XXXIII Cycle PhD course in Science and Technologies of Chemistry and Materials

Curriculum: Drug Discovery & Nanobiotechnologies

Thesis title:

Point-of-care diagnostics for Single Nucleotide Polymorphisms genotyping: applications to food traceability, nutrigenetics and pharmacogenetics

PhD student: Giulia Cibecchini

Supervisors: Dr. Pier Paolo Pompa (IIT)/ Dr.ssa Paola Cecere (IIT), Prof.ssa Renata Riva (UniGe)

.

Abstract

Single Nucleotide Polymorphisms (SNPs) are common genetic variations associated to specific phenotypes. SNPs have high relevance in different biomedical areas including both preventive/personalized medicine and non-clinical contexts, such food traceability. However, current techniques for SNP genotyping are expensive, require specialized laboratories and costly instrumentations, and are time consuming. In this framework, the development of rapid, point-of-care (POC) SNP diagnostic technologies would be of great interest, opening new possibilities for early, on-site screenings in multiple fields. In this PhD thesis, starting from state-of-art approaches, a rapid and portable diagnostic strategy based on Loop Mediated Isothermal Amplification (LAMP) has been developed for SNPs genotyping. In particular, the developed strategy was applied and validated, in a naked-eye colorimetric approach, for food varietal discrimination (i.e. durum wheat variety) in a model case of industrial relevance. Moreover, the POC technology was targeted to different salivary biomarkers, to address nutrigenetics and pharmacogenetics issues. Notably, the possibility to rapidly diagnose the genetic predisposition to lactose intolerance or the impairment of folic acid metabolism was proven. Furthermore, the developed POC-compatible protocols can work with portable, battery-powered devices, allowing for POC operation of such diagnostic tests.

Contents

Chapter 1

Introduction

1.1 In vitro diagnostics and Point-of-Care testing

In Vitro Diagnostics (IVD) includes instruments, techniques and reagents that allow to identify and/or monitor a specific target or a condition.

Over the years, the concept of In-Vitro Diagnostics has evolved, requiring new tools and strategies to satisfy the current need to provide fast, on-site and cost-effective diagnostic responses, ensuring at the same time high performance in terms of sensitivity, specificity and reliability.

In clinical practice, the shift from a curative to a preventive health system has led to an increasing demand for early diagnosis and personalized medicine. Moreover, among the others, the development of new genetic tests is strengthening the concept of predictive medicine, that allows to evaluate predisposition and susceptibility of the patient to a certain physiological or pathological condition, facilitating targeted treatments.

New diagnostics tools are advantageous also in other fields, such as in quality control, in industrial and agri-food sectors. In fact, nowadays, government regulations and people consciousness about food quality standards and environmental issues, are calling for increasingly stringent quality controls, with frequent and massive checks along the entire production chain.

Overall, this situation has resulted in a huge expansion of *In Vitro* Diagnostics (IVD) market with a constantly growing future trend [1]. Indeed, considerable efforts are being doing by scientists and researchers to develop new pioneering technologies for diagnostics, focusing in particular on on-site and Point-of-Care (POC) diagnostics, that allows testing at the point of need, in centralized laboratories as in decentralized settings, such as pharmacy, the physician's office or even at home, raising patient compliance and promoting rapid-testing,

or directly at food processing step in industrial contexts. Traditional diagnostics requires multiple and time-consuming processes, involving expensive instrumentation and qualified personnel. On the contrary, on-field diagnostics is fast, and it can be performed in portable devices, providing reliable and sensitive responses. Furthermore, POC tests speed up the turnaround time of the analysis and as a consequence, the decision time [2]. This is a key point not only in medicine, but also for example for food supply chain because the perishable nature of food and the food market related economy raise the need for rapid tests that allow quality control checks and a prompt non-compliant batch withdrawals.

Among all the emerging POC tests, cellphone (CP)-based technologies provide measurement of biological markers or environmental contaminants with a portable smartphone-like device [3, 4, 5]. Two examples of commercially available and commonly used POC tests are pregnancy test and glucose meters [6, 7]. These systems provide an incredible advantage to the home-based monitoring of a therapy enabling a clear and unique interpretation of results and the possibility to transmit clinical data.

