACGATACAGCTAATTCAGA-	6 nt	<b>20 nt</b>	-	-
		GAGATCCGAATGCGCGTACT	-0.6	<b>-19.0</b>	-	-
	H1	AATGCGCGTACTGATACA-	24 nt	4 nt	<b>21 nt</b>	-
	H2	AGTACGCGCATTGGATCTC	-11.1	-4.6	<b>-23.0</b>	-
<i>NRAS</i> gene	Link	TGTATCAGTACGCGCATTG-	24 nt	4 nt	16 nt	<b>20 nt</b>
		GAGATCCGAATGCGCGTA	-14.9	-4.6	-18.9	<b>-20.9</b>
	Link	AATACATGAGGACAGGCGAA-	6 nt	<b>20 nt</b>	-	-
		GAGATCCGAATGCGCGTACT	-0.6	<b>-20.3</b>	-	-
H1	AATGCGCGTACTGATACA-	24 nt	7 nt	<b>21 nt</b>	-	
	AGTACGCGCATTGGATCTC	-11.1	-4.0	<b>-23.0</b>	-	
H2	TGTATCAGTACGCGCATTG-	24 nt	3 nt	16 nt	<b>20 nt</b>	
	GAGATCCGAATGCGCGTA	-14.9	-3.3	-18.9	<b>-20.9</b>	

T-DNA: template region; nt: number of complementary nucleotides in the loop or double-strand structure. DG variation of free energy (kcal/mol) considering  $[Na^+] = 0.15$  M.

**Table SI.2.** Sequences of the oligonucleotides used for the AS-HCR method.

	Sequence (5'-3')	Length (nt)	%GC	T <sub>m</sub> (°C)
<i>KRAS gene (Kirsten rat sarcoma-2 viral oncogene, codon 12-13)</i>				
Primer 1	CTGAATATAAACTTGTGGTAGTTG	24	33	49.7
Primer 2	CTCTATTGTTGGATCATATTCGT	23	35	50.4
Probe WT	[BtnTg]-T10-GTTGGAGCTGGTGGCGTAGGC	21	67	59.0
Probe p.G12C	[BtnTg]-T10-AGTTGGAGCTG <u>I</u> TGGCGTAGGC	21	57	57.1
Probe p.G12S	[BtnTg]-T10-AGTTGGAGCTG <u>A</u> TGGCGTAGGC	21	57	56.8
Probe p.G12R	[BtnTg]-T10-AGTTGGAGCTG <u>C</u> TGGCGTAGGC	21	62	59.1
Probe p.G13D	[BtnTg]-T10-GTTGGAGCTGGT <u>G</u> ACGTAGGC	21	62	56.8
<i>NRAS gene (neuroblastoma ras viral oncogene, codon 61)</i>				
Primer 1	GTTTGTGGACACTGGATA	21	38	47.1
Primer 2	TTCGCCTGTCCTCATGTATTG	21	48	53.4
Probe WT	[BtnTg]-T10-ATACAGCTGGAC <u>A</u> AGAAGAGTACA	24	42	55.0
Probe p.Q61K	[BtnTg]-T10-GATACAGCTGGAA <u>A</u> AGAAGAGTAC	24	42	53.2
Probe p.Q61R	[BtnTg]-T10-TACAGCTGGAC <u>G</u> AGAAGAGTAC	22	50	54.8
Control ( $\beta$ -actin gene, <i>ACTB</i> gene)				
Primer 1	GCACCACACCTTCTACAATGAG	22	50	53.6
Primer 2	GGCCACCAGAAGAGGTAGC	19	63	53.3
Control probe	[BtnTg]-T10-AACCGCGAGAAGATGACCCAGATCA	25	52	60.8
HCR amplification				
Link 1	GACGATACAGCTAATTCAGAGAGATCCGAATGCGCGTACT	40	48	70.2
Link 2	AATACATGAGGACAGGCGAAGAGATCCGAATGCGCGTACT	40	50	71.1
H1	[Dig]AATGCGCGTACTGATACAAGTACGCGCATTCGGATCTC	38	50	70.4
H2	[Dig]TGTATCAGTACGCGCATTCGGAGATCCGAATGCGCGTA	38	53	71.8

[Dig]: digoxigenin-labeled; [BtnTg]: biotinTEG-labeled; T10: tail of 10 thymine tails; T<sub>m</sub>: melting temperature calculated by the Nearest Neighbor model<sup>9</sup>. The underlined bases indicate the target SNVs.

### HCR method setup

***Fabrication of chips and probe immobilization.*** The hybridization slides (25 mm x 75 mm) were prepared by immobilizing the allele-specific probes on planar polycarbonate chips (Makrolon, Germany). Briefly, 30 nL of each biotinylated probe (200 nM) in printing buffer (10 mg/L streptavidin, 50 mM carbonate buffer, pH 9.6) were spotted on chips in a microarray format. To this end, a noncontact printer (AD 1500 BioDot Inc., USA) was employed that worked at 90% relative humidity. The array layout was four probes per array and three replicates per probe for the general tests. For multiplexing assays, seven probes and three replicates per probe were employed. The array also contained positive and negative probes as the controls. The layout enabled the simultaneous analysis of 12 samples. The spotted chips were incubated for 16 h at 37 °C. Then they were washed with PBS-T, rinsed with water, and air-dried.

A bioaffinity strategy was selected for the probe immobilization on the polycarbonate chip given the excellent features reported by several authors.<sup>4,5,9</sup> Under the selected conditions, the biotin-functionalized probes were attached via streptavidin to the chip surface by optimizing the dispensed concentration (Figure SI.3a). The calculated density of the immobilization probe was about 0.5 pmol/cm<sup>2</sup>, and binding was stable for at least 4 weeks.

Reagents. The concentration of each oligonucleotide also played an important role in HCR system performance. We varied concentrations of H1 and H2, remained at an equimolar ratio, to maximize the formation of long H1/H2 hybrids. The best response was obtained using 100 nM of the link and 500 nM of H1 and H2, which implies that the saturation of HCR product growth was controlled by the opening and bonding of hairpins.

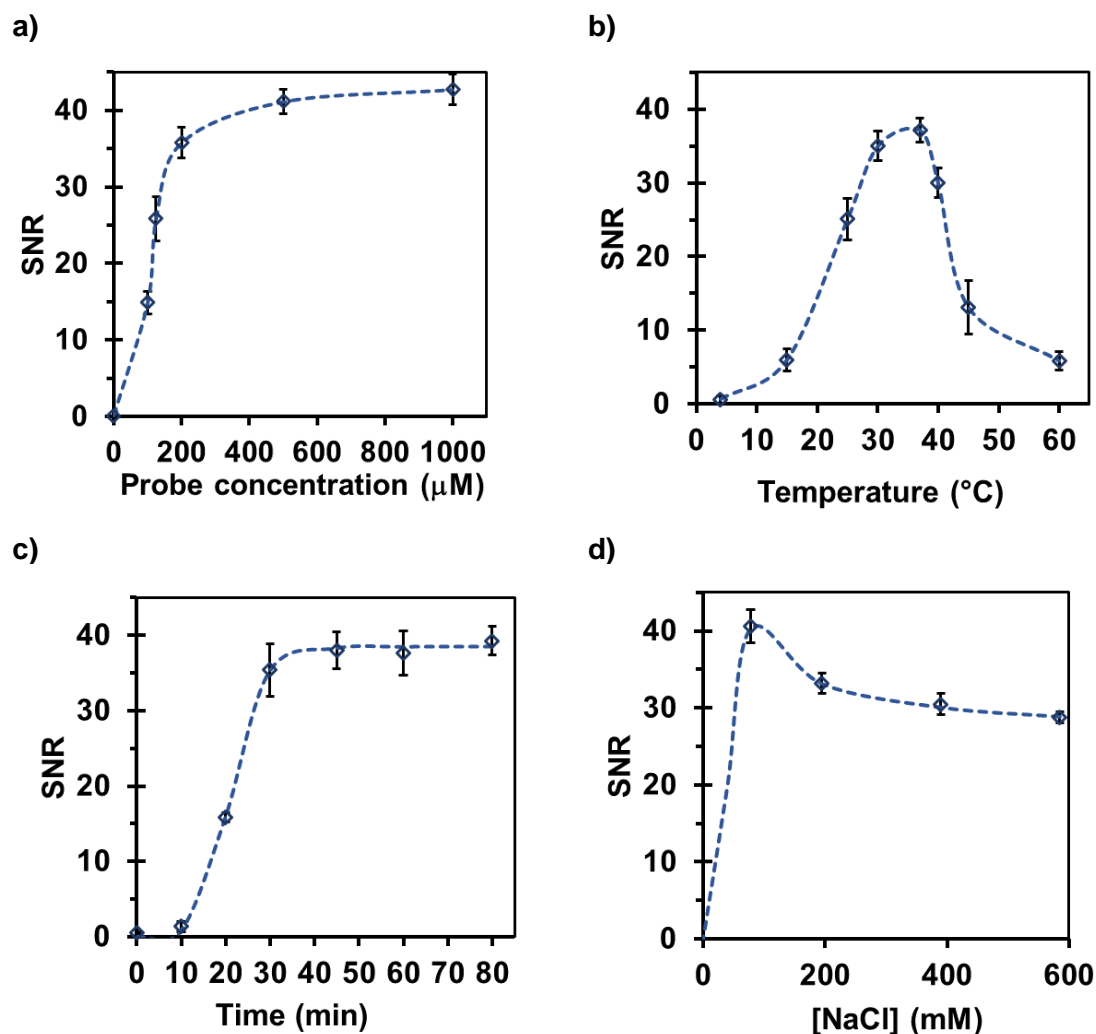
The essential components in the HCR reaction were checked after removing one reagent in each reaction solution. For assay optimization purposes, SNR signals were used to compare the results. The one-factor ANOVA reported significant differences in spot signal (p-value of  $6.13 \cdot 10^{-9}$ ). When all the components were present, the DNA target was hybridized with the link, and the link was able to open one hairpin (H1) after hybridization between complementary regions to expose a new single-stranded region which was able to open the other hairpin (H2) via an unbiased strand-displacement interaction. Therefore, a long dsDNA polymer structure was formed and easily detected (SNR>30). When DNA was absent, the link was closed and could not open the stem structure of H1 (SNR<3). When the link or H1 was absent, non-unions between the target DNA and hairpins occurred, and the two hairpins were concomitant and stable in a closed conformation (SNR<3).

Only a detectable signal was obtained when DNA, link, and H1 were present, indicating a residual HCR process occurred (SNR>13). Although it was possible to bind the DNA/link hybrid to the stem structure of H1, no fully HCR-based amplification could be performed due to the unavailability of H2.

Incubation conditions. In order to enhance the signal, the incubation temperature, time, and hybridization buffer for the HCR method were studied. No response lower than 4°C was observed because H1 and H2 remained in a stable hairpin form. As the temperature rose, hairpins opened, which facilitated the union with the DNA/link hybrid, with a maximum of 37°C (Figure SI.3b). Rapid amplification kinetics agrees with previous studies based on conventional HCR.<sup>10</sup> The highest spot intensity was achieved in 30 min (Figure SI.3c).

Regarding buffer composition, the NaCl concentration correlated with the registered signal (Figure SI.3d). The presence of salt ions (78 mM) reduced the electrostatic repulsion between the phosphate backbone of single-strand DNA and

avored the formation of long double-strand DNA polymers. These results agree with the NaCl effect on hairpin folding and, consequently, on HCR efficiency<sup>2</sup>. The results from these experiments agree with the experimental conditions described for several HCR approaches for short templates.<sup>11</sup> No scientific studies were found in the literature that used HCR directly with genomic DNA.



**Figure SI.3.** Evaluation of conventional HCR reaction conditions. a) probe b) Temperature. c) Time. d) Salt content in buffer solution. SNR: signal-to-noise ratio. Sample: wild-type *KRAS* template from cell culture (5 nM).

Amplification factor estimation. Two model systems were designed for monitoring the recognition process but without HCR amplification (Table SI.3). Control 1 (or direct) consisted of two oligonucleotides called hybrid and marker, the latter of which was functionalized with the reporter molecule. The hybrid oligo sequence was complementary to the immobilized probe (3'-region) and the marker (5'-region). Control 2 (or hairpin-mediated) consisted of three oligonucleotides called hybrid, linker, and H1-marker, the last of which was functionalized with the reporter molecule. The hybrid oligo

sequence was complementary to both the immobilized probe (3'-region) and the linker (5'-region). The linker sequence was also complementary to the hairpin structure.

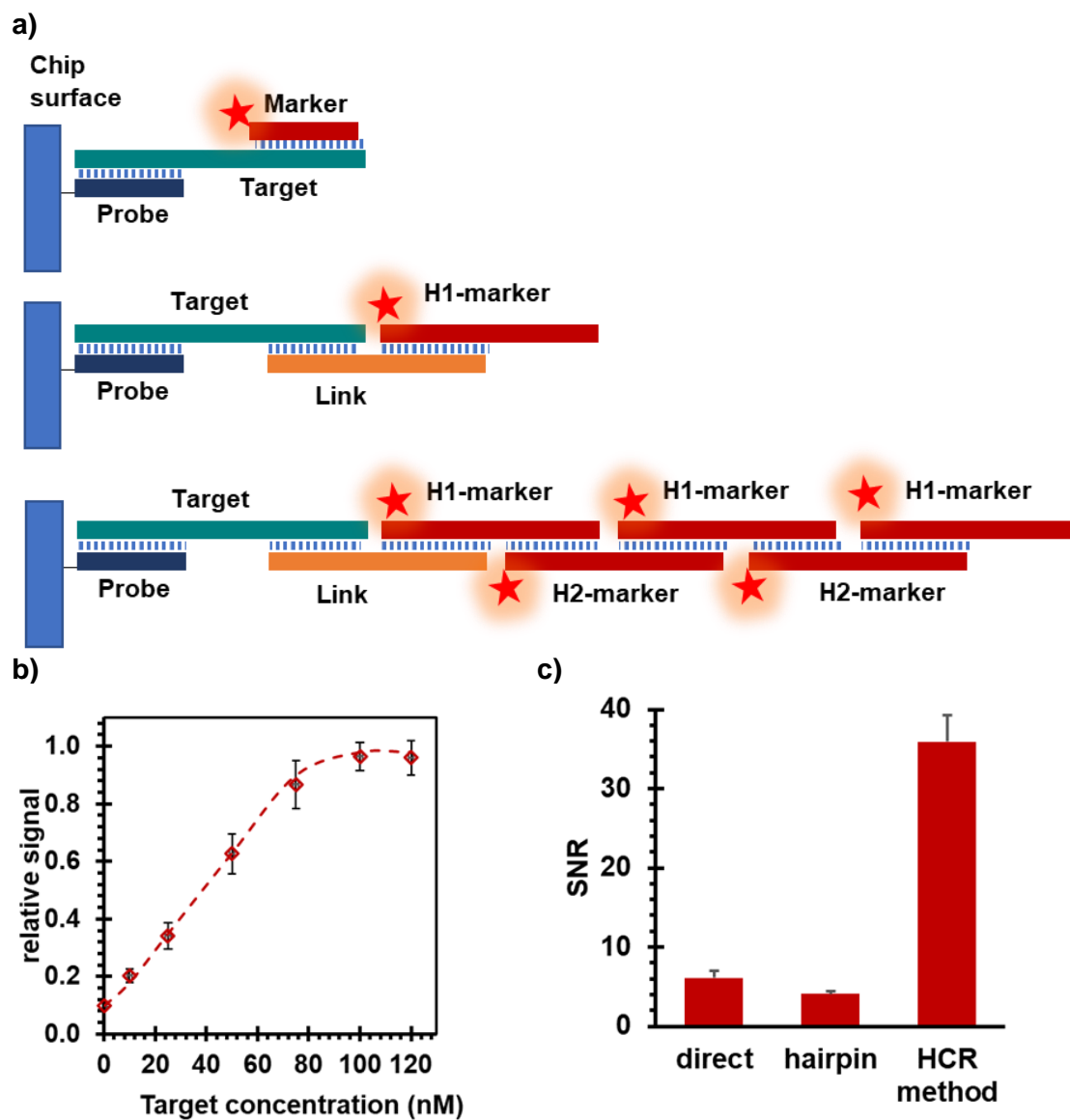
In both controls, the stoichiometric ratio between the target molecule and the reporter molecule was 1:1, i.e., the systems worked as nonamplification controls of the biosensing process (Figure SI.4a). Linear transduction of the target hybridization was obtained until saturation because the response variation increased directly according to the target concentration (Figure SI.4b).

**Table SI.3.** Sequences of the oligonucleotides used as biosensing process controls.

		<b>Sequence (5'-3')</b>	<b>Length (nt)</b>	<b>%GC</b>	<b>T<sub>m</sub> (°C)</b>
Control 1	probe	[BtnTg]T10-GTTGGAGCTGGTGGCGTAGGC	31	67	67.3
	hybrid	GCTTCCTCTGTGTATTTGCCA-TTTTTT-CCTACGCCACCAGCTCCAAC	47	49	81.5
	marker	[Dig]-TGGCAAATACACAGAGGAAGC	21	48	59.5
Control 2	probe	[BtnTg]T10-GTTGGAGCTGGTGGCGTAGGC	31	67	67.3
	hybrid	TCTGAATTAGCTGTATCGTC-AAGGCACTCTT-GCCTACGCCACCAGCTCCAAC	52	52	84.5
	Link	GACGATACAGCTAATTCAGA-GAGATCCGAATGCGCGTACT	40	48	70.2
	H1-marker	[Dig]-AATGCGCGTACTGATACA-AGTACGCGCATTTCGGATCTC	38	50	70.4

[Dig]: digoxigenin-labeled; [BtnTg]: biotin TEG-labeled; T<sub>m</sub>: melting temperature.

The following experiments focused on estimating the amplification factor by the HCR method. The objective was to obtain the number of recognition/hybridization events between a pair of complementary DNA sequences that occurred in succession and yielded a long double strand hybrid. The assay results obtained by the HCR process were compared to the linear transduction of the target hybridization (Figure SI.4c). The signal intensity of the HCR products was 8.8-fold higher than the direct and hairpin-mediated approaches. This confirms that several reporter molecules were attached to each target DNA. Therefore, exponential HCR performance has a very high potential in the signal amplification of multiple targets.



**Figure SI.4.** Linear vs. HCR approaches. a) Scheme of the tested biosensing methods: direct (top), hairpin-mediated (middle), and HCR (bottom). b) Effect of the control concentration for linear systems. c) Comparison of the spot intensities obtained for the studied biosensing approaches. SNR: signal-to-noise ratio.

### Smartphone-based detection setup

Following the guidelines in the literature about point-of-care detection using smartphones,<sup>4,12</sup> a reading platform and methodology were designed and applied (Figure SI.5). The system can obtain a high-definition image of a conventional slide (25 mm×75 mm). A color pattern (grayscale) was added to the image array to correct variations in illumination.



**Figure SI.5.** Assembly for the smartphone-based detection of HCR products. (1) Smartphone for detection (iPhone 11 Pro). (2) Homemade capture chamber (9×11.5×8 cm). The top layer had a 3×3 cm central window that fitted the protruding camera. The bottom was a translucent platform where the chip was located. (3) Illumination 5 W flash LED, e.g., smartphone torch. (4) Hybridization polycarbonate chip. A color palette is included and placed close to all the arrays to correct variations in illumination.

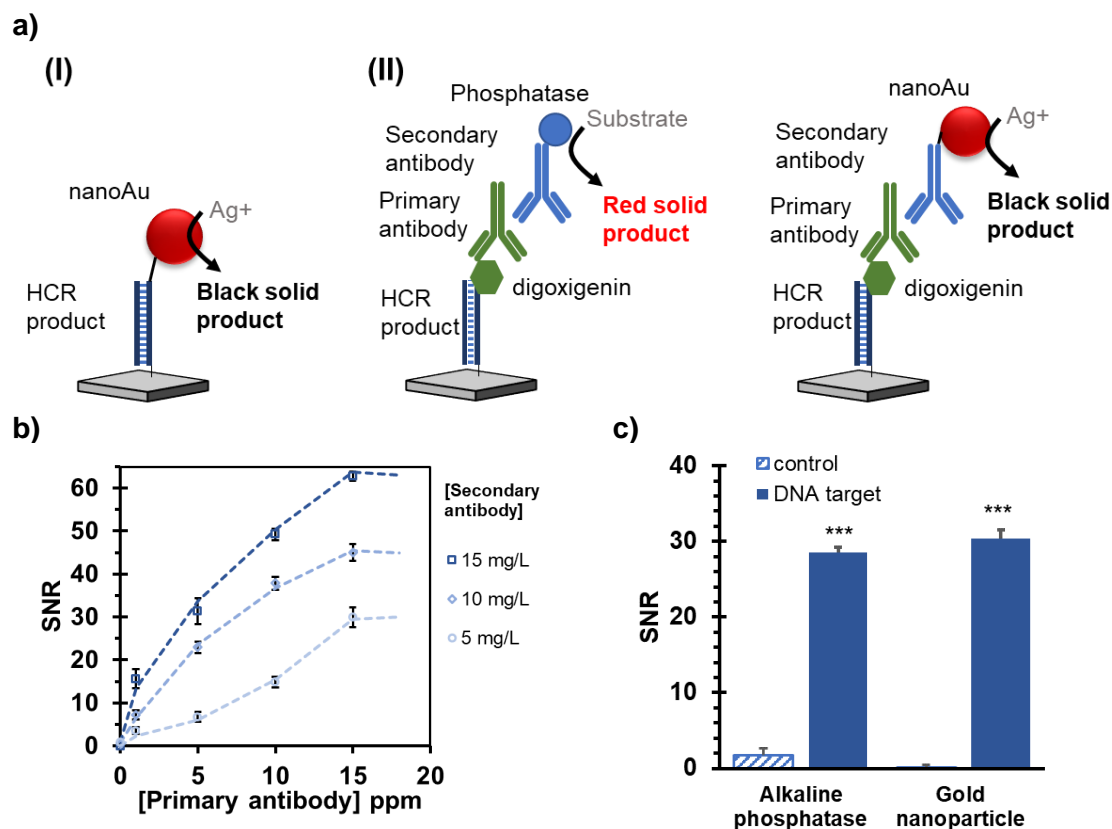
Several options for labelling nucleic acid samples in DNA microarray-based detection methods have been described.<sup>13</sup> Considering the assay requirements, HCR products were labelled using nanoparticles or digoxigenin for colorimetric detection (Figure SI.6).