Among all the wide range of analysis that can be performed through *In Vitro* Diagnostics, there are genetic tests. Genetic tests are almost exclusively performed in specialized laboratories, and they require a long turnaround time and high costs. For this reason, the availability of a genetic pre-screening system in decentralized settings would be very innovative.

Indeed, POC for Nucleic Acid Testing (NAT) are emerging among POC tests. POC-NATs allow to identify the DNA or RNA of a specific target. POC-NATs are applied in several fields enabling, for example, the detection of bacterial and, viral infections, allowing an easy management of a therapy or even, in industrial field, facilitating the monitoring of potential adulteration of end products [8]. Two examples of POC-NAT complete technologies that present the possibility to automate the entire process are the GeneXpert and the IQuum LIAT systems that have been studied for the diagnosis of tuberculous pericarditis and HIV infection, respectively [9, 10].

At the basis of a POC-NAT test usually there is a nucleic acid amplification step. The gold-standard amplification technique is the PCR (Polymerase Chain Reaction). However, since PCR requires thermal cycles, it is commonly performed with a specific instrument, the thermalcycler. Moreover, PCR is susceptible to inhibitors, hence it requires an elaborate

extractive phase to obtain an extremely purified nucleic acid target. For these reasons, the development of POC-NAT technologies based on PCR requires the design of a complex bench top technology able to perform DNA extraction, DNA amplification and detection.

An alternative to the use of PCR-based methods is the employment of isothermal amplification reactions, where the amplification of the nucleic acid target occurs at a single reaction temperature. In this way, such amplification reactions can be performed in a simple heater device reducing costs and complexity. For all these features, the use of an isothermal amplification method appears the strategic choice for the development of a POC platform.

1.2 Isothermal amplification techniques

Polymerase chain reaction (PCR) is the gold standard strategy for nucleic acid amplification in molecular biology [11]. This technique operates the amplification through thermal cycles and consists of three stages: denaturation, annealing and extension. The first step, typically at 94-95 °C, is the heat denaturation step that determines the opening of the double strand of the DNA target molecule and prepares the target for the subsequent stages; in the second step the primer anneals to the complementary region on the single-strand target; during the third step, the primer is elongated and a new molecule of DNA is produced. All these stages are repeated one after the other leading to an exponential amplification that permits to amplify up to 10^6 copies of the starting template.

Since the thermal cycles can be operated only through an expensive instrument, the thermal-cycler, the use of the PCR is not convenient for a POC test. On the contrary, isothermal amplifications represent the most suitable techniques for portable applications. These strategies permit to exponentially amplify nucleic acids at one unique temperature bypassing more complex procedures. Several isothermal amplification methods for both RNA and DNA templates have been proposed in literature (Figure 1) [12].

Nucleic Acid Sequence-Based Amplification (NASBA) and Transcription Mediated Amplification (TMA) are both examples of techniques for isothermal amplification of RNA. These approaches exploit the function of a Reverse Transcriptase enzyme which synthetize a double stranded DNA (cDNA) starting from the target RNA. The cDNA produced represents then the starting point for the later amplification of the RNA product by a DNA dependent-

RNA polymerase. The starting primer includes both a sequence complementary to the target RNA and a t7 promotor sequence. Together the activity of an RNase and two primers generate a double stranded cDNA containing the site for initiating RNA synthesis [13, 14, 15].

Another method, named Signal-Mediated Amplification of RNA Technology (SMART), works with the generation of a three ways junction structure thanks to the special design of primers. The specific design of the probe forces the direction of the DNA synthesis towards the strand, which is perpendicular to the target sequence. The newly formed double stranded DNA which contains a t7 promotor sequence enables the exponential RNA synthesis [16].

Furthermore, many others isothermal amplification methods for DNA amplification are reported in literature, such as Strand Displacement Amplification (SDA), Helicase-dependent amplification (HDA), Rolling circle amplification (RCA) and Recombinase polymerase amplification (RPA). For example, SDA method is based on the use of a restriction enzyme and a 5'-3' exonuclease deficient DNA polymerase with strand displacement activity. In the first proposal of the strategy two primers containing the binding site for a restriction enzyme were used. The 3' end of the template is then extended using the attached primer as template. Nucleotides with a thiophosphate modification are used for DNA synthesis, thus the restriction enzyme is able to cut only the unprotected strand of the DNA. The enzyme nicks the unmodified strand, the DNA polymerase, lacking 5'-3'exonuclease activity, repairs the gap and displaces the downstream strand. Cleavage, extension and displacement are repeated several time triggering isothermal exponential amplification. In the first proposal of SDA method the initial steps require both a denaturation step to obtain a single strand nucleic acid and a pre-cleavage of the DNA sample operated by a restriction enzyme in order to generate a DNA fragment with determined 5' and 3' ends [17]. This step was overcome by the addition of two displacement primers that generate a fragment with two restriction sites at both ends. This optimization permits to obtain a template immediately available for following stages of the amplification [18].

Rolling cycling amplification (RCA) is another strategy that exploits the ability of DNA polymerase to utilize circular templates. This method is based on the use of a special probe, which, after its hybridization to the target sequence, causes the closeness of the ends, circularizing the target nucleic acid. Subsequently a DNA ligase binds both ends and completes the circularization. Upon the circularization of the template, a second primer hybridizes the

sequence. Then, it is extended by a DNA polymerase, producing a long repeating single stranded sequence [19].

However, all the methods described above require a starting heat denaturation step, while later stages are performed at one unique temperature, thanks to the strand displacement activity of the polymerase or, in case of RCA, to the generation of a linear sequence product.

Another approach described in literature is Helicase-dependent amplification (HDA). In this method, double stranded DNA is uncoiled by a DNA helicase. Then two primers anneal to the opened double stranded and a DNA polymerase extends the primers. Two molecules of double stranded DNA that are new substrates for DNA helicase are generated, fueling the cycle. This technique provide 10^6 fold amplification in 3 hours and the entire process is performed at one temperature not requiring a pre-heating step [20].

Recombinase Polymerase Amplification (RPA) is a strategy that exploit the activity of a Recombinase. The complex Recombinase-primer scans the target sequence and binds to the site complementary to the primer. The primer is extended by a DNA polymerase and the new generated DNA strands acts as starting point for the exponential amplification. Also in this case the use of a polymerase with strand displacement activity is required and the entire reaction is performed at one unique temperature [21].

Nowadays, among all the isothermal amplification techniques, Loop-mediated isothermal amplification (LAMP) is emerging and attracting great interest due to its rapidity, its high amplification efficiency and its large tolerance to interfering substances.

Table 2 Summary of isothermal amplification/detection the platforms

Platform	Advantage	Disadvantage	Components	Application
NASBA	10^7 fold amplification in 2 h, commercial kits available	Not ideal for DNA analyte	2 primers, reverse transcriptase, RNase H, RNA polymerase, dNTP, rNTP	HIV-1 genomic RNA,[34] hepatitis C virus RNA,[35] human cytomegalovirus mRNA,[36] 16S RNA in many bacterial species,[37] and enterovirus genomic RNA, [38] *etc.*
SDA	10^7 fold amplification in 2 h	Less efficient on long targets, sample preparation is required	4 primers, DNA polymerase, REase HincII, dGTP, dCTP, dTTP, dATPαS	Mycobacterium tuberculosis genomic DNA[81]
LAMP	~ 10^9 amplification in less than 1 h in the presence of interfering sequences	Complicated primer design, not suitable for short analytes	4 primers, DNA polymerase, dNTP	Viral pathogens, including dengue,[47] Japanese encephalitis,[48] Chikungunya,[49] West Nile,[50] *Severe acute respiratory syndrome* (SARS),[51] and highly pathogenic avian influenza (HPAI) H5N1, [52] *etc.*
RCA	ssDNA product makes this compatible with other isothermal amplification strategies	Ligation using RNA analyte could be problematic	1 probe, ligase, polymerase and dNTP	Cystic fibrosis gene, Epstein–Barr virus in,[70] influenza A H1N1 and H3N2 mutations,[71] porcine circovirus type 2,[72] and Listeria monocytogenes,[73] *etc.*
SMART	50 pM analyte detected in ~ 4 h, tolerates crude sample	Two-step process	2 probes, DNA polymerase, RNA polymerase, dNTP, rNTP	DNA and tRNA of *E. coli*,[79] DNA and mRNA of marine cyanophage virus[80]
HDA	10^6 fold amplification in 3 h, commercial kits available	Complicated buffer optimization	2 primers, helicase, SSB, DNA polymerase, dNTP	*Treponema denticola* gene,[81] *Brugia malayi* gene[82]
RPA	100 copies of analyte were detected in 1 h, selective	Stringent reaction condition	2 primers, recombinase, SSB, DNA polymerase, REAse Nfo, dNTP	*Bacillus subtilis* genome[84]

Figure 1. Isothermal amplification techniques [12].