(i) The direct detection of H1/H2 oligos labelled with gold nanoparticles (5 nm) led to unsatisfactory reflection-mode detection results (a low SNR). By including a silver developer reaction, enhanced detection was promoted because the reduction of silver ions to metallic silver generated an insoluble precipitate in only a few minutes (8 min). A significant attenuation, proportional to the analyte amount, was observed from the captured optical density on the positive spots.

(ii) The assay involved that labelled anti-digoxigenin antibody recognized digoxigenin-functionalized oligos. Later, two staining systems were compared: alkaline phosphatase combined with fast-red colorimetric substrate and gold nanoparticle



combined with the silver developer. In both cases, an insoluble precipitate was obtained in the positive tests, which enabled accurate high-sensitivity optical detection. However, the nanoparticle system offered advantages for the enzyme conjugate by obtaining higher responses due to silver amplification, lower background signals, and a broad spectrum of use given its enzyme-free nature<sup>4</sup>.

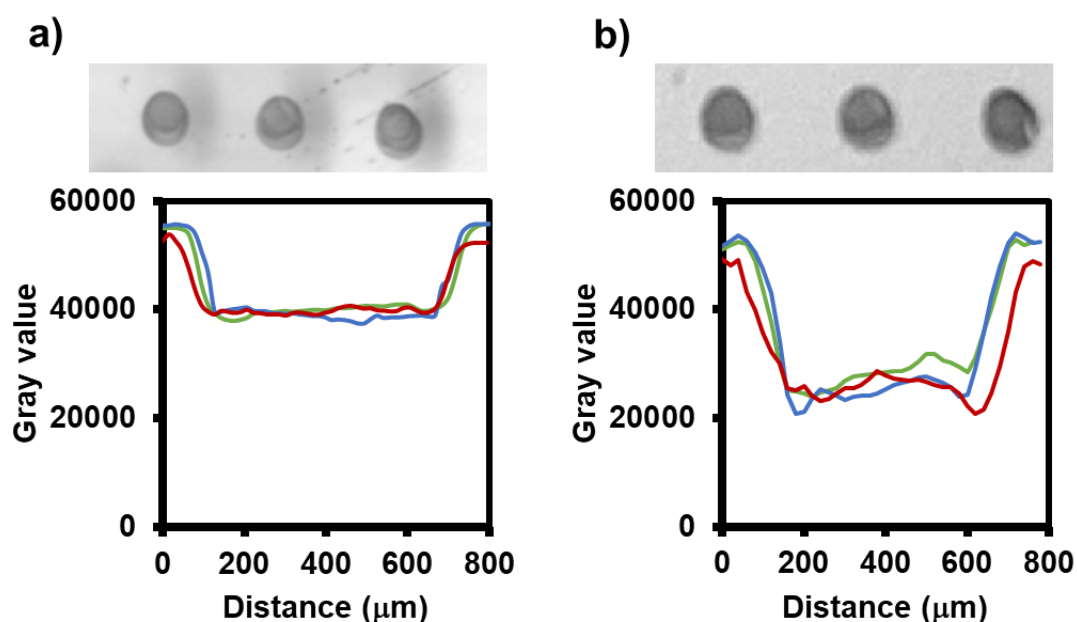


**Figure SI.6.** Optimization of chip immunostaining for smartphone-based detection in the colorimetric mode. a) Studied strategies: Oligonucleotides H1 and H2 were labeled at the 5'-end by 5-nm nanogold (i) or digoxigenin (ii) to generate a solid substrate. b) Effect of antibody concentrations. c) Comparison of developer agents for generating a colored solid deposit after dispensing the cocktail of antibodies.

Therefore, the selected staining of chips was an immunoassay. The working buffer was PBS-T, prepared as 10 mM of sodium phosphate, 0.15 M NaCl, and 0.05% Tween 20 at pH 7.4. The immunostaining solution consisted of an anti-digoxigenin rabbit antibody (Invitrogen, USA) at 5 ppm and a goat antirabbit-antibody labeled with gold (Abcam, UK) at 10 ppm diluted in PBS-T, and incubated for 30 min. Silver enhancer solution, obtained from Sigma-Aldrich (USA), was incubated for 5 min.

To date, HCR products have been labeled mostly with electrochemical and fluorescent markers, which require specific instrumentation.<sup>11,14,15</sup> In the present study, proper integration into easy-to-use detection technologies took place, such as smartphones. This is an appealing HCR approach to run point-of-care tests.

In order to check chip imaging, a conventional documental scanner (Epson, USA) was used to record the array pattern. The scanner was connected to a personal computer via a universal device (USB 2.0). The image was digitized on the grayscale, 16-bits, 1000 ppm, and the file was saved as TIFF using specialized software. Despite having a worse optical resolution than scanners, CMOS sensor chips embedded in phone cameras offer adequate imaging characteristics and wide availability, making them ideal detectors for cost-effective assays. Our results support the technical capabilities of smartphones as analytical readers for molecular diagnostic systems. High-resolution images were recorded and enabled the accurate, sensitive, and simple quantification of the spots generated based on a selective DNA assay (Figure SI.7). Thus smartphone-based detection is an attractive solution for low-cost, ubiquitous portable HCR applications.

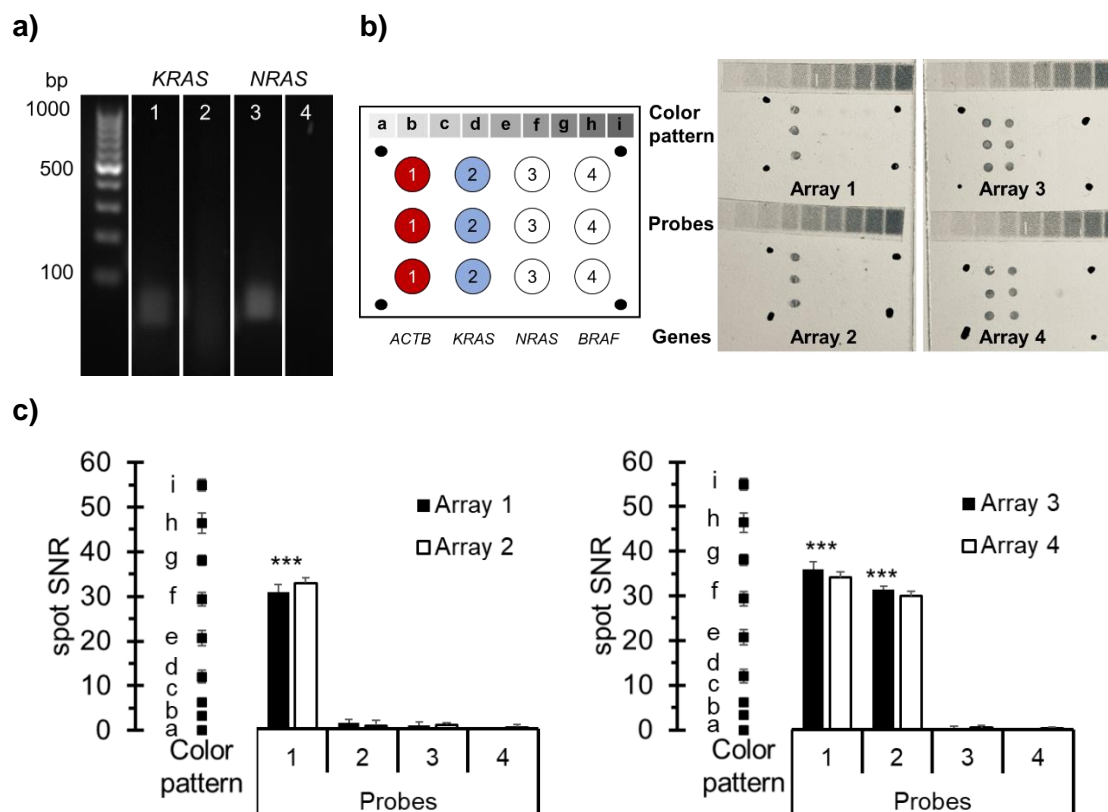


**Figure SI.7.** Assessment of the array quality obtained by a scanner (a) and a smartphone (b). (top) chip images; (bottom) cross-section profiles of spots. Software for the image analysis: Image J.

### HCR method from RPA products

From the genomic DNA of human cell lines (SK-N-AS), target regions were amplified using specific primers and analyzed by gel electrophoresis. The separation of RPA products without purification was conducted in agarose gel 2% with SYBR safe diluted 1:10000 as a staining agent. For that, amplification solutions were mixed with loading buffer and transferred to gel wells. After separation at 90 V for 45 min, a gel image was captured (Figure SI.8a). The observed bands were 70 and 76 pb for *KRAS* and *NRAS* reactions, respectively. Thus, the results confirmed the correct amplification

of target regions because the expected products were obtained. Based on the HCR method, RPA products were successfully detected (Figure SI.8b-c). The images captured by the smartphone showed a clear array pattern depending on the studied products. Also, spots intensities were high and reproducible.



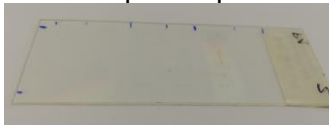



**Figure SI.8.** HCR method from RPA products. a) Agarose gel electrophoresis of RPA products using *KRAS* primers. Lanes 1 and 3: genomic DNA from human cells. Lanes 2 and 4: negative control. b) Qualitative results: Layout (left) and resulting images (right) taken by the smartphone using negative samples (arrays 1 and 2) and positive samples (arrays 3 and 4). Positional markers were put to delimit each array. b) Quantitative results: Spots signals arrays 1 and 2 (left) and arrays 3 and 4 (right). SNR: signal-to-noise ratio. \*\*\* Student's t-test p-values < 0.05. The color pattern ranges from a (lower intensity) to i (higher intensity).

### Allele-specific HCR

PCR combined with allele-specific hybridization was used to compare the analytical performance of AS-HCR (Table SI4). To do so, a reaction mixture was prepared with 1xDNA polymerase buffer, 3 mM of  $MgCl_2$ , 200  $\mu M$  of each deoxynucleotide triphosphate, 300 nM of each primer, 4 ng of genomic DNA, and 1 unit of DNA polymerase (Biotools, Spain). For labeling purposes, 0.01 mM of digoxigenin-11-deoxyuridine triphosphate (Jena Bioscience, Germany) was added.

**Table SI.4.** Comparison of the PCR-chip and RPA-HCR methods

	<b>PCR-chip</b>	<b>RPA-HCR</b>
<b>Amplification</b>	PCR Thermocycler Labeled-dUTPs 35 cycles (95°C/57°C/72°C) 90 min	Fast RPA Heater  37 °C 10 min
<b>Chip</b>	 Polycarbonate 25x75 cm Allele-specific probes	 Polycarbonate 25x75 cm Allele-specific probes
<b>Hybridization</b>	 Restrictive conditions  37 °C 60 min	 Restrictive conditions Reagents: Link, H1, H2 37 °C 30 min
<b>Staining</b>	Colorimetric immunostaining	Colorimetric immunostaining
<b>Detector</b>	Smartphone	Smartphone

PCR was performed in a thermocycler (VWR, USA) under the following conditions: initial denaturation cycle of 95°C for 5 min, 35 denaturation cycles at 95°C for 30 s, annealing at 57°C for 30 s and elongation at 72°C for 60 s and, finally, one cycle extension at the end for 5 min. The amplified products (6 µL) were mixed with hybridization buffer (24 µL), heated (92°C, 10 min), and dispensed on sensing arrays. After incubation (37°C, 60 min), arrays were rinsed with washing buffer. Finally, colorimetric staining and detection were carried out following the same previously described procedure.

### Application to clinical samples

Genomic DNA extraction. The genomic DNA of the cell lines was extracted using the PureLink Genomic DNA kit (Invitrogen, USA). The genomic DNA of the metastatic colorectal cancer samples was obtained with the QIAamp DNA Investigator Kit (Qiagen, Germany). DNA content was quantified in a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). A 260/280 nm absorbance ratio above 1.8 was considered to determine adequate purity.

Next-generation sequencing. Ion Torrent PGM technology (ThermoFisher Scientific, USA) was followed to validate the somatic mutations detection of patient samples. The OncoPrint Solid Tumor DNA kit simultaneously analyzed hotspot mutations in 22 genes (including the *KRAS* and *NRAS* genes). A multiplex PCR amplification of 10 ng of genomic DNA generated the DNA barcoded libraries. The data from the sequencing runs were aligned to the hg19 human reference genome and variant calling.

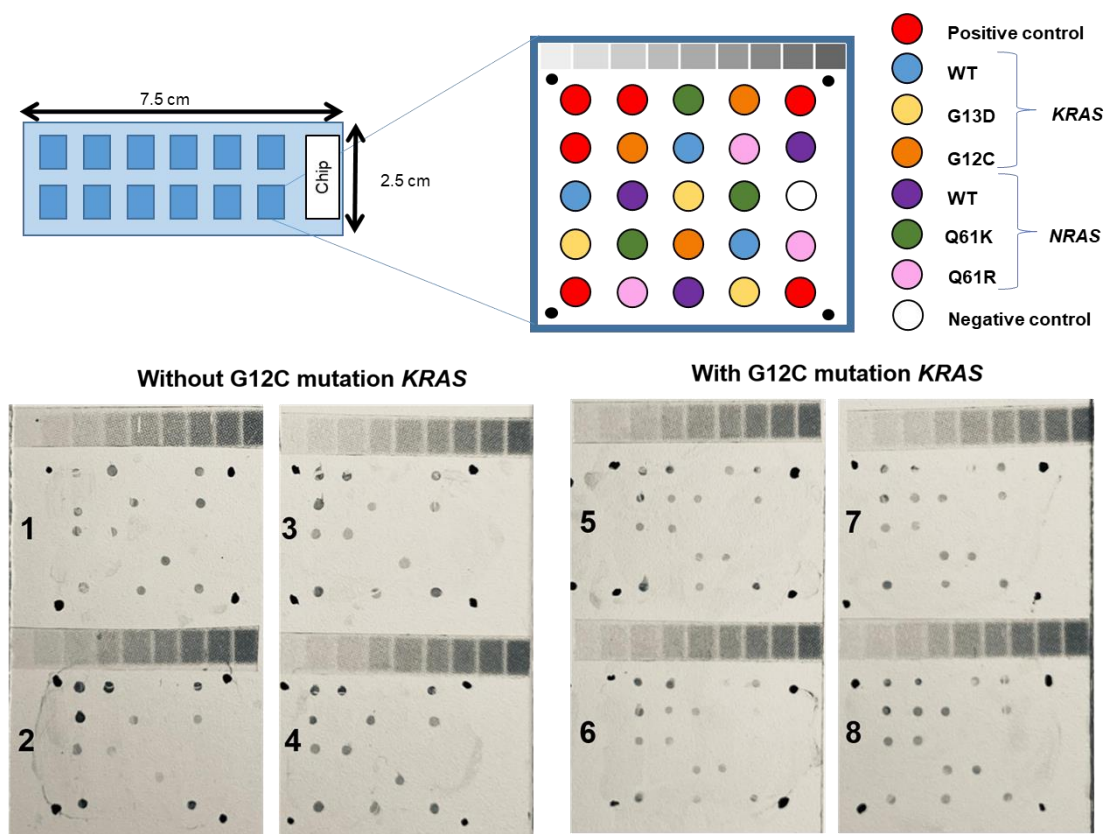
AS-HCR method for single-nucleotide genotyping. The ability of the biosensing system to distinguish SNVs in real samples related to solid cancer screening was tested. As proof of concept, the proposed method was applied to identify the single-point mutations located at a specific locus of the *KRAS* gene (Genomic location 12: 25245347 (GRCh38)). Selective hybridization was achieved. Positive signals (SNR>30) on the probes associated with the specific mutations were recorded for the mutant genomic DNA from cell cultures and patients. Negative or low signals (SNR<10) were obtained for the rest of the probes.

AS-HCR method for multiple targets. Chips contained 12 arrays for the simultaneous analysis of 12 samples. Each array was composed of the anchored probes, which were specific to the main intended mutations for the *KRAS* gene (codons 12 and 13) and *NRAS* gene (codon 61). The studied nucleotide changes were p.G12C (c.34G>T), p.G13D (c.38G>A), p.Q61K (c.181C>A) and p.Q61R (c.182A>G). The array layout was a square with positive controls in corners and three replicates for each target (Figure SI.9). The analyzed samples were obtained from patients with metastatic colon cancer (n=36) as a double-blind assay.

The AS-HCR results obtained from patient samples showed the genetic profile for the studied single-nucleotide variants, which provides valuable information about genes' mutational status. The capture images for the wild-type patients and mutant patients (p.G12C variant) are shown in the example. A clear discrimination pattern was observed, allowing the diagnosis, prognosis, and assignment of appropriate treatment.

Hence, AS-HCR is suitable for centralized laboratories and point-of-care applications considering the employed resources.

Assay accuracy was validated by a parallel analysis of the patient samples following the reference method. Patients suffering from cancer and confirmed by their clinical history and other oncological tests were recruited, including cancer confirmation. A double-blind assay was performed, analyzing tissue samples. The novel method and genomic DNA sequencing results are shown in Table SI.5. Thus, the correct detection of single-nucleotide mutations was confirmed.



**Figure SI.9.** Images obtained from patient samples. In each chip, 12 samples were simultaneously analyzed by discriminating the presence of certain target SNVs.

**Table SI.5.** The mutation analysis of the clinical samples by PCR-based method and NGS technique for the target SNVs: *KRAS* gene (codon 12-13) and *NRAS* gene (codon 61).

Patient sample	<i>KRAS</i> Genotype	<i>NRAS</i> Genotype	Patient sample	<i>KRAS</i> Genotype	<i>NRAS</i> Genotype	DNA code	Protein code
1	WT	WT	22	M	WT	c.35G>A	p.G12D
2	WT	WT	23	M	WT	c.35G>A	p.G12D
3	WT	WT	24	M	WT	c.35G>A	p.G12D
4	WT	WT	25	M	WT	c.35G>A	p.G12D
5	WT	WT	26	M	WT	c.35G>T	p.G12V
6	WT	WT	27	M	WT	c.35G>T	p.G12V
7	WT	WT	28	M	WT	c.35G>T	p.G12V
8	WT	WT	29	M	WT	c.35G>C	p.G12A
9	WT	WT	30	WT	M	c.182A>T	p.Q61L
10	WT	WT	31	M	WT	c.34G>T	p.G12C
11	WT	WT	32	M	WT	c.34G>T	p.G12C
12	WT	WT	33	M	WT	c.34G>T	p.G12C
13	WT	WT	34	M	WT	c.34G>T	p.G12C
14	WT	WT	35	WT	M	c.181C>A	p.Q61K
15	WT	WT	36	WT	M	c.181C>A	p.Q61K
16	WT	WT					
17	WT	WT					
18	WT	WT					
19	WT	WT					
20	WT	WT					
21	WT	WT					

WT: wild-type  
M: mutant

The developed HCR method is a powerful, cost-effective alternative for clinical applications and other scientific fields. In this study, we made several important research advances compared to other HCR-based methods (Table SI.6) or other methods for the genotyping of SNVs (Table SI.7). The advantages include being applied to complex clinical samples, ultrasensitive and selective genotyping of several SNVs, high-throughput analysis, excellent portability, and availability due to its compatibility with smartphone detection.

**Table SI.6.** Examples of different HCR sensing approaches classified according to the sensing principle.

Analytical platform	Labelled hairpins	Signal developer	Target	LOD*	Biological sample	Ref
<b>Fluorescence</b>						
Solution	Both with pyrene excimer	-	Target DNA	256 fM	Cell media	16
Solution	H1-H2 with poly-loop	Ag nanocluster	Let-7a miRNA	0.8 nM	No	17
Chip surface	4 hairpins with non-modification	SYBR Green I	Malaria 18sRNA and GAPDH mRNA	1 pM	Cell lysate	18
Au substrate	Both with azide at their ends	NIR fluorescent dyes	miRNAs-21	1 pM	Cell cultures	19
<b>Electrochemical</b>						
Au electrode	Non-modification	AgNPs	miRNA-17	2 aM	Cell lysates	20
ITO electrode	Non-modification	Fc-PNA probes	Target DNA	100 fM	No	21
AuNP biocathode	Non-modification	[Ru(NH <sub>3</sub> ) <sub>6</sub> ] <sup>3+</sup>	p53 gene fragment	20 aM	Cell lysate	22
<b>Electrochemiluminescence</b>						
Au electrode	Both with biotin at their ends	SA-AuNPs catalyze luminol	HIV-1 DNA fragment	5 fM	Artificial human serum	23
Au electrode	Non-modification	[Ru(NH <sub>3</sub> ) <sub>6</sub> ] <sup>3+</sup>	E. coli DNA fragment	15 fM	No	24
<b>Colorimetric</b>						
Solution	Both with ssDNA sticky ends	AuNPs	Target DNA	50 pM	No	25
Solution	Non-modification	AuNPs	miR-10b, miR-21 and miR-141	20 fM	Cell line	26
Solution	H1 with biotin	Avidin-GOD mediated plasmonic triangular Ag nanoprism etching	p53 gene fragment	6 fM	No	27
Chip surface	H1 with G-quadruplex in the loop and in 1/3 of the stem	DNAzyme catalyze ABTS <sup>2-</sup> oxidation after binding to hemin with H <sub>2</sub> O <sub>2</sub> .	Target DNA and sRNA.	7.5 nM	No	28
Solution	Both with 3/4 and 1/4 of the HRP-mimicking DNAzyme	Hemin/G-quadruplex nanowires catalyze ABTS <sup>2-</sup> oxidation by H <sub>2</sub> O <sub>2</sub>	BRCA1 gene fragment	100 fM	No	29
Solution	Non-modification	Peroxidase mimics of Fe <sub>3</sub> O <sub>4</sub> NPs and AuNPs	Y. pestis DNA fragment	100 pM	No	30
Solution	G-quadruplex structure in the stem of H1	Peroxidase activity by the binding between G-quadruplex and hemin	miR-21, miR-125b	5.6 nM	Diluted sample	31
Chip surface	H1-H2 with digoxigenin	Immunostaining. Smartphone detection	Multi-SNVs	100 fM	Patient tissues	Current paper

\*LOD: Limit of detection



**Table SI.7.** Currently available platforms and closely related SNV strategies for *RAS* genotyping.

Detection method	LOD <sup>1</sup>	Time	Multiplexing (samples/run)	Equipment	CE-IVD <sup>2</sup>	Comments	Ref.
<b>Sanger sequencing</b>	10-25 %	2 days	1-10	Sequencers	CRC RASseq	Requires a high amount of mutated DNA	32
<b>Pyrosequencing</b>	3-7 %	1 day	up to 96	Sequencers	Therascreen KRAS and NRAS Pyro	High error rate in homopolymer readout (>5 ntd)	33
<b>MALDI-TOF-MS</b>	5-10 %	4-5 hours	up to 96	Mass spectrometer	Agena Bioscience OncoCarta	Needs multiple preparation steps and high purity samples	34
<b>dHPLC</b>	3-5 %	Few hours	1	HPLC equipment	-	Cannot detect homozygous mutations directly nor determine the mutation type	35
<b>SNaPshot</b>	1-5 %	7 hours	10	Capillary genetic analyzer	SNaPshot Multiplex Kit	Multi-step process and each assay run separately	36
<b>PCR-RFLP</b>	5 %	4-5 hours	up to 96	Thermal cyclers	-	Limit to targets containing sequences recognizable by the endonucleases	37
<b>Microarrays</b>	0.1-1 %	3 hours	hundreds	Chip reader	Radox biochip array	Low-cost. Challenging to optimize probes design and hybridization conditions	38
<b>HRM</b>	1-5 %	3 hours	up to 96	Real-time thermal cyclers	LightMix NRAS ex2-4 KRAS ex4	Requires a previous PCR amplification step	39
<b>Real-time PCR and TaqMelt</b>	2.5-5 %	8 hours	up to 96	Specific real-time thermal cyclers	COBAS 4800 KRAS assay	The activity of the probe influences sensitivity	40
<b>Scorpion ARMS</b>	1 %	8 hours	up to 96	Specific real-time thermal cyclers	Therascreen KRAS	Scorpion probe design is very complex and laborious	41
<b>Blocked PCR</b>	0.1-1 %	2 hours	up to 96	Real-time thermal cyclers	PNAClamp	Tedious blocker design	42
<b>AS-PCR</b>	0.1-1 %	3 hours	up to 96	Real-time thermal cyclers	EnteroGen KRAS	May lead to false-positives results	43

(continue)

(continued)

<b>COLD-PCR</b>	0.01-1 %	< 3 hours	up to 96	Real-time thermal cycler	-	Requires precise control of working temperatures	44
<b>Droplet digital PCR</b>	0.001-0.1 %	2 hours	up to 96	Sophisticated thermal cycler	-	Absolute quantification. Costly for both the equipment and the reagents,	45
<b>Magnetic-assisted bioelectrocatalytic cycling</b>	0.005 %	2 hours	20	Chrono-amperemeter	-	Isothermal, low-cost. Multi-step protocol	46
<b>Electrochemical clamp assay</b>	1 fg· $\mu\text{L}^{-1}$	1 hour	23	Voltammeter	-	Isothermal, low-cost. Multi-step protocol	47
<b>SHERLOCKv2</b>	2 aM	2 hours	1-8	Fluorescence detector and lateral flow strips reader	-	Isothermal. Multi-step protocol. Limited access to reagents	48
<b>AS-HCR</b>	0.2 %	1 hour	12-100	Smartphone	-	Isothermal, low-cost, versatile transduction and POC detection. Two-step process RPA and detection of HCR products	Current paper

<sup>1</sup> LOD: Limit of detection expressed in percentage of mutated DNA respect to total DNA (%) or concentration ( $\text{fg}\cdot\mu\text{L}^{-1}$  or aM), according to the results of each work.

<sup>2</sup> CE-IVD: In vitro diagnostic devices in European CE marking

**Table SI.8.** Potential applications of the AS-HCR and examples showing the clinical relevance of correct genotyping.

Diseases	Target gene	Target variation	Clinical sample	Examples
Cancer	KRAS NRAS	multiple	Biopsy	49, 50
Retinitis pigmentosa	GARP2	Arg86Gln	Blood or buccal samples	51, 52
Hearing loss	GJB2 LRTOMT	Gly12Val Trp77Arg Arg81Gln	Blood or buccal samples	53, 54
Hepatitis C	RASSF 1A CTLA-4	rs2073498 rs5742909	Blood or buccal samples	55, 56
Pharmacogenetic applications	STAT4 VKORC1 CYP2C9 CYP2C9	rs7574865 rs9923231 rs1799853 rs1057910	Blood or buccal samples	57, 58

## References

- (1) Dirks, R. M.; Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci.* **2004**, *101*, 15275-15278.
- (2) Choi, H. M.; Beck, V. A.; Pierce, N. A. Next generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano.* **2014**, *8* (5), 4284-4294.
- (3) Figg, C. A.; Winegar, P. H.; Hayes, O. G.; Mirkin, C. A. Controlling the DNA hybridization chain reaction. *J. Am. Chem. Soc.* **2020**, *142*, 8596-8601.
- (4) Yamanaka, E. S.; Tortajada-Genaro, L. A.; Maquieira, Á. Low-cost genotyping method based on allele-specific recombinase polymerase amplification and colorimetric microarray detection. *Microchim. Acta.* **2017**, *184*, 1453-1462.
- (5) Tortajada-Genaro, L. A.; Puchades, R.; Maquieira, Á. Primer design for SNP genotyping based on allele-specific amplification—Application to organ transplantation pharmacogenomics. *J. Pharm. Biomed. Anal.* **2017**, *136*, 14-21.
- (6) Martorell, S.; Palanca, S.; Maquieira, Á.; Tortajada-Genaro, L. A. Blocked recombinase polymerase amplification for mutation analysis of PIK3CA gene. *Anal. Biochem.* **2018**, *544*, 49-56.
- (7) Ang, Y. S.; Yung, L. Y. L. Rational design of hybridization chain reaction monomers for robust signal amplification. *Chem. Commun.* **2016**, *52*, 4219-4222.
- (8) Markham, N. R.; Zuker, M. DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* **2005**, *33*, W577-W581.
- (9) Bumgarner, R. Overview of DNA microarrays: types, applications, and their future. *Curr. Protoc. Mol. Biol.* **2013**, *101*, 22-1.
- (10) Li, F.; Tang, Y.; Traynor, S. M.; Li, X. F.; Le, X. C. Kinetics of proximity-induced intramolecular DNA strand displacement. *Anal. Chem.* **2016**, *88*, 8152-8157.
- (11) Yang, D.; Tang, Y.; Miao, P. Hybridization chain reaction directed DNA superstructures assembly for biosensing applications. *TrAC, Trends Anal. Chem.* **2017**, *94*, 1-13.
- (12) Kanchi, S.; Sabela, M. I.; Mdluli, P. S.; Bisetty, K. Smartphone based bioanalytical and diagnosis applications: A review. *Biosens. Bioelectron.* **2018**, *102*, 136-149.
- (13) Gibriel, A.A. Options available for labelling nucleic acid samples in DNA microarray-based detection methods. *Brief. Funct. Genomics*, **2012**, *11*, 311-318.
- (14) Zhang, C.; Chen, J.; Sun, R.; Huang, Z.; Luo, Z.; Zhou, C.; Wu, M.; Duan Y ...Li, Y. The recent development of hybridization chain reaction strategies in biosensors. *ACS Sens.* **2020**, *5*, 2977-3000.
- (15) Kramer, E. E.; Steadman, P. E.; Epp, J. R.; Frankland, P. W.; Josselyn, S. A. Assessing individual neuronal activity across the intact brain: using hybridization chain reaction (HCR) to detect arc mRNA localized to the nucleus in volumes of cleared brain tissue. *Curr. Protoc. Neurosci.* **2018**, *84*, e49.

- (16) Huang, J.; Wu, Y.; Chen, Y.; Zhu, Z.; Yang, X.; Yang, C.J.; Wang, K.; Tan W. Pyrene-excimer probes based on the hybridization chain reaction for the detection of nucleic acids in complex biological fluids. *Angew. Chem., Int. Ed.* **2011**, *50*, 401–404.
- (17) Qiu, X.; Wang, P.; Cao, Z. Hybridization chain reaction modulated DNA-hosted silver nanoclusters for fluorescent identification of single nucleotide polymorphisms in the let-7 miRNA family. *Biosens Bioelectron.* **2014**, *60*, 351–357.
- (18) Xu, Y.; Zheng, Z. Direct RNA detection without nucleic acid purification and PCR: Combining sandwich hybridization with signal amplification based on branched hybridization chain reaction. *Biosens Bioelectron.* **2016**, *79*, 593–599.
- (19) Yin, F.; Liu, H.; Li, Q.; Gao, X.; Yin, Y.; Liu, D. Trace MicroRNA Quantification by Means of Plasmon-Enhanced Hybridization Chain Reaction. *Anal Chem.* **2016**, *88*, 4600–4604.
- (20) Miao, P.; Tang, Y.; Yin, J. MicroRNA detection based on analyte triggered nanoparticle localization on a tetrahedral DNA modified electrode followed by hybridization chain reaction dual amplification. *Chem Commun.* **2015**, *51*, 15629–15632.
- (21) Xuan, F.; Fan, T. W.; Hsing, I. M. Electrochemical interrogation of kinetically-controlled dendritic DNA/PNA assembly for immobilization-free and enzyme-free nucleic acids sensing. *ACS Nano.* **2015**, *9*, 5027–5033.
- (22) Gu, C.; Kong, X.; Liu, X.; Gai, P.; Li, F. Enzymatic Biofuel-Cell-Based Self-Powered Biosensor Integrated with DNA Amplification Strategy for Ultrasensitive Detection of Single-Nucleotide Polymorphism. *Anal Chem.* **2019**, *66*, 37–9.
- (23) Wang, X.; Ge, L.; Yu, Y.; Dong, S.; Li, F. Highly sensitive electrogenerated chemiluminescence biosensor based on hybridization chain reaction and amplification of gold nanoparticles for DNA detection. *Sens. Actuators, B Chem.* **2015**, *220*, 942–948.
- (24) Chen, Y.; Xu, J.; Su, J.; Xiang, Y.; Yuan, R.; Chai, Y. In Situ Hybridization Chain Reaction Amplification for Universal and Highly Sensitive Electrochemiluminescent Detection of DNA. *Anal. Chem.* **2012**, *84*, 7750–7755.
- (25) Liu, P.; Yang, X.; Sun, S.; Wang, Q.; Wang, K.; Huang, J.; Liu, J.; He, L. Enzyme-free colorimetric detection of DNA by using gold nanoparticles and hybridization chain reaction amplification. *Anal Chem.* **2013**, *85*, 7689–7695.
- (26) Rana, M.; Balcioglu, M.; Kovach, M.; Hizir, M.S.; Robertson, N.M.; Khan I.; Yigit, M.V. Reprogrammable multiplexed detection of circulating oncomiRs using hybridization chain reaction. *Chem. Commun.* **2016**, *52*, 3524–3527.
- (27) Yang, X.; Yu, Y.; Gao, Z. A highly sensitive plasmonic DNA assay based on triangular silver nanoprism etching. *ACS Nano.* **2014**, *8*, 4902–4907.
- (28) Dong, J.; Cui, X.; Deng, Y.; Tang, Z. Amplified detection of nucleic acid by G-quadruplex based hybridization chain reaction. *Biosens Bioelectron.* **2012**, *38*, 258–263.
- (29) Shimron, S.; Wang, F.; Orbach, R.; Willner, I. Amplified detection of DNA through the enzyme-free autonomous assembly of hemin/G-quadruplex DNAzyme nanowires. *Anal Chem.* **2012**, *84*, 1042–1048.
- (30) Zeng, C.; Lu, N.; Wen, Y.; Liu, G.; Zhang, R.; Zhang, J.; Wang, F.; Liu, X.; Li, Q.; Tang, Z.; Zhang, M. Engineering Nanozymes using DNA for catalytic regulation. *ACS Appl Mater Interfaces.* **2019**, *11*, 1790–1799.
- (31) Park, C.R.; Rhee, W.J.; Kim, K.W.; Hwang, B.H. Colorimetric biosensor using dual-amplification of enzyme-free reaction through universal hybridization chain reaction system. *Biotechnol. Bioeng.* **2019**, *116*, 1567–1574.
- (32) El Bairi, K. Illuminating Colorectal Cancer Genomics by Next-Generation Sequencing. Springer Nature Switzerland. **2020**.
- (33) Araujo, L. H.; Souza, B. M.; Leite, L. R.; Parma, S. A.; Lopes, N. P.; Malta, F. S.; Freire, M. C. Molecular profile of KRAS G12C-mutant colorectal and non-small-cell lung cancer. *BMC Cancer.* **2021**, *21* (1), 1-8.
- (34) Sherwood, J. L.; Müller, S.; Orr, M. C.; Ratcliffe, M. J.; Walker, J. Panel based MALDI-TOF tumour profiling is a sensitive method for detecting mutations in clinical non small cell lung cancer tumour. *PLoS One.* **2014**, *9* (6), e100566.
- (35) Norman, R. L.; Singh, R.; Langridge, J. I.; Ng, L. L.; Jones, D. J. The measurement of KRAS G12 mutants using multiplexed selected reaction monitoring and ion mobility mass spectrometry. *Rapid Commun. Mass Spectrom.* **2020**, *34*, e8657.
- (36) Penson, R. T.; Sales, E.; Sullivan, L.; Borger, D. R.; Krasner, C. N.; Goodman, A. K.; Growdon, W. B.; Schorge, J. O.; Boruta, D. M.; Birrer, M. J. A SNaPshot of potentially personalized care: Molecular diagnostics in gynecologic cancer. *Gynecol. Oncol.* **2016**, *141* (1), 108-112.

- (37) Li, W. M.; Hu, T. T.; Zhou, L. L.; Feng, Y. M.; Wang, Y. Y.; Fang, J. Highly sensitive detection of the PIK3CA H1047R mutation in colorectal cancer using a novel PCR-RFLP method. *BMC cancer*. **2016**, *16* (1), 1-11.
- (38) Damin, F.; Galbiati, S.; Soriani, N.; Burgio, V.; Ronzoni, M.; Ferrari, M.; Chiari, M. Analysis of KRAS, NRAS and BRAF mutational profile by combination of in-tube hybridization and universal tag-microarray in tumor tissue and plasma of colorectal cancer patients. *PLoS One*. **2018**, *13* (12), e0207876.
- (39) Suhaimi, N. A. M.; Foong, Y. M.; San Lee, D. Y.; Phyo, W. M.; Cima, I.; Lee, E. X. W.; Lim, W. Y.; Chia, K. S.; Likong, S.; Gong, M.; Lim, B.; Hillmer, A. M.; Koh, P. K.; Ying, J. Y.; Tan, M. H. Non-invasive sensitive detection of KRAS and BRAF mutation in circulating tumor cells of colorectal cancer patients. *Mol. Oncol.* **2015**, *9* (4), 850-860.
- (40) Botezatu, I. V.; Nechaeva, I. O.; Stroganova, A. M.; Senderovich, A. I.; Kondratova, V. N.; Shelepov, V. P.; Lichtenstein, A. V. Optimization of melting analysis with TaqMan probes for detection of KRAS, NRAS, and BRAF mutations. *Anal. Biochem.* **2015**, *491*, 75-83.
- (41) Matsunaga, M.; Kaneta, T.; Miwa, K.; Ichikawa, W.; Fujita, K. I.; Nagashima, F.; Furuse, J.; Kage, M.; Akagi, Y.; Sasaki, Y. A comparison of four methods for detecting KRAS mutations in formalin-fixed specimens from metastatic colorectal cancer patients. *Oncol. Lett.* **2016**, *12* (1), 150-156.
- (42) Lázaro, A.; Tortajada-Genaro, L. A.; Maquieira, Á. Enhanced asymmetric blocked qPCR method for affordable detection of point mutations in KRAS oncogene. *Anal. Bioanal. Chem.* **2021**, *413* (11), 2961-2969.
- (43) Barbano, R.; Pasculli, B.; Coco, M.; Fontana, A.; Copetti, M.; Rendina, M.; Vanna, M. V.; Graziano, P.; Maiello, E.; Fazio, V. M.; Parrella, P. Competitive allele-specific TaqMan PCR (Cast-PCR) is a sensitive, specific and fast method for BRAF V600 mutation detection in Melanoma patients. *Sci. Rep.* **2015**, *5* (1), 1-11.
- (44) Galbiati, S.; Damin, F.; Burgio, V.; Brisci, A.; Soriani, N.; Belcastro, B.; Resta, C.-D.; Gianni, L.; Chiari, M.; Ronzoni, M.; Ferrari, M. Evaluation of three advanced methodologies, COLD-PCR, microarray and ddPCR, for identifying the mutational status by liquid biopsies in metastatic colorectal cancer patients. *Clin. Chim. Acta.* **2019**, *489*, 136-143.
- (45) McEvoy, A. C.; Wood, B. A.; Ardakani, N. M.; Pereira, M. R.; Pearce, R.; Cowell, L.; Robinson, C.; Lacopetta, F. G.; Spicer, A. J.; Amanuel, B.; Ziman, M.; Gray, E. S. Droplet digital PCR for mutation detection in formalin-fixed, paraffin-embedded melanoma tissues: a comparison with sanger sequencing and pyrosequencing. *J. Mol. Diagn.* **2018**, *20* (2), 240-252.
- (46) Koo, K. M.; Trau, M. Direct enhanced detection of multiple circulating tumor DNA variants in unprocessed plasma by magnetic-assisted bioelectrocatalytic cycling. *ACS Sens.* **2020**, *5* (10), 3217-3225.
- (47) Das, J.; Ivanov, I.; Montermini, L.; Rak, J.; Sargent, E. H.; Kelley, S. O. An Electrochemical Clamp Assay for Direct, Rapid Analysis of Circulating Nucleic Acids in Serum. *Nat. Chem.* **2015**, *7*, 569-575.
- (48) Gootenberg, J. S.; Abudayyeh, O. O.; Kellner, M. J.; Joung, J.; Collins, J. J.; Zhang, F. Multiplexed and Portable Nucleic Acid Detection Platform with Cas13, Cas12a, and Csm6. *Science*. **2018**, *360*, 439-444.
- (49) Yu, R. T. D.; Garcia, R. L. NRAS mutant E132K identified in young-onset sporadic colorectal cancer and the canonical mutants G12D and Q61K affect distinct oncogenic phenotypes. *Sci. Rep.* **2020**, *10* (1), 1-13.
- (50) Daver, N.; Schlenk, R. F.; Russell, N. H.; Levis, M. J. Targeting FLT3 mutations in AML: review of current knowledge and evidence. *Leukemia*. **2019**, *33* (2), 299-312.
- (51) Gibriel, A. A.; Tate, R. J.; Yu, Y.; Rawson-Lax, E.; Hammer, H. M.; Tettey, J. N.; Pyne, N. J.; Converse, C. A. The p. Arg86Gln change in GARP2 (glutamic acid-rich protein-2) is a common West African-related polymorphism. *Gene*. **2013**, *515*(1), 155-158.
- (52) Daiger, S. P.; Sullivan, L. S.; Bowne, S. J. Genes and mutations causing retinitis pigmentosa. *Clin. Genet.* **2013**, *84*(2), 132-141.
- (53) Gibriel, A. A.; Abou-Elw, M. H.; Masmoudi, S. Analysis of p. Gly12Valfs\* 2, p. Trp24\* and p. Trp77Arg mutations in GJB2 and p. Arg81Gln variant in LRTOMT among non syndromic hearing loss Egyptian patients: implications for genetic diagnosis. *Mol. Biol. Rep.* **2019**, *46*(2), 2139-2145.

- (54) Sliwinska-Kowalska, M.; Pawelczyk, M. Contribution of genetic factors to noise-induced hearing loss: a human studies review. *Mutat. Res. - Rev. Mutat. Res.* **2013**, 752(1), 61-65.
- (55) Ali, N. A.; Hamdy, N. M.; Gibriel, A. A.; Mesallamy, H. O. E. Investigation of the relationship between CTLA4 and the tumor suppressor RASSF1A and the possible mediating role of STAT4 in a cohort of Egyptian patients infected with hepatitis C virus with and without hepatocellular carcinoma. *Arch. Virol.* **2021**, 166(6), 1643-1651.
- (56) Coppola, N.; Minichini, C.; Starace, M.; Sagnelli, C.; Sagnelli, E. Clinical impact of the hepatitis C virus mutations in the era of directly acting antivirals. *J. Med. Virol.* **2016**, 88(10), 1659-1671.
- (57) Lazaro, A.; Yamanaka, E. S.; Maquieira, A.; Tortajada-Genaro, L. A. Allele-specific ligation and recombinase polymerase amplification for the detection of single nucleotide polymorphisms. *Sens. Actuators B Chem.* **2019**, 298, 126877.
- (58) Rost, S.; Fregin, A.; Ivaskevicius, V.; Conzelmann, E.; Hörtnagel, K.; Pelz, H. J.; Lappegard, K.; Seifried, E.; Scharrer, I.; Tuddenham, E. G. D.; Müller, C. R.; Strom T. M.; Oldenburg, J. Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2. *Nature*, **2004**, 427(6974), 537-541.