1.3 Loop-mediated isothermal amplification (LAMP)

In 2000, Notomi and co-workers proposed a new isothermal amplification strategy, the Loop-mediated isothermal amplification (LAMP) method. This approach revealed to be efficient and fast presenting a lot of advantages if compared to others strategies [22]. With LAMP technique DNA is amplified 10^9-10^{10} times in 15-60 minutes.

The amplification proceeds at a constant temperature exploiting a DNA polymerase with a high strand-displacement activity and four specially designed primers, complementary to six distinct regions on the target sequence. In PCR-based technique, exponential amplification is operated by thermal cycles, while in LAMP it is obtained thanks to the strand displacement activity of DNA polymerase together with the formation of single stranded loop, in the stem-loop structures. A stem-loop structure or hairpin is a secondary structure that can occur in single-stranded DNA or RNA. This kind of structure is formed when two regions of the same strand are one identical and one complementary to another region of the same strand. The annealing of these complementary regions enable the formation of a double strand portion (stem) and a single strand structure (loop). To form a stem-loop structure, the length and the composition of the strand are fundamental. Moreover, this structure is formed if it has enough thermodynamic stability. In particular, between the two complementary regions, a sufficient number of nucleotides is required to obtain the folding and generate a stable loop structure. In addition, also the composition of the complementary portions influences the stability, affecting the thermodynamic balance between the linear and the stem-loop structure.

In general, four primers are involved: forward and backward inner primers (FIP and BIP), and forward and backward outer primers (F3 and B3), designed on six regions of the target sequence: F3c, F2c e Fc1 at 3’ end, and B1, B2 e B3 at 5’ end.

Forward Inner Primer (FIP) consists of the F2 region (at the 3' end) that is complementary to the F2c region, and the same sequence as the F1c region at the 5' end. Backward Inner Primer (BIP) is designed in the same way on the complementary strand.

Forward Outer Primer consists of the F3 region that is complementary to the F3c region. Backward Outer Primer is designed in the same way on the complementary strand (Figure 2).

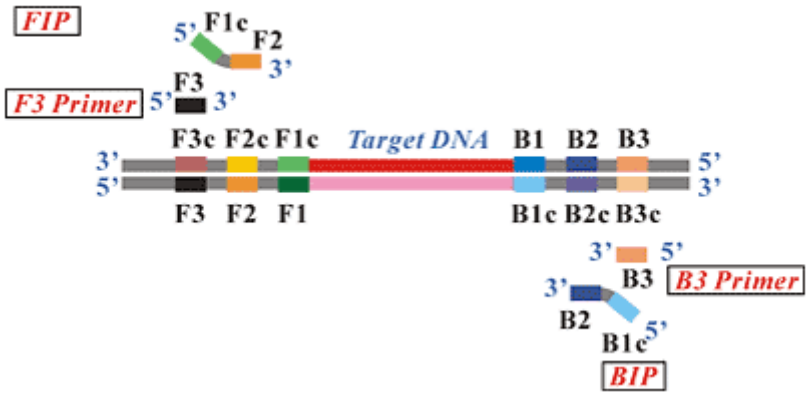

Figure 2. Schematic representation of four primers for LAMP reaction and their corresponding six binding regions on the target sequence [23].

At the temperature around 60-65°C the double stranded DNA is in the condition of dynamic equilibrium. Indeed, one of the LAMP primers can anneal to its complimentary sequence on the target gene, without requiring an initial heath denaturation step at higher temperature.

The reaction starts with the hybridization of F2 portion of FIP with F2c region in the target sequence (Figure 3).