## **CHAPTER 4:**

# **Development of a BRET-based intercalating dye for DNA biomarker detection**

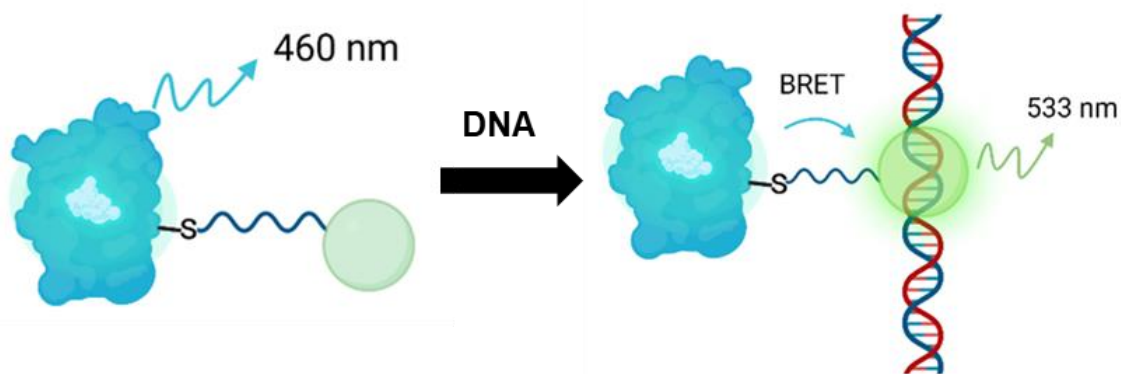




### 3.4 CHAPTER 4: DEVELOPMENT OF A BRET-BASED INTERCALATING DYE FOR DNA BIOMARKER DETECTION

Following the research line described in the previous two chapters, RPA is used as a critical isothermal amplification tool. In this study, we devote our efforts to developing new transduction. Our hypothesis is that bioluminescence meets ASSURED requirements: affordable, sensitive, specific, easy to use, fast and robust, equipment-free, and point-of-care. Here, the novelty is the synthesis of a new bioluminescent reagent, combining a luciferase and a fluorescent intercalating dye, which enables bioluminescent resonance energy transfer (BRET).

One challenge was to study the best conjugation strategy between NanoLuc, a mutant luciferase and the sensitive intercalating agent, such as thiazole orange succinyl ester. Another difficulty is determining the reaction conditions under which only the BRET effect is obtained with dsDNA to achieve a universal and versatile detection platform. As a proof of concept, the reagent is applied to monitor the RPA reaction of the *KRAS* oncogene (codon 12-13) from genomic DNA samples by BRET in end-time and real-time modes without the requirement of an internal excitation source.



Summary figure. Overview chapter 4.

## Development of a BRET-based intercalating dye for DNA biomarker detection

Ana Lázaro,<sup>a</sup> Ángel Maquieira<sup>a, b</sup>, Maarten Merk<sup>c</sup>, Luis A. Tortajada-Genaro,<sup>a, b, \*</sup>

<sup>a</sup> Instituto Interuniversitario de Investigación de Reconocimiento Molecular y Desarrollo Tecnológico (IDM), Universitat Politècnica de València, Universitat de València, Camino de Vera s/n, Valencia, Spain.

<sup>b</sup> Unidad Mixta UPV-La Fe, Nanomedicine and Sensors, Valencia, Spain

<sup>c</sup> Biomedical Engineering department of Eindhoven University of Technology, Eindhoven, The Netherlands.

### 3.4.1 Abstract

The reliable detection of DNA biomarkers supported by simple portable technologies is crucial for molecular diagnosis close to the point-of-need. Due to unique features, bioluminescence-based methods are appropriate for the efficient transduction of molecular events. This work aims to develop a biosensing method based on isothermal DNA amplification and bioluminescence resonance energy transfer (BRET) transduction. The assay uses a novel bioluminescent reagent composed of a mutant NanoLuc luciferase and an intercalating dye, such as thiazole orange. Several conjugation strategies were compared, including the direct coupling to lysine residues and the indirectly coupling to cysteine residues with 1-(2-aminoethyl) maleimide hydrochloride as a crosslinker. These conjugated were characterized by gel electrophoresis and HPLC-MS. In the presence of double-strand DNA, the BRET phenomenon was observed registering an intense emission at 533 nm, enabling optical detection without needing an excitation source. As the reagent signal correlates to DNA concentration, the functionalized NanoLuc is a universal bioluminescence dye with excellent performances (detection limit of 10 nM and standard deviation <3%). The biosensing system was applied to monitor the products from an isothermal amplification technique, such as recombinase polymerase amplification (RPA) in the end-point and real-time modes. The RPA-BRET method successfully detected a specific region located at the *KRAS* gene (codon 12-13) from DNA genomic samples. In conclusion, a simple detection and quantification of DNA biomarkers is possible thanks to a versatile bioluminescence reagent.

**Keywords:** bioluminescence sensing; BRET; NanoLuc luciferase; isothermal amplification; DNA biomarker detection.

### 3.4.2 Introduction

The specific determination of DNA biomarkers provides essential genetic information to implement personalized medicine (1). Among the different techniques for their detection, fluorescent methods in solution using molecular beacons or intercalating agents stand out. Nevertheless, the sensitivity of these methods is limited due to the background fluorescence light scattering, which can severely hamper applications in strongly absorbing or scattering media, such as complex medium. Another drawback is the requirement of excitation light sources, making it necessary to use more complicated equipment (2,3). Hence, challenging research is the development of alternative optical transduction principles to improve the features of fluorescent dyes (4).

A powerful detection strategy for molecular diagnosis is bioluminescence. In this technique, emission light is mediated by biological reagents or entire organisms such as bacteria, fireflies, and marine life during a chemical reaction. An interesting approach is supported on the oxidation of a luciferin substrate by an enzyme called luciferase. The electronically excited oxyluciferin produced returns to the ground state releasing a photon (visible light energy) (5). A relevant example is the combination of isothermal DNA amplification with bioluminescent detection. Here, a product of amplification reaction, i.e., pyrophosphate, is catalyzed into adenosine triphosphate, which is converted to bioluminescence through a thermostable luciferase reaction (6,7).

Over time, different luciferin-luciferase combinations were discovered with various characteristics such as emission wavelength, intensity, type, and duration signals. The Renilla (Rluc) and Firefly (Fluc) are well-known luciferases but are predominantly used as gene reporters in cellular assays and in vivo imaging due to their low emission intensity (8). In 2012, Promega introduced a new, very bright luciferase, NanoLuc. This luciferase was obtained from the deep-sea shrimp *Oplophorus gracilirostris*. Extensive re-engineering of this tetrameric protein led to a functional 19 kDa monomeric luciferase. NanoLuc exhibits higher thermal and chemical stability and can function optimally between pH 7-9, whereas the other luciferases have a narrower range. Further engineering of the original substrate, coelenterazine, resulted in furimazine, which had higher luminescence, more excellent stability, and less autofluorescence than coelenterazine (9). Combining this optimized substrate and NanoLuc luciferase produces a glow-like signal at 460 nm that is 100-fold brighter than the bioluminescence produced by other systems. NanoLuc has been applied on molecular biorecognition and research applications (10,11).

A novel optical transduction is the bioluminescence resonance energy transfer (BRET) phenomenon, which is similar to the principle of Förster resonance energy

transfer (FRET). Both events arise when the relaxation of an excited donor chromophore is achieved by non-radiative energy transfer to an acceptor chromophore through dipole-dipole interactions (12). The principle of BRET is based on the transfer of energy between a bioluminescent donor, luciferase, and an acceptor fluorophore. The transition to the excited state of the bioluminescent donor occurs via an oxidation reaction catalyzed by the luciferase. The high-energy state drops to the basal state, resulting in the characteristic emission band of the luciferase substrate. BRET occurs when part of the energy released by the excited donor is transferred to a nearby acceptor fluorophore. Then, relaxation of the excited acceptor results in the emission of its characteristic wavelength (13). For BRET to occur is crucial that a spectral overlap occurs between emission and excitation spectra of the bioluminescent donor and the acceptor, respectively. In addition, the efficiency of the energy transfer depends on several parameters, such as the dipole-dipole alignment, the distance between the two chromophores, and the quantum yield (14).

BRET-based biosensing systems have been obtained by coupling NanoLuc to proteins and applied to molecular recognition and cellular imaging (15). The most widely used strategy was cysteine-maleimide conjugation (16). In 2016, five different color variants coupling NanoLuc to fluorescent proteins were reported, resulting in highly efficient BRET effect for bioimaging (17). Also, antibody/NanoLuc hybrids have been constructed for developing bioluminescent immunoassays to detect interferon- $\gamma$  (18) and to distinguish between cell lines (19). However, the employment of the BRET principle applied to DNA sensing is a recent concept. Yoshida et al. developed a sensing molecule based on a zinc finger fused to firefly luciferase and the BOBO-3 intercalating dye (20). Merx et al. designed a platform to detect short synthetic single-strand DNA or RNA by coupling the NanoLuc with molecular beacons and using Cy3 fluorophore as a BRET acceptor (21). However, these solutions lack sensitivity for practical applications (5,22).

In this work, we have synthesized a new NanoLuc conjugate that combines the universal and straightforward performances of DNA intercalating dyes with the benefits of bioluminescence detection. Hence, our research hypothesis was that this solution enables the sensitive detection of double-strand DNA molecules, such as amplification products. Moreover, these unique features will be especially suitable for isothermal amplification techniques aimed to develop point-of-care platforms. Among all isothermal methods described, recombinase polymerase amplification (RPA) is the proposed reaction due to its excellent performances, such as low working temperature, short time, simple primer design, high amplification yield, and low effect of inhibitors (23).

### 3.4.3 Materials and methods

**Synthesis of NanoLucs.** Mutated NanoLucs were obtained based on a protocol including mutagenesis, expression, and purification described in the literature (21). Two pET expression vectors (pET28a plasmid) with the gene coding for kanamycin resistance and NanoLuc with an N-terminal streptavidin tag (strep-tag) and a C-terminal polyhistidine affinity tag (his-tag) was ordered from Genscript. Multisite-direct mutagenesis was performed by PCR using QuickChange (Agilent Technologies) according to the manufacturer's protocol, and the correct sequence was confirmed by sequencing (Figure SI.1).

The plasmids encoding the NanoLuc mutants were transformed in *E. coli* BL21 (DE3) (Novagen) and cultured in 8 mL of LB medium. Since the plasmids contain a gene for antibiotic resistance, bacteria cultures were supplemented with 8  $\mu$ L of kanamycin to select the cells with an insert plasmid. The cells were cultured until the productive growth phase ( $OD_{600} = 0.6$ ) using an Eppendorf Biophotometer. Then, 1 mM of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) was added, which triggered the transcription of the lac operon, inducing *E. coli* NanoLuc expression. The cells were incubated overnight at 20  $^{\circ}$ C. Later, proteins were harvested by centrifuging for 10 min at 10,000 rcf at 4 $^{\circ}$ C and resuspending the pellet in the detergent BugBuster and benzonase endonuclease (Novagen). The lysed cells were centrifugated for 30 min at 40,000 rcf at 4 $^{\circ}$ C, and the cleared lysate (supernatant containing proteins) was transferred for further processing. NanoLucs were purified through Ni<sup>2+</sup>-affinity chromatography. In short, the reaction mixture was loaded on a His-bind resin column so that the his-tag of NanoLucs bound selectively while other proteins or contaminants were washed away. Afterward, 250 mM of imidazole was added to elute the Ni-bound NanoLucs. Subsequently, they were passed through a streptavidin column and eluted by E-buffer (100 mM Tris, 150 mM NaCl, 1 mM EDTA, pH 8.0) to ensure the removal of any contaminants and obtain pure proteins.

After protein expression, the proteins were incubated with 200 mM of Tris (2-carboxyethyl) phosphine for 1 h at 25 $^{\circ}$ C in shaking at 450 rpm to avoid disulfide bridging dimers. Later, they were desalted afterward using PD-10 columns (Merck). Their concentration was increased using 0.5 mL amicon ultra cell-10 K (Merck). The protein concentration was calculated by measuring the absorbance at 280 nm and the extinction coefficient using the ExPASy-ProtParam tool. SDS-PAGE gel electrophoresis was used to confirm the synthesis of the mutant Nanolucs (see Supplementary Information).

**Conjugation of NanoLuc and intercalating agent.** The bioluminescent reagent luciferase-dye was synthesized conjugating NanoLuc to the thiazole orange succinimidyl

ester (Biotum). Different strategies were approached depending on the used protein residues. The conjugate synthesis was characterized by SDS-PAGE gel electrophoresis and liquid chromatography-mass spectrometry (see Supplementary Information). As a control of the conjugation, the luciferase was also conjugated to Cyanine 3 maleimide (Cy3) (Lumiprobe). Following the literature (21), a 20-fold excess of the dye was mixed to 10  $\mu$ M NanoLuc in PBS at pH=7. The solution was incubated for 2 h at 4 °C with continuous shaking and darkness.

**Tests based on BRET measurements.** Bioluminescence assays were performed in white 384-microwell plates (Thermo Fisher Scientific). Blanks were buffered solutions without DNA, negative contained synthetic single-strand DNA (ssDNA, 108 nt), and positive contained synthetic double-strand DNA (dsDNA, 108 bp) at different concentrations. Solutions (10  $\mu$ L) were mixed with NanoLuc conjugates (10  $\mu$ L) in the microwells. After incubation for 10 min at room temperature in darkness, 1  $\mu$ L of the NanoLuc substrate, furamizine (Promega), was added. The emission at 460 nm (furamizine-NanoLuc emission) and 533 nm (BRET effect) was measured by a plate reader (Tecan Spark II). The ratio between the emission of the luciferase and intercalating dye was related to DNA concentration.

**RPA-BRET method. Analysis of human samples.** The studied target was a region of the *KRAS* gene, codon 12-13, associated with many types of solid cancer (24). Specific designed primers were supplied by Eurofins. The RPA reagents used for isothermal DNA amplification were the TwistAmpBasic RPA kit TwistDx. Dilutions of human genomic DNA (Thermo Fisher Scientific) were used.

RPA reaction mixtures (50  $\mu$ L) were prepared with rehydrated buffer, 14 mM of magnesium acetate, 200 nM of forward primer (5'-CTGAATATAAACTTGTGGTAGTTG-3'), 200 nM of reverse primer (5'-TCTGAATTAGCTGTATCGTCAAG-3'), 5  $\mu$ L of genomic DNA, and the enzyme pellet. Solutions were mixed with NanoLuc-dye conjugate at 30 nM (10  $\mu$ L) and 1  $\mu$ L of furimazine in each microwell. The plates were incubated at 37 °C for 30 min in darkness. The reaction was monitored at end-point mode (1 measurement at minute 30), and real-time mode (1 measurement per 1 min) using a plate reader (Tecan Spark II). The registered signals were emissions wavelength at 460 nm (donor) and 533 nm (acceptor) with an integration time of 100 ms.

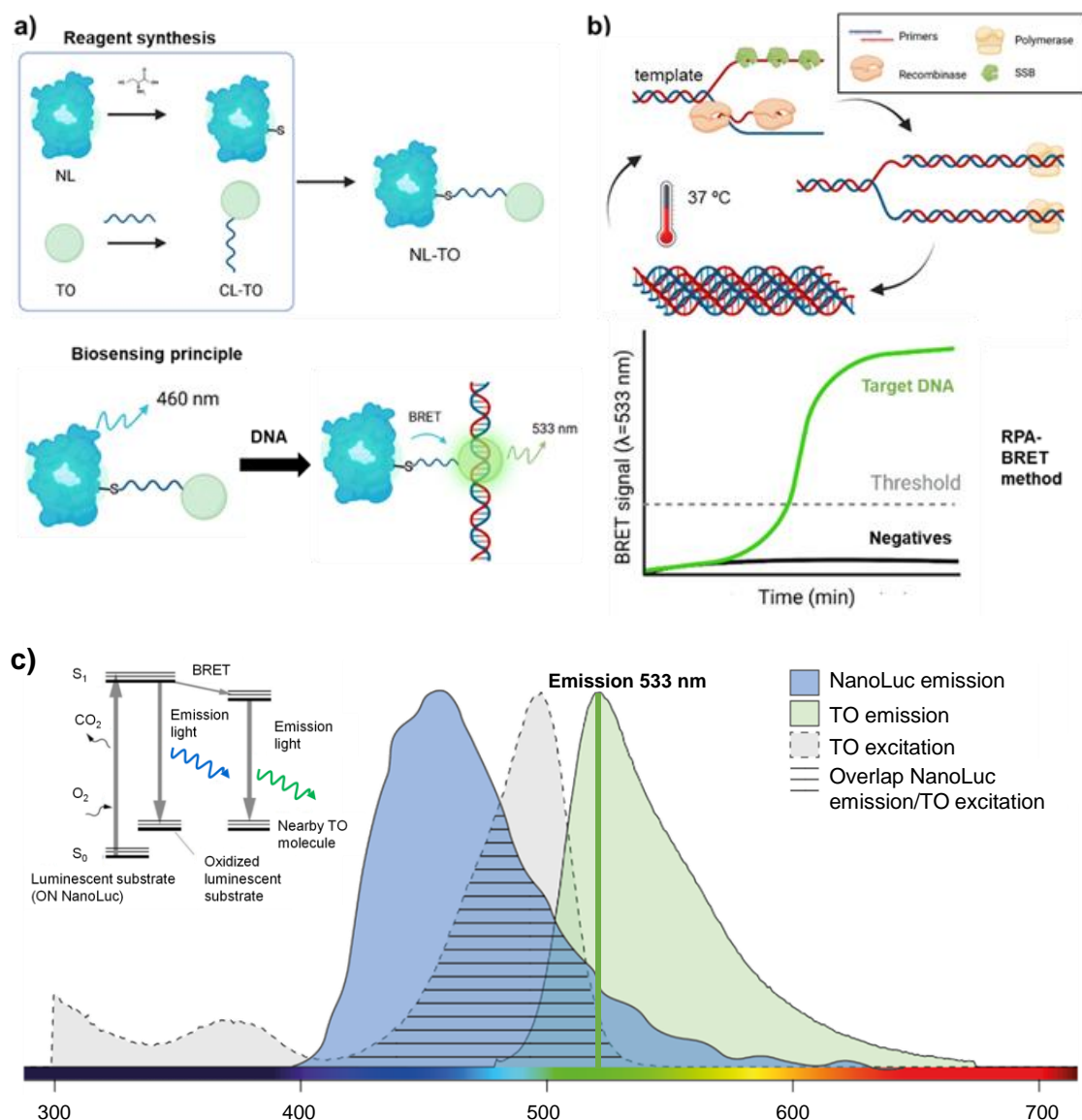
### 3.4.4 Results and discussion

#### Principle of the detection system

The detection principle was based on BRET phenomena activated by the presence of DNA biomarker as shown in Figure 1. For this purpose, the bioluminescence reagent was synthesized from the conjugation between the thiazole orange succinyl ester, a dsDNA-intercalating dye, and NanoLuc, a luciferase with high luminescence intensity and chemical stability. In absence of the target, the conjugate only emitted at 460 nm, corresponding to the background bioluminescence. As the luciferin substrate (furimazine) is oxidized catalyzed by the NanoLuc, an extremely unstable, high-energy peroxide intermediate is formed. Then, the peroxide moiety of the intermediate undergoes homolytic cleavage and decarboxylation, thus forming a singlet excited oxyluciferin (14). The excited molecule forms an oxyanionic version of oxyluciferin (furimamide), and energy is released in light form.

The presence of DNA molecules produces the formation of a complex with the reagent. The excited-state energy generated by substrate conversion is transferred to a nearby thiazole orange molecule of the conjugate. When the excited fluorescent molecule falls back to the ground state, the characteristic light of the fluorescent protein is observed as there is an optimal spectral overlap between NanoLuc substrate emission spectrum and the thiazole orange excitation. Thus, the intramolecular energy transfer occurs from the donor region, luciferase, to the acceptor region, thiazole ring, resulting in light emission at higher wavelengths (band at 533 nm, green).

As the reagent is a dsDNA binding molecule that intercalates nonspecifically into the double-strand structure, the proposed system is a universal bioluminescence dye. A relevant application of this biosensing principle is the real-time monitoring of isothermal DNA amplification, such RPA. As the reaction was happening, the copy number increases and, consequently, the number of complexes NanoLuc-dye conjugate/DNA. Measuring the resulting emission, the detection system registers the reaction kinetics. Therefore, a correlation between the number of template copies and the registered signal can be established. This assay simplifies the required instrument compared to qPCR or isothermal approaches using fluorescent dyes. The needed elements are thermoblock for heating at a constant temperature and photodetector at 533 nm.



**Figure 1.** Scheme of DNA biomarker detection using bioluminescence sensing. a) Synthesis of the assembled BRET-intercalating dye and the detection principle of DNA. b) Real-time bioluminescence monitoring of isothermal DNA amplification based on RPA. c) Spectrums emission/excitation and emission spectra of NanoLuc and thiazole orange. NL: NanoLuc luciferase, CL: crosslinker, TO: thiazole-orange.

### Synthesis of mutant NanoLucs

A relevant feature of the proposed transducer is the effective conjugation of NanoLuc and the candidate intercalating reagent. The limitations are the number of available reactive sites, such as lysine (-NH<sub>2</sub>) and cysteine (-SH), and the losing activity after the protein conjugation. In fact, the conjugation of NanoLuc using the unique native cysteine (Cys164) leads to the loss of its bioluminescence emission (11). Therefore, mutants NanoLuc were designed and obtained to increase the coupling options to the intercalating agent (Figure SI.1). The first selected modification was to replace the native

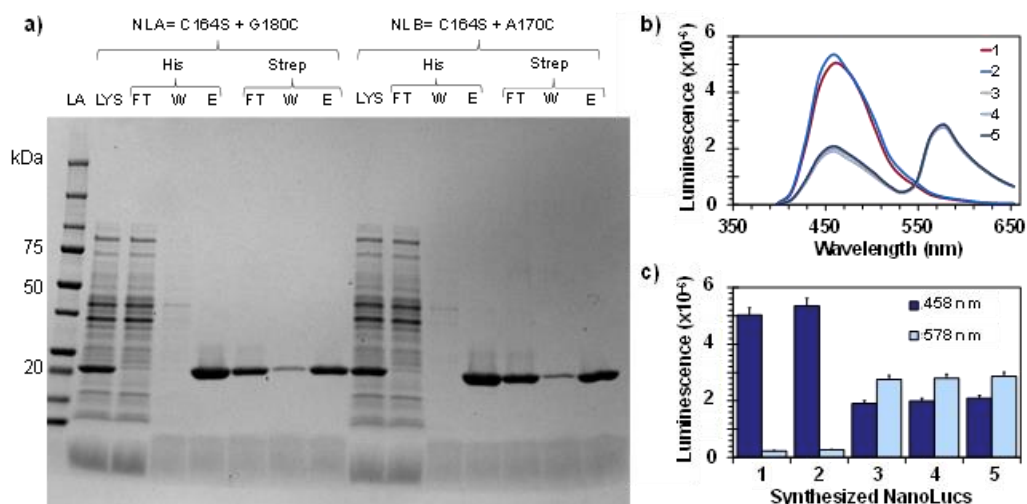


cysteine with a serine leading to mutation C164S. Both amino acids are polar amino acids showing structural similarity; thus, the estimated protein folding was similar. The second modification was to increase the number of thiol groups. So, a new cysteine residue into the protein sequence was introduced via site-directed mutagenesis. In order to not interfere with the protein function, a tail of 11 amino acids was also incorporated at the C-terminal end. Two different luciferases were designed depending on the position of the cysteine in the tail, being the first residue for NLA variant (G180C) and the last for NLB (A170C). The third modification was the inclusion of purification tails (streptavidin at the N-terminal and polyhistidine at the C-terminal).

Mutant luciferases were isolated, and extracts were characterized by gel electrophoresis. Two purification strategies were tested, a single purification using the luciferase 6-histidine tag tail and a sequential purification protocol based on the 6-histidine tag tail and the streptavidin tail. As the gel electrophoresis showed (Figure 2a), a single purification with the Ni-NTA column was enough to obtain pure proteins, judging from the absence of other bands. The quality of extract was similar after the second purification step, considering the flow-through and the elution bands. The observed bands correspond to the expected size for both NanoLucs (22kDa), so the introduced mutations did not affect protein expression.

The activity of the synthesized NanoLuc (NLA and NLB) was tested compared to the wild-type enzyme. For this purpose, furimazine, a substrate with a quantum yield of 5% (25) was added and the luminescence was registered. An overlapping spectrum was obtained with the maximum peak at 460 nm for the mutants and wild-type (Figure 2b). The comparison of emission signals indicated no significant differences were obtained from two mutant luciferases (Student's t-test p-values < 0.05). Therefore, the position of the incorporated modifications did not influence the luciferase spectrum and protein activity.

On the other hand, the mutants NanoLuc were conjugated to Cy3 by cysteine-maleimide coupling and the luminescence spectrum was registered (Figure 2c). The BRET effect was observed indicating a dye efficient conjugation and NanoLuc mutagenesis or tail modifications did not impact on the energy transfer. Experiments performed for blanks and DNA standards provided similar signals confirming Cy3 was only an acceptor fluorophore but it is invalid for recognizing DNA biomarkers.



**Figure 2.** Control of NanoLuc synthesis. a) SDS-PAGE gel of the purified mutated luciferases: NLA (C164S + G180C) and NLB (C164S + A170C). LA: ladder in kDa as a reference point. LYS: the lysate bacterial culture. FT: the flow-through that came out of the columns after adding bind buffer. W: the wash with buffer containing imidazole. E: Elution fractions with the purified proteins. The purity of the final product based on relative band intensities was estimated to be > 99%. The clear bands around 22 kDa corresponded to the mutant NanoLucs (Mw=19 kDa) and their tags attached ( $\approx$ 3 kDa). b) Spectrum luminescence of synthesized NanoLucs. c) Signal luminescence intensity for luciferase (458 nm) and fluorophore (578 nm) emission. Keeping luciferases at 10 nM. 1: NanoLuc wild-type, 2: Mutant NanoLuc (NLA and NLB), 3: Mutant NL-Cy3 + NTC, 4: Mutant NL-Cy3 + ssDNA, 5: Mutant NL-Cy3 + dsDNA. Mutant NanoLucs corresponded to NLA and NLB. The mutants NL-Cy3 were incubated with DNA (1000 nM).

### Features of NanoLuc and intercalating agent

The choice for intercalating dye was based on excitation spectrums, dsDNA affinity, fluorescent properties, and coupling capabilities (Table SI.1). The selected option was the thiazole orange, a cyanine dye with high fluorescence quantum yield upon binding to dsDNA and it has functional groups available for coupling (Figure SI.2) (26). Also, the experiments with the dye and DNA standards confirmed the excellent correlation between fluorescence signal and DNA concentration.

In order to obtain a high-performance BRET-based system for DNA recognition, the energy transfer efficiency of NanoLuc to thiazole orange was evaluated, estimating the Förster distance ( $R_0$ ), i.e., the distance where the energy transfer efficiency is 50%. The value of  $R_0$  depends on the orientation factor ( $\kappa$ ), the quantum yield of the donor (QD), the refractive index of the medium ( $n$ ), and the spectral overlap of donor and acceptor (J).  $R_0$  can be calculated using the following equation:

$$R_0 = 0.21[\kappa^2 Q_D n^4 J(\lambda)]^{1/6}$$

The major determinant for the  $R_0$  of BRET pairs includes its spectral overlap (J). The thiazole orange has the expected emission properties to act as a BRET acceptor

since its excitation spectrum overlaps with NanoLuc emission with a desired optical window from 450 to 520 nm. In addition, the Förster distance equation was used to estimate the maximal linker distance that allows for efficient BRET. Comparing their spectral overlap to known BRET pairs, indicated that the  $R_0$  should lay within the range of 3.2-4.4 nm, and therefore, linker length should preferably be shorter than this value (27).

The reactivity of NanoLuc and thiazole orange was examined (Figure SI.3). Regarding to the protein, the isoelectric point under the reaction conditions (PBS buffer, pH=7) was calculated by PypKa software being 5.37. Modeling with ChimeraX software reported that many nonpolar and aromatic residues were located on the protein surface. Structural analysis of the intercalating dye was also performed using MarvinTool software. The dye is a flat molecule with a polar extreme and an extensive aromatic region and reactive groups, such as the N-alkylated and the NHS ester. The conclusions were that the dye can be conjugated to the newly synthesized luciferase by three conjugation ways, though, (i) the lysine residues, (ii) the cysteine residue, and (iii) aggregation around the cysteine residue.

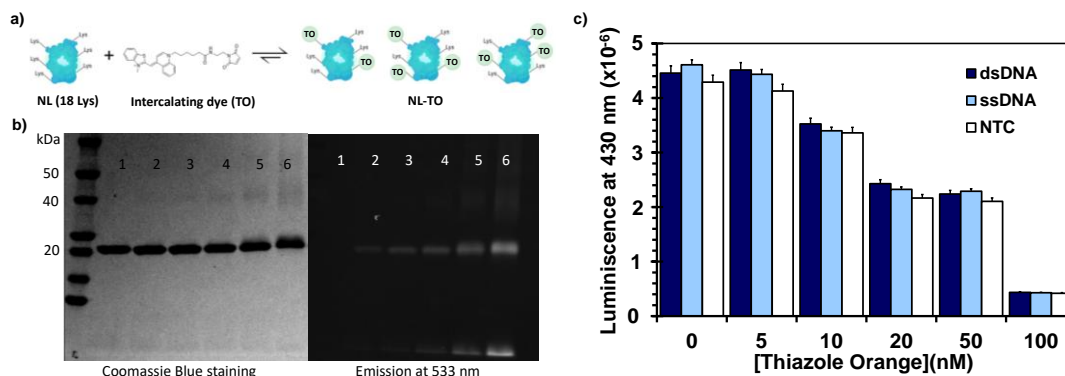
#### **Conjugation via lysine residues**

The primary amines of lysine residues offered a simple strategy to conjugate the dye, forming a stable peptide bond under a slightly basic solution (Figure 3). As NanoLuc has 18 possible attachment points (Lys), the conjugation was non-specific. The products were separated by SDS-PAGE and Coomassie blue staining, and the expected band of protein was distinguished in 22 kDa (lane 1). Increasing molar ratio of the dye, bands shifted to higher mass, indicating that multiple dye molecules were conjugated (lanes 2-6).

Imagining the gel at dye emission wavelength (533 nm), no fluorescence background was detected in the absence of reagent (lane 1). For conjugate solutions, bands at the same locations as in the protein staining were observed (lanes 2-6). In addition, another band below was distinguished corresponding to free reagent. Hence, results confirmed the correct coupling and the formation of a heterogeneous product in terms of dye molecules per protein.

The conjugates were incubated with DNA, but BRET emission did not happen and any differences between the ssDNA and dsDNA were distinguished (Figure 3c). Increasing the initial dye concentration, the background luciferase emission decreased; thus, a quenching effect was happening. The same experiments were performed with the unconjugated NanoLuc and dye, and the luciferase signal did not diminish (Figure

SI.4). Results suggested that the quenching effect was intramolecular rather than intermolecular due to the free reagent excess in the solution (Figure SI.5).



**Figure 3.** NanoLuc-thiazole orange conjugates using lysine residues. a) Scheme of the reaction between primary amines and NHS ester using PBS at pH=8 b) SDS-PAGE analysis of the products. Concentration of NanoLuc (22 kDa) was 10  $\mu$ M and variable concentration of thiazole orange (0.5 kDa) 1: 0  $\mu$ M; 2: 5  $\mu$ M, 3: 10  $\mu$ M, 4: 20  $\mu$ M; 5: 50  $\mu$ M; 6: 100  $\mu$ M. c) Luciferase emission peak of the reaction solutions employing different types of DNA.

### Conjugation via a cysteine residue

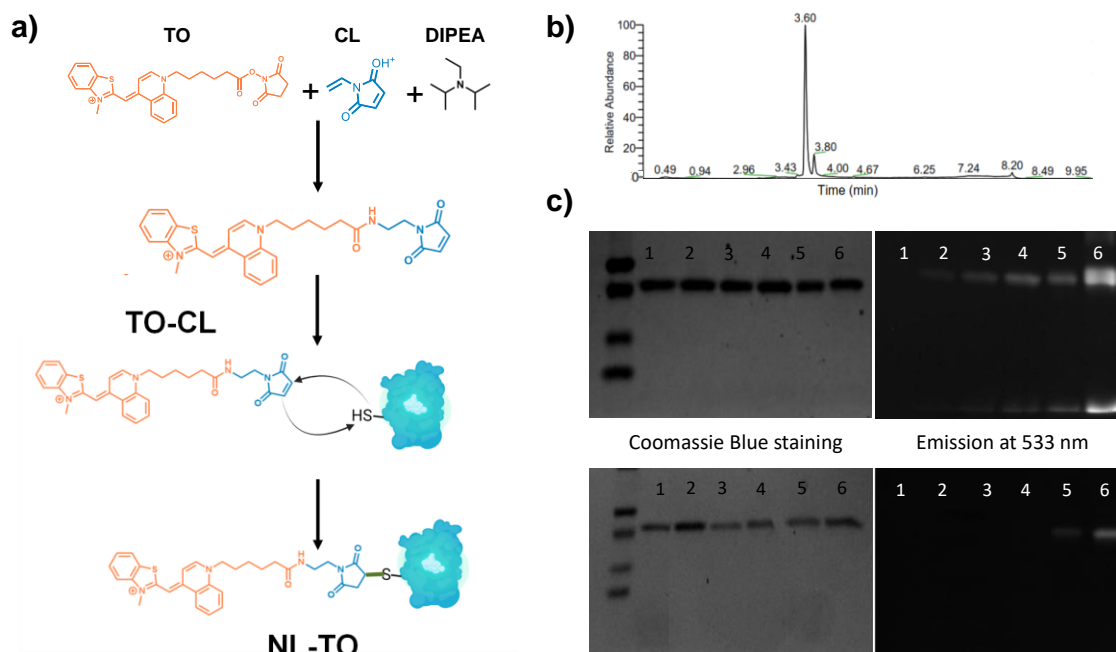
The second strategy was focused on achieving a specific conjugation through the cysteine residue of the mutant luciferases (Figure 4a). 1-(2-Aminoethyl) maleimide hydrochloride was chosen as a crosslinker because this heterobifunctional linker offers a thiol-reactive maleimide and a carboxyl-reactive hydrazide group. The total linker length is approximately 3-4 nm, which was still beneath the estimated  $R_0$ , being compatible with a BRET effect (27).

The first step was the formation of the intermediate between the crosslinker and the dye. The products formed under different conditions were separated by LC-MS. The crosslinker, dye, and intermediated produced peaks at 0.43, 3.74, and 3.60 min, respectively (Figure SI.6). In excess of crosslinker or thiazole orange, many impurities were detected or high amount of free reagents (Figure SI.7). The best chromatogram corresponded when the crosslinker acted as the limiting reagent with a stoichiometric ratio of 1:0.8 in presence of N,N-Diisopropylethylamine (DIPEA) at 3 mM (Figure 4b).

The next step was the addition of Tris buffer to deactivate the unspecific conjugation between the lysine groups and free dye excess. The aim was to avoid the consequent quenching phenomenon in the luciferase emission, as described before. Later, the coupling was produced between the maleimide group of intermediate with the thiol group of the luciferase tail under near-neutral conditions. The conjugates were checked by SDS-PAGE (Figure 4c). For unpurified products, bands associated with the free dye were detected, indicating an important reagent excess. For purified products,

the band intensity of the conjugated products depended on the stoichiometric ratio, being the best result a 10-fold excess of thiazole orange.

The BRET effect was studied by incubating the conjugated products with DNA standards. The complex between NanoLuc-thiazole orange conjugate and DNA molecules emitted low signal being insufficient intensity to detect biomarkers at low concentrations.



**Figure 4.** NanoLuc-thiazole orange conjugates using cysteine residue. a) Scheme of the reaction using PBS at pH=7. b) HPLC chromatogram obtained from the conjugation between crosslinker and dye using a stoichiometric ratio of 1:0.8 in 3 mM of DIPEA. c) SDS-PAGE analysis of unpurified products (up), and purified products (down). The concentration of NanoLuc (22 kDa) was 10  $\mu$ M and concentration of thiazole orange (0.5 kDa) was 0  $\mu$ M (1), 5  $\mu$ M (2), 10  $\mu$ M (3), 20  $\mu$ M (4), 50  $\mu$ M (5), and 100  $\mu$ M (6).

### Conjugation via a cysteine residue and complex formation

A way to increase the number of intercalating agent molecules per protein was the formation of the protein-dye complex. Cyanine dyes are well known to form supramolecular aggregates in aqueous solution. Due to the hydrophobicity and polarizability of the molecules,  $\pi$ - $\pi$  stacking of the conjugated systems of the dyes is favored (28). Parameters such as dye concentration, structure, ionic strengths, temperature influence the ability to form aggregates in the solutions (29). Hence, considering the structure of thiazole orange, the non-covalent intermolecular interactions between intercalating dye molecules for multiple  $\pi$ - $\pi$  stacking were predicted.

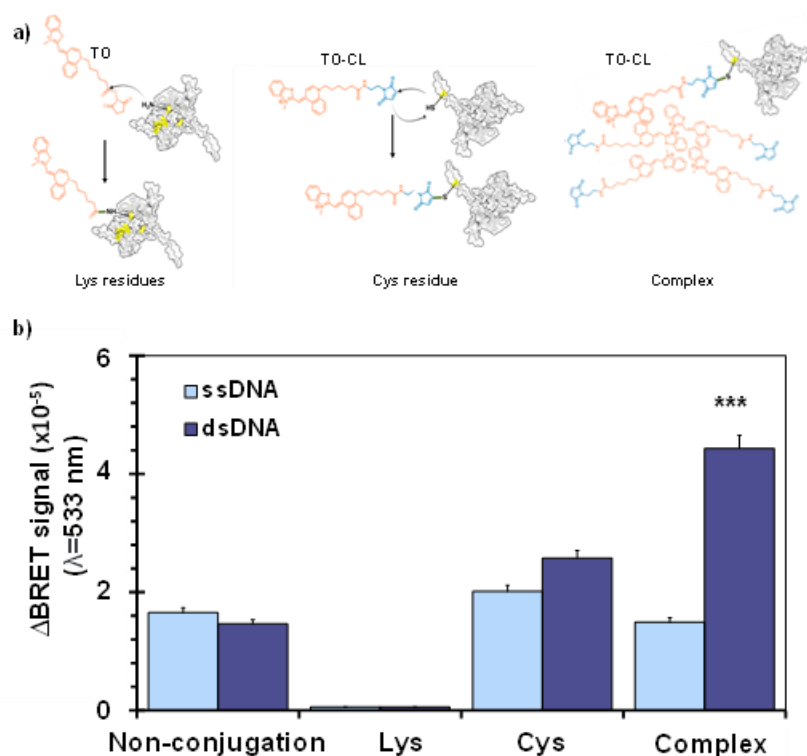
Conjugates based on the coupling of cysteine residues through a crosslinker were prepared as described above but in presence of a 100-fold excess of dye. Thus, multiple dye molecules interacted with NanoLuc through the coupled dye, based on

intermolecular non-covalent bonds. In this case, the number of dye molecules arranged around the luciferase was undefined. With this approach, a specific BRET effect was achieved since dsDNA was distinguished from ssDNA.

As a process control, solutions of free thiazole orange and unconjugated luciferase in the same proportions were prepared. The objective was to discriminate the direct non-covalent interaction arranged around the protein and the interaction mediated by the Cys-coupled thiazole orange. The study of BRET emission in the presence of DNA concluded that a lower emission than complex products was registered. Therefore, the energy transfer was favored through the  $\pi$ - $\pi$  stacking to conjugated dye molecule.

### **Comparison of BRET effect**

The synthesized bioluminescent reagents showed differences in the conjugation position and the maximum number of intercalating agent molecules per protein, being 18 for lysine conjugation, 1 for cysteine conjugation, and undefined for complex formation. The evaluation of their biosensing capability was performed by a paired bioluminescent assay. The BRET effect was studied by incubating the NanoLuc-thiazole orange conjugates with the ssDNA and dsDNA standards and registering emission signals at 533 nm (Figure 5). Good biosensing systems were those that provided a high signal variation between dsDNA (specific response) and ssDNA (unspecific response). Non-signal variation was observed for the solution composed of free thiazole orange and unconjugated luciferase. Signal variation was null using conjugates via lysine residues and low for conjugates via cysteine residue in mutant NanoLuc. The greater of signal variation corresponded to the complex product in high excess of intercalating dye (t-test p-values < 0.05). In conclusion, BRET energy transfer was achieved due to the proximity of NanoLuc to dye molecules, based on the multiple intermolecular interactions mediated by the cystein-coupled thiazole orange.

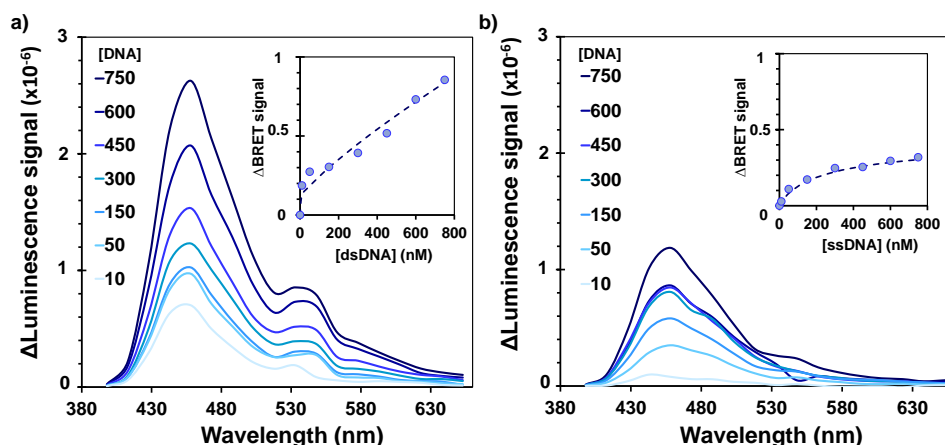


**Figure 5.** Comparison of NanoLuc-thiazole orange conjugates obtained following different reaction pathways. a) Scheme of conjugation reactions. b) BRET emission using 108-pb DNA (1  $\mu$ M) and 10 nM of the conjugate. Complex formation using 1:100 excess of thiazole orange. The control was the non-conjugated mix composed of NanoLuc at 10 nM and thiazole orange at 1:100 excess.

The reaction conditions were examined in order to enhance the BRET signal. The studied variables were the concentration of bioluminescent reagent and the stoichiometric ratio of dye (Figure SI.8). The best results were obtained for 30 nM of the conjugate and a dye excess of 1:10,000.

The biosensing sensitivity of the new reagent was evaluated by dilution series of DNA standards, registering the bioluminescent spectrum and calculating the calibration curves (Figure 6). Although with ssDNA signal was detected, the system depicted a high BRET intensity for dsDNA with a detection limit of 10 nM (approx.  $10^{10}$  copies). Data were fitted to 4-parameter logistic regression equation.

Reproducibility was determined from replicated assays and expressed as a relative standard deviation. The values ranged from 1.6 to 3%, confirming the signal robustness produced NanoLuc-thiazole orange complex. In addition, the conjugate had a highly thermal stability as it retained the activity at least 30 min below 50  $^{\circ}$ C. In these experiments, a plate reader registered BRET emission, but the readout could be performed with a digital camera due to its high brightness.



**Figure 6.** Effect of DNA concentration in the luminescence signal using the conjugate complex. a) 108-bp dsDNA. B) 108-bp ssDNA. The signal was obtained by subtracting the conjugates with different DNA amounts minus the conjugate with the blank (non-DNA). Calibration curve was  $\text{Signal} = 1.04 \cdot 10^5 + 1.05 \cdot 10^{10} / (1 + ([\text{DNA}] / 8.76 \cdot 10^7)^{-0.82})$ ,  $R^2 = 0.95$  for dsDNA.

#### End-point RPA-BRET method. Application to oncogene detection

One of the main applications of intercalating agents is quantifying the copy number of a DNA biomarker, especially copies generated after finishing a replication process. This study examined the detection of products from an amplification reaction by measuring the BRET effect based on NanoLuc for the first time. The proposed strategy is called RPA-BRET, consisting of the addition of NanoLuc-thiazole orange complex to the solution will all reagents required to perform this isothermal amplification at 37 °C. As proof-of-concept, the bioluminescence detection of the selective amplification of the *KRAS* gene (codon 12-13) from serial dilutions of genomic DNA (gDNA) was studied. RPA-BRET method was applied to samples with different copy number of the human gDNA (Figure 7a). After 30 min of amplification, negative samples showed similar responses to controls. A quantitative response for the BRET band at 533 nm was observed according to the initial template copies. The calculated detection limit was 26 copies of gDNA, confirming that the amplification efficiency of RPA was about 8-9 orders (30,31). The sensitivity was similar or better than some fluorescent dyes used in end-point RPA (Figure 7b), demonstrating that analytical performances of RPA-BRET method were suitable for many applications.

#### Real-time RPA-BRET method. Application to oncogene detection

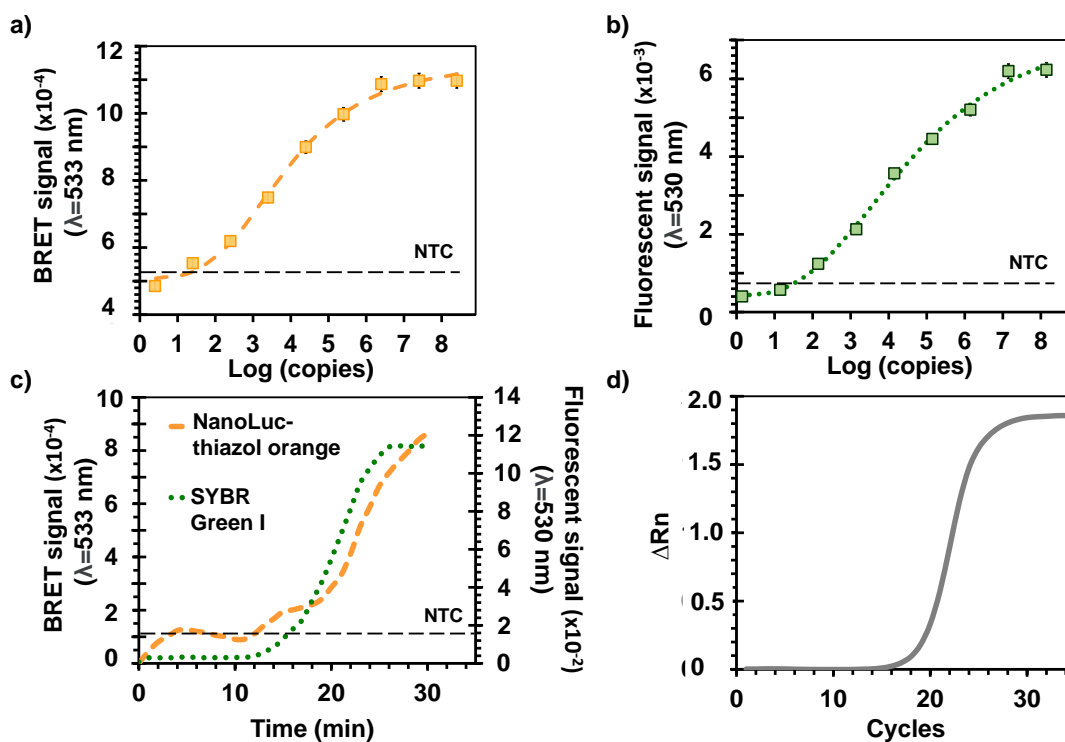
The new bioluminescent reagent was applied to monitor the isothermal amplification in real-time mode, quantifying the copy number generated during the entire amplification. As proof-of-concept, a bioluminescence method, called real-time RPA-BRET, was developed and applied to monitor the products of the *KRAS* gene (codon 12-



13) from serial dilutions of gDNA. As shown in Figure 7c, bioluminescence signal ( $\lambda_{\text{emission}}=533$  nm) increased, as a function of time, enabling real-time monitoring of the RPA reaction. In parallel, the RPA products were detected using the intercalating agent SYBR Green I and a fluorescence detector ( $\lambda_{\text{excitation}}=498$  nm and  $\lambda_{\text{emission}}=522$  nm). Both sensing methods reported similar behavior, fitting to a sigmoid growth curve, but the light emitted from the novel developed system was 66-fold more intense than the traditional dye.

The real-time RPA-BRET method was compared to qPCR, the gold-standard method of DNA biomarkers. Figure 7d shows the real-time PCR graph was similar than the obtained by real-time RPA-BRET. The difference lay in the amplification technology, the transduction principle, and consequently, the requirements of each case. An essential drawback of PCR-based methods is that a high percentage assay time is due to temperature cycling rather than amplification, slowing down the assay time (<1 h) (32). Besides, RPA-based methods provide a faster response since they work at constant low temperature (37 °C), obtaining the results in a shorter time frame (30 min). For qPCR, a real-time thermal cycler and fluorescent readout were required. RPA detected by NanoLuc-thiazole orange complex needed a heater and an emission reader. In this way, the requirement for an external excitation source was eliminated, avoiding potential photobleaching, generation of toxic radicals, and high background signal caused by autofluorescence and scattering. Therefore, equipment is low-cost and can be transported in a mobile suitcase laboratory and battery-powered, being a competitive solution even compared to portable qPCR instruments.

Real-time RPA-BRET method also overcomes the current state-of-art for monitoring of isothermal amplifications (4,6,30). The bioluminescence method called LAMP-BART is less sensible because it does not rely on the BRET phenomenon and needs a thermostable firefly luciferase. Also, the method is less simple, considering to a complex primer design and a higher working temperature (60-65°C).



**Figure 7.** Different transducers modes of biomarker DNA amplification. a) Assay sensitivity of the end-point RPA-BRET method b) Assay sensitivity by conventional end-point fluorescence method. c) Real-time RPA monitoring BRET signal by NanoLuc-thiazole orange (orange) and fluorescence by SYBR Green I (green). d) Real-time qPCR monitoring with fluorescence by SYBR Green I. Amplification product of *KRAS* gene codon 12-13 from human gDNA.

### 3.4.5 Conclusions

A novel BRET reagent obtained from the conjugation of a mutant NanoLuc, a very bright luciferase, and a thiazole orange, an intercalating dye, has been developed. One of the attractive features is that it can be applied for a universal DNA biomarker detection as the intercalating dye binds nonspecifically. The number of potential applications in molecular diagnosis is broad. Nevertheless, considering analytical performances and operational requirements, the relevant challenge is detecting certain regions and determining copy number for specific genes. As far as we know, the developed methods are the first bioluminescent assays based on an isothermal amplification technique using standard bioluminescent proteins compatible with point-of-care demands. In fact, this proof-of-concept study demonstrates the effective detection of RPA products in end-point and real-time modes, but the extension to other amplification techniques is feasible. The main limitation is the conjugate stability because temperatures higher than 50 °C denaturalize the enzyme reducing its activity. Thus, the new reagent is valid for all amplifications (PCR or isothermal) in end-point mode, and isothermal techniques in real time if the working temperature is low. On the other hand,

future experimentation can be aimed to exploit this new approach, improving features of the bioluminescent transduction. For instance, the new studies can be focused on constructing multiples cysteine in the luciferase sequence for increasing the conjugation capacity.

The detection of DNA biomarkers plays a crucial role as it allows determining the individual's genetic profile to support disease prevention, diagnosis, and treatment. Following the developed methodology, DNA biomarker detection can be performed using reduced resources, facilitating the implementation of personalized medicine at decentralized laboratories.

### 3.4.6 References

1. Ziegler A, Koch A, Krockenberger K, Großhennig A. Personalized Medicine Using DNA Biomarkers: A Review. *Hum Genet.* 2012;131(10):1627–38.
2. Zhang P, Beck T, Tan W. Design of a Molecular Beacon DNA Probe with Two Fluorophores. *Angew Chem.* 2001;113(2):416-9.
3. Campbell R E. Fluorescent-Protein-Based Biosensors: Modulation of Energy Transfer as a Design Principle. *Anal Chem.* 2009;81(15):5972–9.
4. Tian CJ, Lin ZX, He XM, Luo Q, Luo CB, Yu HQ, et al. Development of a fluorescent-intercalating-dye-based reverse transcription loop-mediated isothermal amplification assay for rapid detection of seasonal Japanese B encephalitis outbreaks in pigs. *Arch Virol.* 2012;157(8):1481-8.
5. Sharifian S, Homaei A, Hemmati R, Luwor RB, Khajeh K. The Emerging Use of Bioluminescence in Medical Research. *Biomed Pharmacother.* 2018;101:74–86.
6. Hardinge P, Baxani DK, McCloy T, Murray JAH, Castell OK. Bioluminescent Detection of Isothermal DNA Amplification in Microfluidic Generated Droplets and Artificial Cells. *Sci Rep.* 2020;10(1):1–14.
7. Mirasoli M, Bonvicini F, Lovecchio N, Petrucci G, Zangheri M, Calabria D, et al. On-chip LAMP-BART reaction for viral DNA real-time bioluminescence detection. *Sens Actuators B Chem.* 2018;262:1024-33.
8. Hiblot J, Yu Q, Sabbadini MDB, Reymond L, Xue L, Schena A, et al. Luciferases with Tunable Emission Wavelengths. *Angew Chem.* 2017;129(46):14748–52.
9. Hall MP, Unch J, Binkowski BF, Valley MP, Butler BL, Wood MG, et al. Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. *ACS Chem Biol.* 2012;7(11):1848–57.
10. England CG, Ehlerding EB, Cai W. NanoLuc: A Small Luciferase Is Brightening Up the Field of Bioluminescence. *Bioconjug Chem.* 2016;27(5):1175–87.
11. Biewenga L, Rosier BJHM, Merckx M. Engineering with NanoLuc: A Playground for the Development of Bioluminescent Protein Switches and Sensors. *Biochem Soc Trans.* 2020;48(6):2643–55.
12. Boute N, Jockers R, Issad T. The Use of Resonance Energy Transfer in High-Throughputscreening: BRET versus FRET. *Trends Pharmacol Sci.* 2002;23(8):351–4.
13. Bacart J, Corbel C, Jockers R, Bach S, Couturier C. The BRET Technology and Its Application to Screening Assays. *Biotechnol J.* 2008;3(3):311–24.

14. Saito K, Nagai T. Recent Progress in Luminescent Proteins Development. *Curr Opin Chem Biol.* 2015;27:46–51.
15. Xia Z, Rao J. Biosensing and Imaging Based on Bioluminescence Resonance Energy Transfer. *Curr Opin Biotechnol.* 2009;20(1):37–44.
16. Renault K, Fredy J W, Renard PY, Sabot C. Covalent Modification of Biomolecules through Maleimide-Based Labeling Strategies. *Bioconjug Chem.* 2018;29(8):2497–513.
17. Suzuki K, Kimura T, Shinoda H, Bai G, Daniels MJ, Arai Y, et al. Five Colour Variants of Bright Luminescent Protein for Real-Time Multicolour Bioimaging. *Nat Commun.* 2016;7(1):1-10.
18. Moutsopoulos A, Hunt E, Broyles D, Pereira CA, Woodward K, Dikici E, et al. Bioorthogonal Protein Conjugation: Application to the Development of a Highly Sensitive Bioluminescent Immunoassay for the Detection of Interferon- $\gamma$ . *Bioconjug Chem.* 2017;28(6):1749–57.
19. Ni Y, Arts R, Merckx M. Ratiometric Bioluminescent Sensor Proteins Based on Intramolecular Split Luciferase Complementation. *ACS Sens.* 2019;4(1):20–5.
20. Yoshida W, Kezuka A, Abe K, Wakeda H, Nakabayashi K, Hata K, et al. Detection of Histone Modification by Chromatin Immunoprecipitation Combined Zinc Finger Luciferase-Based Bioluminescence Resonance Energy Transfer Assay. *Anal Chem.* 2013;85(13):6485–90.
21. Engelen W, van de Wiel KM, Meijer LHH, Saha B, Merckx M. Nucleic Acid Detection Using BRET-Beacons Based on Bioluminescent Protein-DNA Hybrids. *Chem Commun.* 2017;53(19):2862–65.
22. Fleiss A, Sarkisyan KSA. Brief Review of Bioluminescent Systems. *Curr Genet.* 2019;65(4):877–82.
23. Lobato IM, O'Sullivan CK. Recombinase Polymerase Amplification: Basics, Applications and Recent Advances. *Trends Analyt Chem.* 2018;98:19–35.
24. Das V, Kalita J, Pal M. Predictive and Prognostic Biomarkers in Colorectal Cancer: A Systematic Review of Recent Advances and Challenges. *Biomed Pharmacother.* 2017;87:8–19.
25. Weihs F, Wang J, Pflieger KDG, Dacres H. Experimental Determination of the Bioluminescence Resonance Energy Transfer (BRET) Förster Distances of NanoBRET and Red-Shifted BRET Pairs. *Anal Chim Acta X.* 2020;6:100059. doi: 10.1016/J.ACAX.2020.100059
26. Nygren J, Svanvik N, Kubista M. The Interactions Between the Fluorescent Dye Thiazole Orange and DNA. *Biopolymers.* 1998;46(1):39–51.
27. Dacres H, Wang J, Dumancic MM, Trowell SC. Experimental Determination of the Förster Distance for Two Commonly Used Bioluminescent Resonance Energy Transfer Pairs. *Anal Chem.* 2010;82(1):432–5.
28. Armitage BA. Cyanine Dye-DNA Interactions: Intercalation, Groove Binding, and Aggregation. In: Waring MJ, Chaires JB, editors. *DNA Binders and Related Subjects. Topics in Current Chemistry.* Berlin: Springer; 2005.p. 55-76. doi: 10.1007/b100442
29. Place I, Perlstein J, Penner TL, Whitten DG. Stabilization of the Aggregation of Cyanine Dyes at the Molecular and Nanoscopic Level. *Langmuir.* 2000;16(23):9042-8.
30. Santiago-Felipe S, Tortajada-Genaro LA, Puchades R, Maquieira Á. Parallel Solid-Phase Isothermal Amplification and Detection of Multiple DNA Targets in Microliter-Sized Wells of a Digital Versatile Disc. *Mikrochim Acta.* 2016;183(3):1195–1202.
31. Lázaro A, Yamanaka ES, Maquieira Á, Tortajada-Genaro LA. Allele-Specific Ligation and Recombinase Polymerase Amplification for the Detection of Single Nucleotide Polymorphisms. *Sens Actuators B Chem.* 2019;298:126877. doi: 10.1016/j.snb.2019.126877

32. Lázaro A, Tortajada-Genaro LA, Maquieira Á. Enhanced Asymmetric Blocked QPCR Method for Affordable Detection of Point Mutations in KRAS Oncogene. *Anal Bioanal Chem.* 2021;413(11):2961-9.

### 3.4.7 Supplementary Information

#### Selection of reagents

##### Luciferase

The structure of the synthesized luciferases is shown in Figure SI.1. The Chimera X software program (UCSF) has been used for its modeling. To predict the protein structure from its sequence, the Alpha fold was used. It is a computational prediction using Artificial Intelligence leveraging multi-sequence alignments into the design of the deep learning algorithm to predict the protein's structure from its amino acid sequence.

Compared to wild-type protein, distinctive features are: mutagenesis of C164S, streptavidin tag at the N-terminal end, and 16-amino acid tail at the C-terminal end. The tail at the C-terminal end serves as a scaffold to introduce the new mutant cysteine, producing two proteins depending on its position: G180C for NanoLuc A and A170C for NanoLuc B. Also, the last residues are five histidine for the later protein purification.

##### Intercalating dye

Different intercalating agents with fluorescence properties, reviewed in Table SI.1, have been used for nucleic acid detection (1,2). The selection criteria were the optical features and structural properties for an intense emission and an effective coupling between the intercalating agent and the luciferase, respectively.

Considering the quantum yield and the LOD, a candidate molecule was SYBR Green I. However, luciferase coupling was inefficient considering its structure (3). Thiazole orange is a fluorescent asymmetric cyanine dye formed by two heterocyclic ring structures linked through a methine bridge. It presents two different conformations that determine the fluorescence intensity. A planar conformation that allows conjugation between the two aromatic systems, giving rise to the fluorescent form, and a non-planar state produced by rotation of the methine bridge, which is found in solution and does not generate fluorescence.

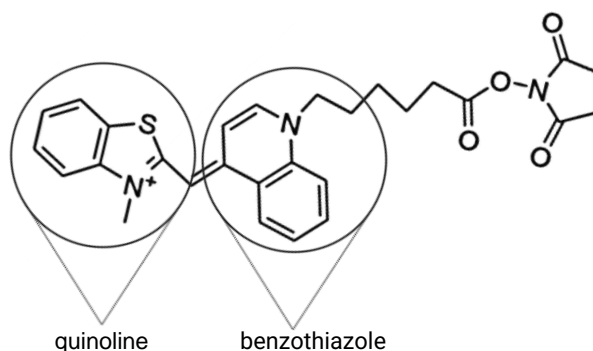


**Table SI.1.** Overview intercalating agents characteristics. b) Structure of thiazole orange N-hydroxysuccinimide ester.

a)

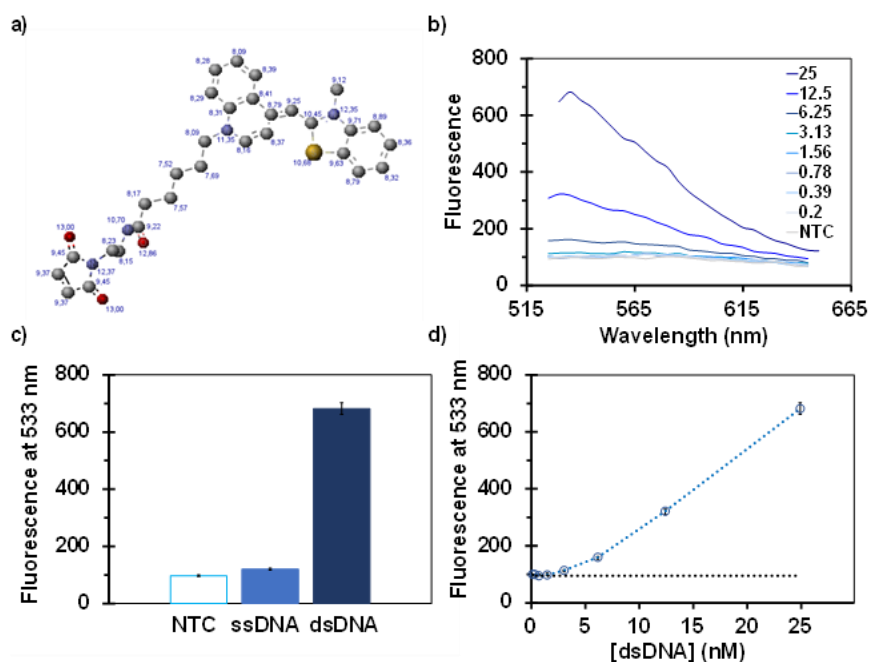
Dye	Thiazole orange	SYBR Green I	Evagreen	YOYO-1	Picogreen
$\lambda_{max}$ , excitation (nm)	512	496	500 (DNA bound), 471 (without DNA)	491	502
Affinity dsDNA	Weak	High	High	Very high, irreversible	High
LOD	$\mu$ M	60 pg/ $\mu$ L	1 copy/ $\mu$ L	2.5 ng/mL	250 pg/mL
Increase in fluorescence (fold)	2000	1000	Unknown	100-1000	1000
Quantum yield	$\approx$ 0.4	$\approx$ 0.8	Unknown	$\approx$ 0.2-0.6	Unknown
Absorbance free in solution	Yes	Yes	Yes	Yes	Yes
Coupling possibilities	Yes (NHS ester)	Low	Low	Low	Low, similar to SYBR Green
The commercially available activated form.	Yes	No	No	No	No

b)



The thiazole orange intercalates between the double helix, stabilizing its fluorescent conformation by stacking between the DNA base pairs (4). The spectrum of has an excitation peak at 514 nm, an emission peak at 533 nm, and it can be excited using a 488 nm laser. The first experiment was the study of the binding affinity. For that, a fluorescent titration experiment using a 108 pb-DNA was performed (Figure SI.2). When double-strand DNA was not present, a negligible signal was reported, which indicated that TO did not show intrinsic autofluorescence, and the fluorescent signal increased based on dsDNA concentration. The signal was higher with significant differences until reached a DNA concentration of 1.56 nM, which was the detection limit (Student's t-test p-values < 0.05). The fluorescence obtained with ssDNA and dsDNA of the same length was compared to confirm the intercalating dye performances. The fluorescence intensity was 5.7- folds higher when dsDNA was present.





**Figure SI.2.** Thiazole orange succinyl ester as an intercalating agent. a) Tridimensional structure showing its atoms and their orbital electronegative. b) Fluorescence spectrum with different dsDNA concentration (nM). c) Affinity of the dye for the different types of DNA (25 nM). d) Calibration curve with dsDNA at TO emission. Using 5  $\mu$ M TO and 108 bp-DNA for single and double-strand. NTC: absence of DNA.

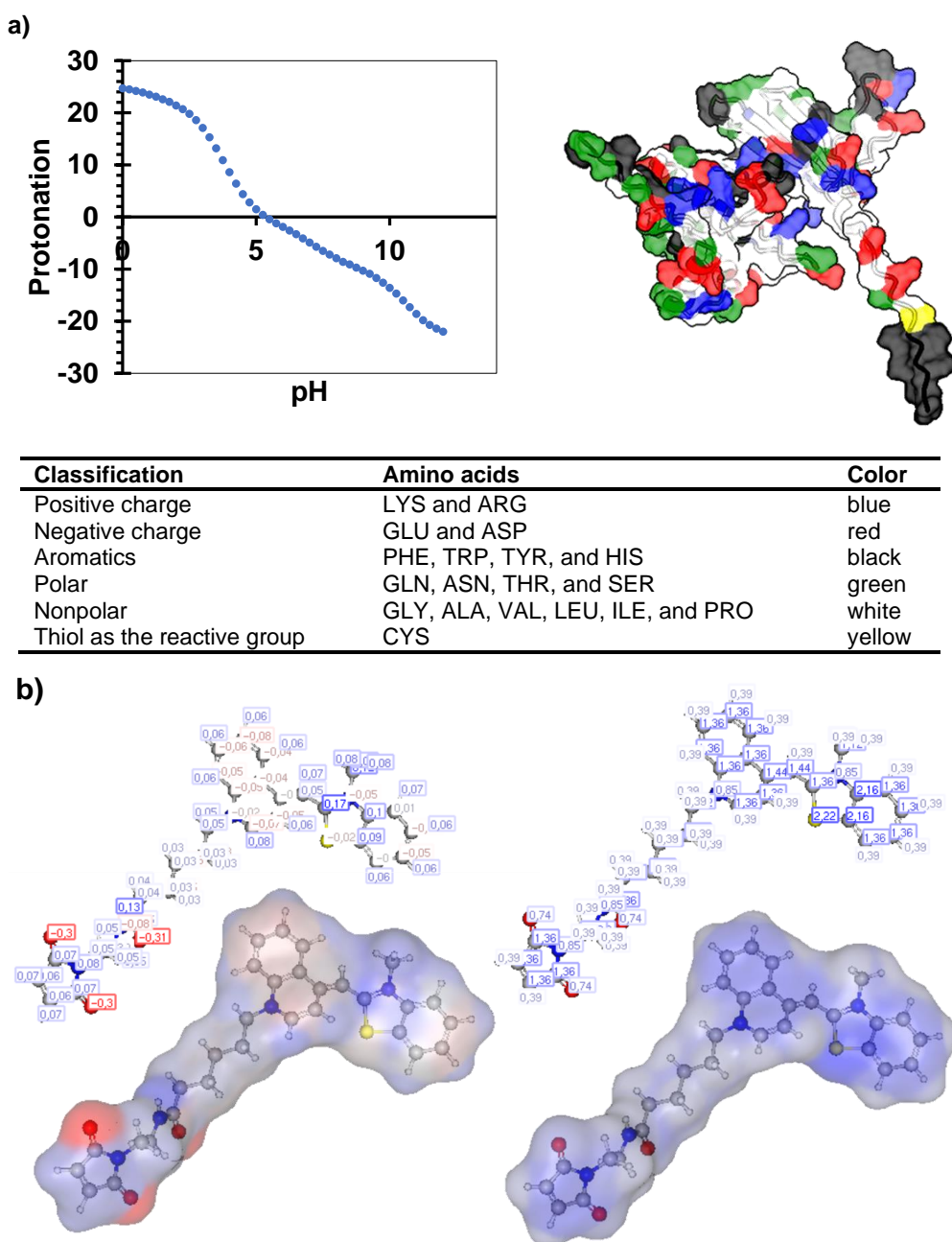
### Reactivity

The conjugation performances of the NanoLuc and thiazole orange were evaluated by analyzing their structures (Figure SI.3). PypKa software calculated that the isoelectric point of synthesized proteins was 5.37. According to the model obtained with ChimeraX, the protein has 38 ionic residues (15 positively and 23 negatively charged) for ionic interactions, 40 polar residues for forming hydrogen bonds (including the cysteine introduced by mutagenesis to attach the thiazole covalently), and 119 nonpolar residues with low reactivity (89 aliphatics and 30 aromatics) (Figure SI.3a).

The intercalating dye was analyzed using MarvinTool software (Figure SI.3b) by studying its polarity and hydrophobicity. It is a planar molecule with two distinct zones, a more polar zone at one end, and a more hydrophobic zone with non-polar and aromatic groups exposed on the surface that allow  $\pi$ - $\pi$  stacking interactions between the dye molecules. Consequently, the interaction of several intercalating agent molecules per protein molecule is possible.

Therefore, there are different reactive sites for covalent coupling the intercalating agent with NanoLuc. First, thiazole orange can couple nonspecifically to the primary amines of the lysine residues (18). Second, the dye can specifically couple to the thiol of

cysteine (1), allowing excellent specificity about the site and number of conjugations. Third, on the dye attached to the cysteine residue, several dye molecules can be coupled via  $\pi$ - $\pi$  stacking.



**Figure SI.3.** Reactivity of the reagents. a) NanoLuc: (left) titration curve calculated by PypKa, (right) surface model colored by charge using ChimeraX, and (bottom) classification of amino acid residues. b) Thiazole orange: (left) charge distribution: negative in red and positive in blue. (right) Polarizability (molecular= 58.30).

### Conjugation via Lysine residues

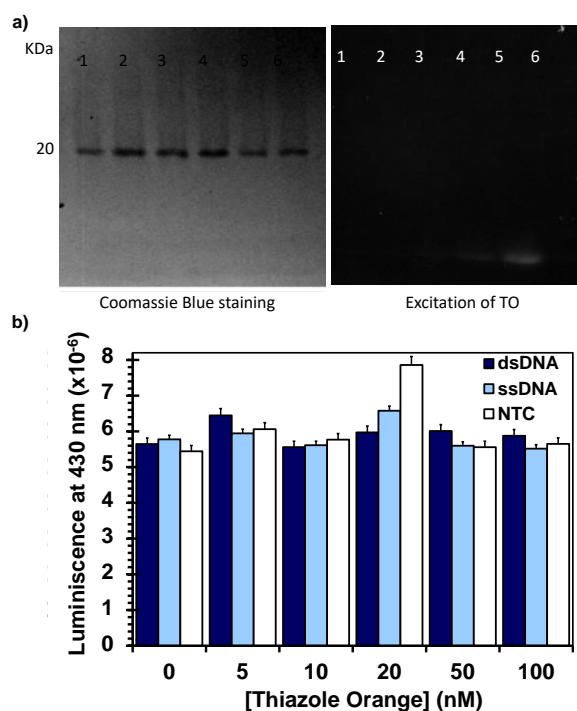
**Procedure.** 10  $\mu$ M NanoLuc and different amounts of thiazole orange (0-100  $\mu$ M) were mixed in the sodium phosphate buffer (PBS) (0.1 M sodium phosphate, 0.15 M NaCl) at pH=8. The conjugation reaction was incubated for 2 h at 4 °C with continuous

shaking and darkness. As a control, for deactivating the dye and avoiding conjugation of the reagents, the thiazole orange was preincubated with Tris buffer (500 mM Tris HCl at pH=8.0) and subsequently with NanoLuc in PBS at pH=8.

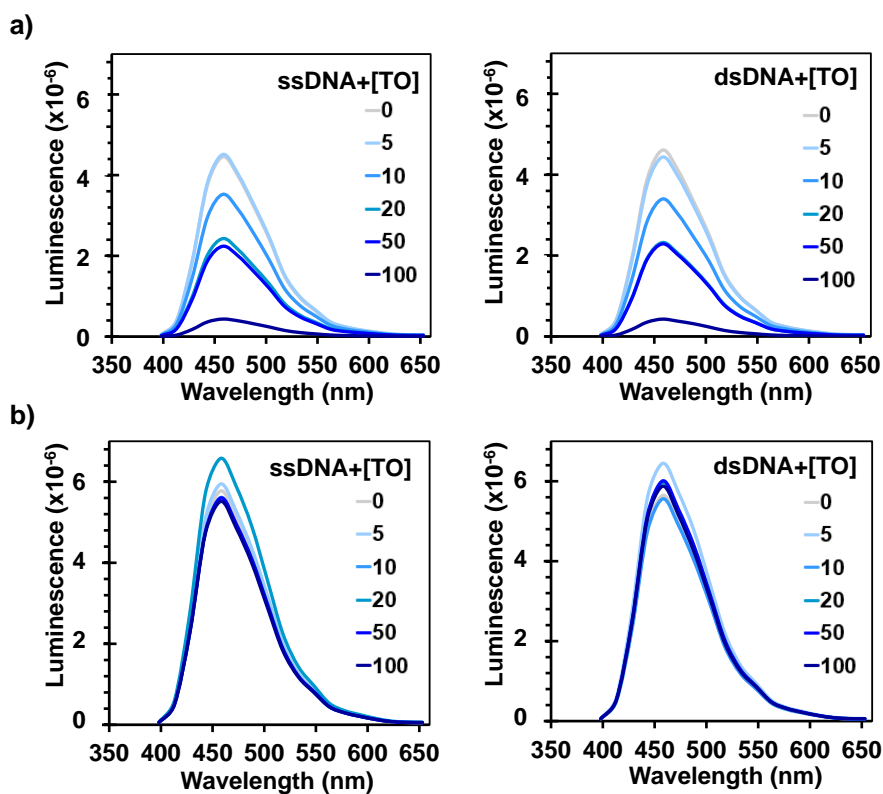
**SDS-PAGE gel electrophoresis.** The loading buffer was 180  $\mu$ L of 2x TGS (Tris-Glicine-SDS Running Buffer) with 20  $\mu$ L DTT at 1 M. The product solutions (10  $\mu$ L) and the loading buffer were mixed and incubated at 95 °C for 5 minutes. After loading the samples in 4-20 % precast protein gel (, the separation was performed for 30 min at 180 V in 1xTGS buffer. The gel was washed with miliQ water and put on a shaking rack for 15 min, changing the water each 5 min. After that, the gels were stained with Biosafe Coomassie Blue G-250 stain and washed with miliQ twice. ImageQuant 350 was used to take a picture of the gels in white light mode.

**Interpretation.** A quenching effect in the luciferase emission peak occurred when the dye amount increased in the solution. As a control, the same experiments with the unconjugated reagents were carried out deactivating the dye. The primary amines found in the Tris compete with the Lys for coupling with the NHS ester, therefore the conjugate could not be formed. The mixtures were analyzed by gel electrophoresis and the bioluminescence was measured (Figure SI.4). In the Coomassie blue staining, a 22 kDa band corresponding to the NanoLuc was clearly distinguished, but this band did not appear at thiazole orange wavelength. Only the bands corresponding to the free intercalating agent could be distinguished indicating the dye and the luciferase were not attached. On the other hand, the bioluminescence emission was registered, being constant for all deactivated dye solutions.

The luciferase spectrums of the heterogeneous product and the unconjugated reagents were compared in the presence of DNA solutions (Figure SI.5). For each case, same response was obtained regardless of whether ssDNA or sdDNA was added. Since in the control mixtures the quenching was not observed, it was concluded that the quenching effect happened due to the interaction between the Lys of NanoLuc and the NHS ester of the thiazole orange since indeed, these results were interpreted as the intercalating dye in solution was not responsible for the conjugate could not enable BRET. We hypothesized that since the conjugation with Lys is non-specific, they are close to the active center, and the dye is structurally similar to NanoLuc inhibitors (5), its coupling to Lys may block the active center of the enzyme and therefore inhibits substrate conversion.



**Figure SI.4.** Control with the unconjugated NanoLuc and thiazol orange deactivated. a) SDS-PAGE analysis of the reaction solutions. Concentration of NanoLuc (22 kDa) was 10  $\mu$ M and variable concentration of thiazole orange (0.5 kDa) 1: 0  $\mu$ M; 2: 5  $\mu$ M, 3: 10  $\mu$ M, 4: 20  $\mu$ M; 5: 50  $\mu$ M; 6: 100  $\mu$ M. b) Luciferase emission peak of the reaction solutions employing different types of DNA.



**Figure SI.5.** Luminescence results using variable concentration of thiazol orange (TO) in the solution. Concentration of NanoLuc was 10 nM with 1  $\mu$ M of 108 bp ssDNA (left) and dsDNA (right) NanoLuc-thiazole orange conjugates using lysine residues. b) Control with the unconjugated NanoLuc and thiazol orange deactivated.

### Conjugation via Cysteine residue

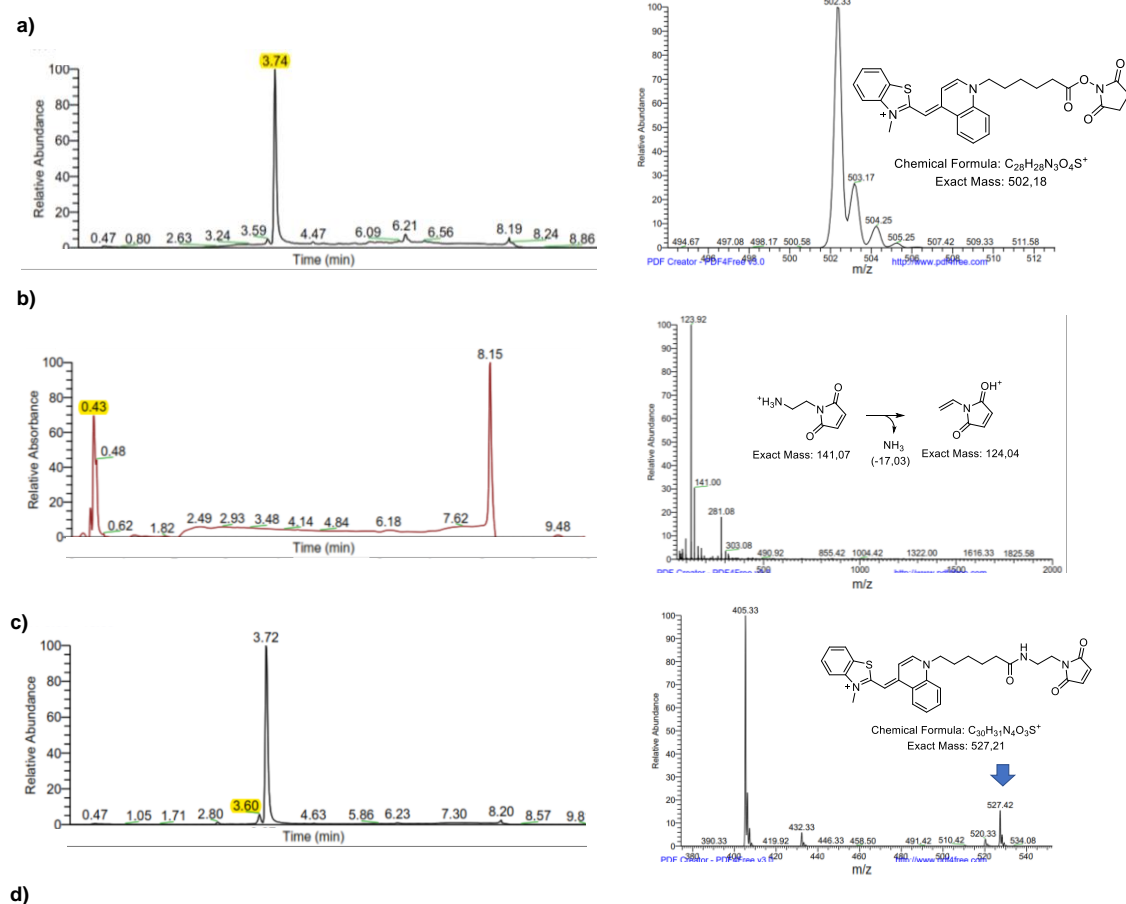
**Procedure.** The crosslinker 1-(2-aminoethyl) maleimide hydrochloride (CL) (Sigma-Aldrich) and NHS modified intercalating dye were mixed. The reagents were incubated at room temperature and darkness under different ratios (0.8-20 mM) and times (2-12 h) in DMSO. Subsequently, different amounts of the resulting intercalating dye (0-100  $\mu$ M) were added to 10  $\mu$ M NanoLuc in PBS at pH=7. The conjugation reaction was incubated for 2 h at 4 °C with continuous shaking and darkness. The conjugates were purified by filtration (Amicon 10k ultra-0.5, Merck) following the manufacturer's instructions. The dye/protein ratio of the purified conjugate was determined with absorbance signals measured by NanoDrop (ThermoFisher) and using extinction coefficients ( $M^{-1} cm^{-1}$ ) for protein 30,940 (280 nm) and for thiazole orange 8,570 (280 nm) and 63,000 (480 nm).

**Liquid chromatography-mass spectrometry.** The samples were incubated at 0.1 mg/mL in a buffer formed of 0.1 % formic acid in darkness for 2 h at 4 °C. Later they were injected on an Agilent Polaris C18A RP column with a flow of 0.3 mL/min and a 15-60% acetonitrile gradient containing 0.1% formic acid. The measuring of the mass spectra was carried on a Xevo G2 QToF mass spectrometer in positive mode. Deconvoluting the m/z spectra was performed with MaxEnt Deconvolution software.

**Interpretation.** The reagents for the synthesis of the intermediate product formed by the thiazole orange and the crosslinker were characterized on the chromatogram (Figure SI.6). Using a 1:1 ratio and 2 h incubation time only a small LC trace of the TO-CL at 3.6 min was obtained, indicating that the reaction almost did not happen. In the mass spectrum, a peak of 405.33 that corresponded to the hydrolysis of the NHS ester was observed, suggesting that NHS of the TO reacted earlier with water of the buffer than with the CL.

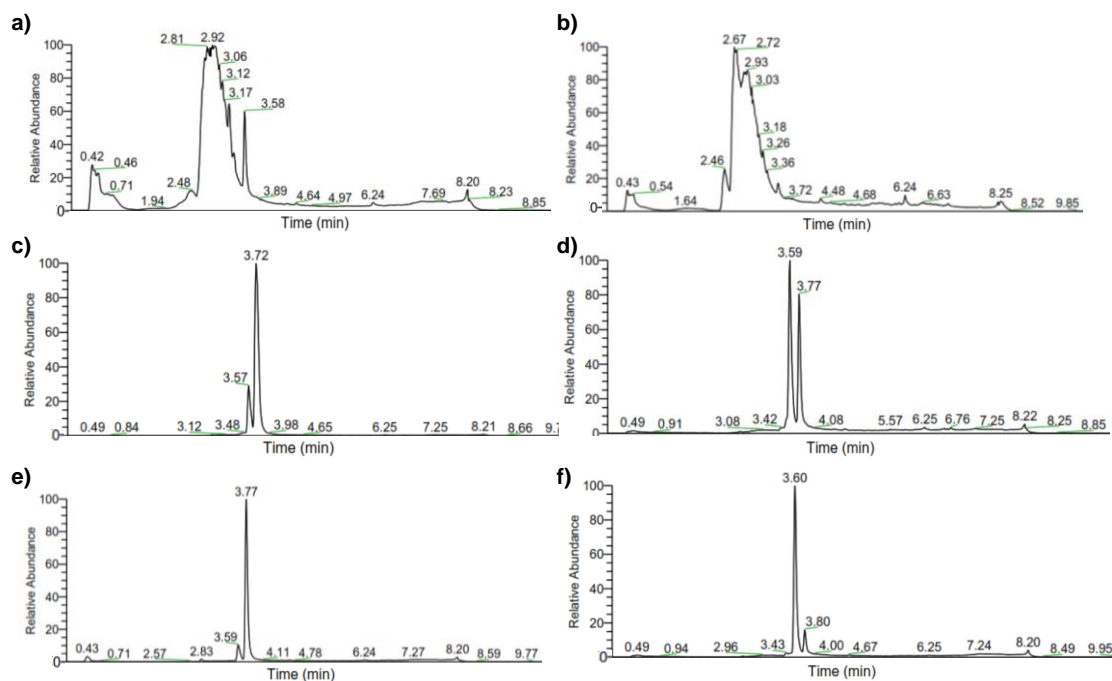
Our interpretation was that since the amine group of the CL was protonated, the reaction could not be performed. For this reason, a base, N, N-Diisopropylethylamine (DIPEA), was incorporated into the reaction solution. In addition, longer incubation time and different conditions increasing the proportion of reagents (TO: CL: DIPEA, expressed as mM) were studied to improve the reaction performance (Figure SI.7). When there was an excess of the crosslinker, a lot of undesired products came out from the chromatogram. A 20x excess of the thiazole orange produced a very small peak of the intermediary product, while a 2x excess produced a bigger peak, but still not the majority one. In the last two experiments, instead of using an equimolar amount of DIPEA, slightly more was added to ensure that all the crosslinker remained deprotonated, enhancing the reaction to occur. It was concluded that the thiazole orange should be in excess but the number of equivalents was reduced. Thus, with these

conditions, the peak area obtained in the chromatograph was small, demonstrating that the base ensured the deprotonation of the crosslinker, and enhancing the formation of the intermediate product.



Compound	Molecular formula	Retention time (min)	Mass of the most abundant isotope (g/mol)
TO-NHS	$C_{30}H_{31}N_4O_3S^+$	3.74	502.33
CL	$C_6H_8N_2O_2$	0.43	123.92
TO-CL	$C_{28}H_{28}N_3O_4S^+$	3.60	527.42

**Figure SI.6.** Functionalization of thiazole orange (TO) with crosslinker (CL). Chromatogram (left) and mass spectrum (right): a) TO-NHS. b) CL. c) TO-CL using 1:1 ratio and 2 h incubation time. d) Summary of the data obtained of each compound.

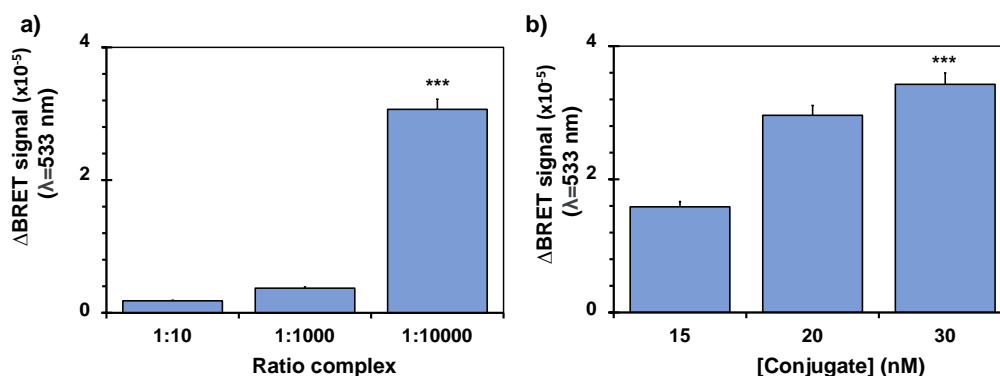


**Figure SI.7.** Chromatograms of reaction products from the functionalization of thiazole orange (TO) with crosslinker (CL). Reaction conditions: overnight incubation, ratios of TO:CL:DIPEA (mM): a) 1:10:10, b) 1:20:20, c) 20:1:3 d) 2:1:3. e) 1:0.8:0 f) 1:0.8:3.

### Conjugation via Cysteine residue and complex formation

**Procedure.** Different amounts of the intermediate product (0-100  $\mu$ M) formed by the intercalating dye and the crosslinker were added to 10  $\mu$ M NanoLuc in PBS at pH=7. The conjugation reaction was incubated for 2 h at 4  $^{\circ}$ C with continuous shaking and darkness. No purification step was performed.

**Interpretation.** This way of conjugation was the most efficient for obtaining the BRET effect between the NanoLuc and the thiazole orange in presence of dsDNA. The conditions of the reactions were tested varying the ratio for the complex formation and the conjugate amount to obtain the enhanced bioluminescent reagent (Figure SI.8). In the absence of dsDNA or the presence of single-strand DNA (ssDNA), the intercalating dye was minimally fluorescent, and only the luciferase emitted light. The conjugate enabled BRET only in the presence of double-strand DNA (dsDNA). Therefore, a POC platform for DNA recognition using a BRET-based system analog for fluorescent intercalating dyes was achieved.



**Figure SI.8.** BRET emission calculated from the difference between ssDNA (1  $\mu$ M) and dsDNA signal (1  $\mu$ M): a) Effect of complex composition. b) Effect of conjugate concentration, keeping 1:10000 excess of thiazole orange in the solution.

## References

1. Ranasinghe RT, Brown T. Ultrasensitive fluorescence-based methods for nucleic acid detection: towards amplification-free genetic analysis. *Chem. Commun.* 2011;47(13):3717-35.
2. Park CK, Hong SK, Kim YH, Cho H. Nucleic Acid-Binding Fluorochromes and Nanoparticles: Structural Aspects of Binding Affinity and Fluorescence Intensity. *Macromol Res.* 2018;26(2):204-9.
3. Dragan AI, Pavlovic R, McGivney JB, Casas-Finet JR, Bishop ES, Strouse RJ, et al. SYBR Green I: Fluorescence properties and interaction with DNA. *J Fluoresc.* 2012;22(4):1189–99.
4. Silva GL, Ediz V, Yaron D, Armitage BA. Experimental and Computational Investigation of Unsymmetrical Cyanine Dyes: Understanding Torsionally Responsive Fluorogenic Dyes. *J Am Chem Soc.* 2007;129(17):5710-8.
5. Chang D, Kim KT, Lindberg E, Winssinger N. Smartphone DNA or RNA sensing using semisynthetic luciferase-Based logic device. *ACS Sens.* 2020;5(3):807-13.



# **4. RESULTS DISCUSSION**



The research carried out in this doctoral thesis had the main objective of developing versatile biorecognition strategies for identifying clinically relevant genetic variants. Summary table compares the main characteristics and properties of the different methods developed. The results have demonstrated that the new DNA discrimination and amplification techniques show excellent selectivity. Also, the developed methods from the combination of innovative supports, assay formats, and transduction modes, are adaptable to many clinical environments. These systems had in common a very high sensitivity even for low-abundant SNV, showed good cost-effectiveness, and allowed the semi-quantitative or quantitative identification of the variants in complex clinical samples. The specific features of each method are summarized and discussed, emphasizing the advantages below.

- Compared to the current methods:

The **asymmetric blocked PCR amplification** strategy (chapter 1), EAB-qPCR using real-time fluorescent detection, was developed. Compared to conventional blocked qPCR (cb-qPCR), EAB-qPCR favored the formation of the blocking-native against primer-native hybrid by controlling the thermocycling (incorporation of a 65 °C stage) and the reaction conditions (asymmetric concentration of primers and addition of a blocking oligonucleotide). Thus, the amplification of native alleles was specifically reduced ( $\approx 8$  cycles of delay, blocking percentage  $> 99.7\%$ ), and minority alleles were enriched. The EAB-qPCR method demonstrated excellent sensitivity (0.5% mutant DNA/total) and reproducibility (RSD  $< 2.8\%$ ), as well as higher genotyping capacity and a 3-fold lower LOD than cb-qPCR. Moreover, it was successfully applied to the mutational analysis of metastatic colorectal carcinoma from biopsied tissues, providing results in agreement with two reference methods (cb-qPCR and NGS) and higher accuracy than cb-qPCR.

Considering that PCR-based methods are routinely employed as standard diagnostic tests in most hospitals due to their robustness, affordability, and general reagents, EAB-qPCR is easy to implement as it does not require substantial additional requirements when using the same instrument (standard qPCR equipment). Only the thermal cycle step must be regulated, which results in a slightly longer duration (1 min per cycle) than the other qPCR variants. The proposed strategy improves the common drawback of blocked PCR approaches by ensuring effective primer/blocker competition. It uses only 3 oligonucleotides without functionalization (2 primers and a ddC-type blocking agent), enabling the assay cost to be lower (about 2.5 € per assay) than approaches using expensive modified oligonucleotides as blocking agents (i.e., PNA, LNA). On the other hand, the oligonucleotide design and optimization process are more accessible to implement than COLD-PCR or ARMS-PCR approaches. In our strategy,

the use of SYBR Green as detection dye makes EAB-qPCR universal and straightforward for the detection of monobasic mutations.

For all these reasons, EAB-qPCR proved to be an accurate and cost-effective approach for extensive use in clinical laboratories. Although to decrease the energy requirement, price and improve portability, other systems were developed in the following chapters of the thesis.

**Ligation-uRPA method** (chapter 2) allows us the multiplex SNV detection, achieving for the first time the combination of ligation with RPA amplification in a universal format. Allele-specific ligation proved to be a very effective strategy based on probes and the high selectivity of the ligase. Efficient discrimination of multiple polymorphisms (six simultaneous probe pairs) in the *VKORK* and *CIP2C9\** genes even with low amounts of target DNA (1000 copies per assay) in a short time (first stage of 35 min for probe hybridization; and the second stage of 20 min for ligation) was achieved. The ligation products were amplified by RPA, increasing their multiplex capacity due to the use, for the first time, of universal primers, a single pair of shared primers for all the targets.

The method applied for POC detection was based on Blu-Ray disc technology. It requires fewer thermal variations and less precise heating/cooling technologies than strategy 1. Its characteristics favor the combination with portable instruments concerning other published PCR-mediated ligation methods. To avoid the heating/cooling cycles of the previous chapters, fully isothermal methods were used in the following approaches.

**RPA-AS-HCR method** (chapter 3) is a rapid amplification (RPA) and allele-specific hybridization chain reaction (AS-HCR) in the array format for SNV genotyping. On the one hand, it is the first study in which HCR is applied to work directly with genomic DNA by rapid sample processing consisting of a short pre-step of amplification by RPA. On the other hand, in addition to detecting the change as in previous studies in the literature, the identification of the nucleotide change in the gDNA was also possible thanks to its integration with the microarray. Its multiplexing capacity was high because the linker triggered the HCR reaction, and a universal pair of oligonucleotides (H1 and H2) formed the superstructures of nicked dsDNA.

The assay using a smartphone as detector was demonstrated useful for *KRAS* and *NRAS* in human cell culture and cancer biopsy tissue samples. An amplification factor of 8.8-fold higher than linear hybridization was obtained, with a LOD of  $10^3$  copies and a percentage of mutant DNA of 0.7% (1.1% for the PCR-based method), and good reproducibility (RSD= 5-17%). Smartphone detection allowed the assignment of the individual molecular profile at low cost and without the need for sophisticated instruments compatible within resource-limited settings.

**RPA-BRET method** (chapter 4) was developed in order to simplify operational requirements, decrease assay time and achieve direct detection of results. A new bioluminescent reagent was synthesized, enabling BRET transduction. It was the product conjugation between a mutant NanoLuc, a very bright luciferase, and thiazole Orange, a dsDNA intercalating dye. A specific BRET effect was achieved, as the dsDNA was distinguished from the ssDNA in the approach when the thiazole orange molecules were added around the mutated cysteine residue of the luciferase. The LOD was 10 nM (approximately  $10^{10}$  copies) without the need for an excitation source. The assay achieved good reproducibility (RSD= 1.6-3%) and high thermal stability, as it retained activity for at least 30 min below 50 °C. This reagent bound nonspecifically to dsDNA, so it had the advantage of universally recognizing DNA biomarkers.

With this reagent, it was possible for the first time to monitor RPA-amplified products of the KRAS gene (codon 12-13) from gDNA by measuring the BRET effect in endpoint and real-time modes. The RPA-BRET method presented a detection limit of 26 copies of gDNA in short times (30 min). Although with both modes, the sensitivity was similar to that obtained by fluorescence monitoring with SG RPA, the light emitted by the newly developed system was 66 times more intense than that of the traditional dye. The real-time RPA-BRET method also outperforms current state-of-the-art isothermal amplification monitoring by bioluminescence (e.g., LAMP-BART, which requires a thermostable luciferase, a complex primer design, and a working temperature of 60-65 °C). The new reagent was valid for all amplification (PCR or isothermal) in endpoint mode and real-time isothermal techniques with low working temperatures (below 50 °C). Therefore, it proved to be a simple and portable technology for molecular diagnostics near the point of need as it required only a thermoblock for constant temperature heating and a photodetector at 533 nm.

- Compared with each other:

According to the **amplification reaction**, one method is based on conventional PCR (chapter 1) and three methods are based on isothermal amplification (chapters 2-4). An essential drawback of PCR-based methods was that a high percentage of the assay time is due to temperature cycling and no amplification, which slows down the assay time (< 1 h). The use of RPA as an amplification technique was successful in raising sensitivity and producing small products for subsequent biorecognition. RPA was an excellent choice due to its high amplification yield in only 10 to 40 min from gDNA samples. Because it operates at very low temperatures close to ambient (37 °C) and its tremendous versatility, it was successfully integrated into multianalyte analysis systems using different assay formats and detection modes. In addition, the oligonucleotide

design and format did not require substantial changes to those of PCR (a single pair of short primers per target).

Regarding the proposed **detection system**, each strategy has its own particularities. The assay format based on solid-phase hybridization on microarrays of the products obtained with oligonucleotide probes was used for chapters 2 and 3. The main advantage was that several SNV were detected simultaneously, and their multiplexing capacity was very high. In contrast, in chapters 1 and 4, hybridization occurred in the reaction solution. In all strategies, selective hybridization to the target was achieved by effective design of the oligonucleotides (primers and probes) and by controlling the reaction conditions (temperature and buffer composition by adjusting the ionic strength or adding destabilizing agents).

On the one hand, the amplified fragments were detected by low-cost indirect colorimetry in chapters 2 and 3. In both systems, an easy immunoassay was chosen to recognize digoxigenin-labeled products, which has a simpler manufacturing process than other haptens, such as biotin, and presented adequate capabilities to recognize immobilized hybridization products. Both assay platforms allowed the effective immobilization of biotinylated DNA probes via affinity to streptavidin anchored to the support by physisorption ( $\approx 0.5\text{-}1 \text{ fmol/mm}^2$ ). In addition, they provided signals whose intensities were related to the target concentration.

In chapter 2, the use of BD technology in both the support (Blu-ray discs) and the detector (disc reader) proved to be an analytical tool of enormous potential. They were inexpensive, portable, sensitive, and rugged devices manufactured under high quality and mass production standards. Due to their large surface area, low non-specific absorption, and proper sensitivity and reproducibility, they were a competitive alternative to traditional microarray surfaces. In addition, they could read microarrays with smaller dots due to their higher resolution capacity, showing higher performance than other CD and DVD optical discs.

In chapter 3, PC chips and consumer electronics were applied for POC detection, reducing the costs. While the cost of a fluorescence microarray scanner can reach from 40,000 to 100,000 €, the price of a mid-range studied in this thesis was in the hundreds range. The advantages of using the smartphone were its widespread presence, portability, and ability to transmit data with an easy-to-use interface as a biomedical reader.

On the other hand, in chapters 1 and 4, the detection was either direct fluorescence (strategy 1) or BRET-based (strategy 4). Both had a high working capacity, and their main advantage was that they allow tracking and quantifying the amplification reactions in real-time, monitoring their kinetics. In addition, amplification and detection

were achieved in a single step. For qPCR, a real-time thermocycler and fluorescent readout were required. Whereas RPA detected by the NanoLuc/thiazol orange reagent required a heater and an emission reader. Thus, the requirement for an external excitation source was eliminated, avoiding potential photobleaching, toxic radical generation, and high background signal caused by autofluorescence and scattering. The analytical performances achieved with RPA-BRET were comparable to those of EAB-qPCR; although it has not been applied to SNV detection, the assay time and energy requirement were decreased.

The results obtained for the systems developed in this thesis demonstrate that the combination of discrimination and amplification strategies with optical detection was a promising tool to analyze genetic variants for molecular diagnostics, being a highly competitive alternative to reference techniques.

This thesis has also opened a path for future improvements in DNA variant genotyping by amplification and POC detection. More outstanding multiplexing capabilities and throughput could be achieved by optimizing multiple oligonucleotides (primers and probes) *in silico* and *in vitro*. In addition, the level of assay integration can be improved in the developed systems by combining two or more steps, such as ligation and amplification or amplification and hybridization. Assay platforms can also be improved using microfluidics, and other 3D printed supports to perform the assays, enhancing miniaturization and reproducibility promoting process automation.

**Summary table.** Comparison between the strategies developed in the thesis.

Strategy	1: EAB-qPCR	2: Ligation-uRPA	3: RPA-AS-HCR	4: RPA-BRET
<b>Discriminatory reaction</b>	Blocker oligo	Ligase	Allele specific HCR	-
<b>Amplification reaction</b>	PCR	uRPA	RPA	RPA
<b>Labelled oligonucleotides</b>	2',3'-dideoxycytidine	digoxigenin dUTPs	Digoxigenin	Nonmodified
<b>Detection</b>	Fluorescence	Colorimetric	Colorimetric	BRET
<b>Developer system</b>	SYBR-Green I	Immunostaining by Alkalyne phosphatase	Immunostaining by Au/Ag nanoparticle	NanoLuc-Thiazole Orange conjugate
<b>Detector</b>	Real-time thermal cycler	Blu-ray reader	Smartphone	Photodetector
<b>Assay monitoring</b>	Real-time	Endpoint	Endpoint	Real-time Endpoint
<b>Sample volume (µL)</b>	1	5	1	5
<b>Reaction volume (µL)</b>	12.5	20	12.5	60
<b>Support platform</b>	PP microplate	Blu-ray disc	PC chip	PP microplate
<b>Assay format</b>	Solution	Microarray	Microarray	Solution
<b>Assay steps</b>	1	3	2	1
<b>Assay time (min)</b>	80	200	60	30
<b>Assay temperature (°C)</b>	40 cycles of 95-65-55	98-65-54-98 (L) 37 (A) 92-37 (H)	37	37
<b>Multiplex gene</b>	1	3	2	1
<b>Throughput</b>	96	36	12-100	96
<b>LOD (%mutant DNA)</b>	0.5%	-	0.7%	-
<b>Sensitivity (template copies)</b>	20	1000	1000	26
<b>RSD (%)</b>	2.2- 2.8	6.0-19	5.0-17	1.6- 3
<b>Cost</b>	Moderate	Low	Low	Low
<b>Portability</b>	Low	Medium	High	High
<b>Energy requirement</b>	High	Low	Low	Low
<b>Target gene</b>	<i>KRAS</i>	CYP2C9*2 CYP2C9*3 <i>VKROC1</i>	<i>KRAS</i> <i>NRAS</i>	<i>KRAS</i>
<b>Clinical samples</b>	Human cell cultures and FFPE biopsy tissues	Buccal sample	Human cell cultures and FFPE biopsy tissues wild-type <i>KRAS</i> codon 12-13 c.38G>A c.34G>T	Blood sample
<b>SNV identified</b>	wild-type <i>KRAS</i> codon 12-13	g.8633C>T g.47639A>C g.3588C>T	wild-type <i>NRAS</i> codon 61 c.181C>A c.182A>G	-

PC: polycarbonate; PP: polypropylene; L: ligation; A: amplification; H: hybridization FFPE: Formalin-fixed paraffin-embedded



# **5. GENERAL CONCLUSIONS**



In the present thesis, new strategies for biorecognition and optical detection of clinically relevant biomarkers for mutational analysis have been developed to facilitate the implementation of personalized medicine in any field. The method performances have been evaluated according to the specific problem and application, achieving the initial objectives. The experimental results have been presented throughout this work support their advantages compared to the current technologies. The most relevant conclusions are summarized below.

**1.** Specific discrimination and detection of target DNA sequences have been achieved, even in cases where the difference between the mutant variant and the wild-type was only a one-base-pair mismatch. For this purpose, different recognition events have been integrated into versatile analytical platforms using simple optical transduction modes:

- The new variant of the blocked PCR, EAB-qPCR, minimized amplification of native DNA variants allowing better selective enrichment of minority alleles by real-time fluorescence.
- Multiple SNVs can be identified by combining allele-specific ligation with uRPA in multiple hybridization assays in microarray format and colorimetric detection using Blu-ray compact disc technology.
- Ultrasensitive and selective multiplex identification of SNV has been achieved by combining RPA with AS-HCR in a microarray format and colorimetric detection using a smartphone.
- The design and synthesis of a bioluminescent reagent enable real-time and end-to-end monitorization of RPA products based on the BRET effect.

**2.** The strategies developed have been successfully applied to the genotyping of variants involved in treating colon cancer and cardiovascular disorders from genomic DNA and clinical samples at femtomolar scale concentrations, using simple genotype calling criteria for patient classification. One of the achieved challenges has been to improve the ability to discriminate the mutant from the native fraction in solid tumor biopsy samples, even those with low percentages of tumor cells or despite their preservation method.

**3.** The comparison with conventional polymerase chain reaction (PCR)-based tests shows the excellent analytical performances of developed methods, such as higher selectivity against single-base mismatch sequences and high reproducibility. In addition, the assays were validated to be 100% concordant with the reference sequencing method, regardless of mutation type and position. Considering state of the art, all our

biorecognition approaches can be classified as a high-moderate sensitivity method for mutational analyses (0.1-1%) useful for solid tissues, second only to ddPCR and ice-COLD-PCR technologies (0.001-0.1%).

4. Compared to the current techniques in clinical laboratories, the proposed strategies are cheaper, simpler, and more robust alternatives. Although efforts must be made to commercialize the developed systems, they can all be implemented in different settings covering any environment. Due to its efficiency and affordability, EAB-qPCR can become a powerful tool to support the diagnosis of patients in the daily routine in hospitals. Ligation-uRPA, which does not require specialized personnel or complex equipment, is very useful to support clinical practice in decentralized laboratories. The last two strategies, AS-HCR and RPA-BRET, based on portable, low-power, and easy to handle systems, are especially suitable for molecular profiling in onsite detection with limited laboratory infrastructure.

5. Since the knowledge generated is based on universal or versatile formats, it is easily extrapolated to other genome screening applications. This research included analysis of the requirements of each method to be applied in other clinical situations, ensuring an effective discrimination, amplification, and hybridization of the desired variations.

These achieved scientific advances are promising for facilitating access to a sustainable universal personalized medicine in routine clinical practice.