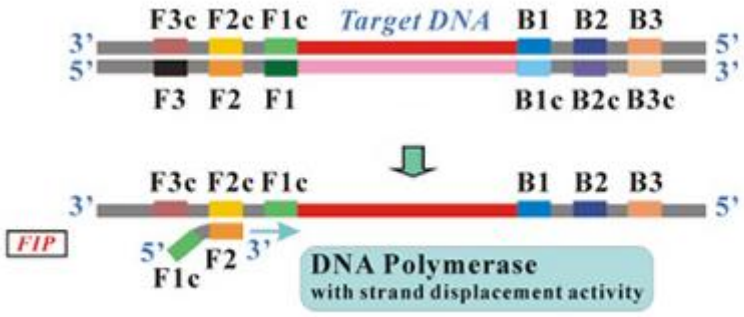

Figure 3.Step 1 of LAMP reaction [23].

DNA polymerase extends the 3' end of FIP synthesizing a DNA strand complementary to the template (Figure 4); meantime, the outer primer F3 binds F3c region of the target sequence (Figure 5).

Figure 4. Step 2 of LAMP reaction [23].

Figure 5. Step 3 of LAMP reaction [23].

F3c region is located upstream to F2c, therefore F3 extension by DNA polymerase with strand displacement activity causes the release of the newly synthetize product (Figure 6 and 7).

Figure 6. Step 4 of LAMP reaction [23].

Figure 7. Step 5 of LAMP reaction [23].

As the first product is displaced, F1c portion of FIP at the 5'end of the newly formed product can self-hybridizes to F1 region, creating a stem-loop structure. This newly generated strand works as template for the BIP-mediated synthesis and subsequent B3-primed strand displacement DNA synthesis (Figure 8).

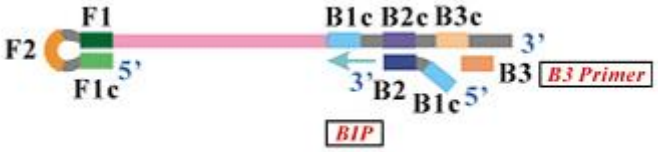

Figure 8. Step 6 of LAMP reaction [23].

BIP through its B2 portion binds to B2c of the strand and, starting from the 3' end of the BIP, the synthesis of complementary DNA takes place. (Figure 8). Through this process, the DNA is reverted into a linear strand.

Subsequently B3 anneals with B3c region, DNA polymerase synthetizes a new strand by the extension of B3 and displaces the product (Figure 8 and 9). The single strand product rapidly forms a structure with stem-loops at each end, named dumb-bell structure, thanks to the self-hybridization of F1 with F1c and B1 with B1c (Figure 10).

Figure 9. Step 7 of LAMP reaction [23].

Figure 10. Step 8 of LAMP reaction [23].

The stem-loop structure contains a single stranded DNA portion (the loop) and a double stranded one (the stem). This structure contains multiple sites for primers annealing and consequently leads to the exponential amplification stage. Starting from this structure, the amplification continues exponentially. The self-priming at the 3' end of F1 portion leads to the extension of the strand. Simultaneously, FIP anneals to the F2c portion of the stem-loop construct triggering the synthesis of a new strand (Figure 11a and 11b). Then, the strand displacement activity of the DNA polymerase causes the release of the strand synthesized by the extension of FIP (Figure 11d). This strand generates a dumb-bell structure with single strands loops at both ends (Figure 11d). The newly generated dumb-bell represents the complementary configuration of the starting structure. Then, B1 portion at the 3' end of the new product behaves as a primer triggering DNA synthesis and BIP anneals to the B2c portion of the stem-loop structures (Figure 11d and 11e). Rapidly, this strand is displaced and it forms a stem-loop structures by the annealing between the newly formed portions, B1 and B1c, and by the hybridization of F1 with F1c portions (Figure 11a). In this way, the amplification restarts [23].

Finally, in the elongation and recycling steps, FIP and BIP continue to hybridize to the template and the DNA polymerase continues to displace newly synthetized strands. The ending products are multiple stem-loop structures, called cauliflower-like structures, with alternately inverted duplicates of the starting template (Figure 11f). The effect of the amplification is the production of a high concentration of complex structures with different length.

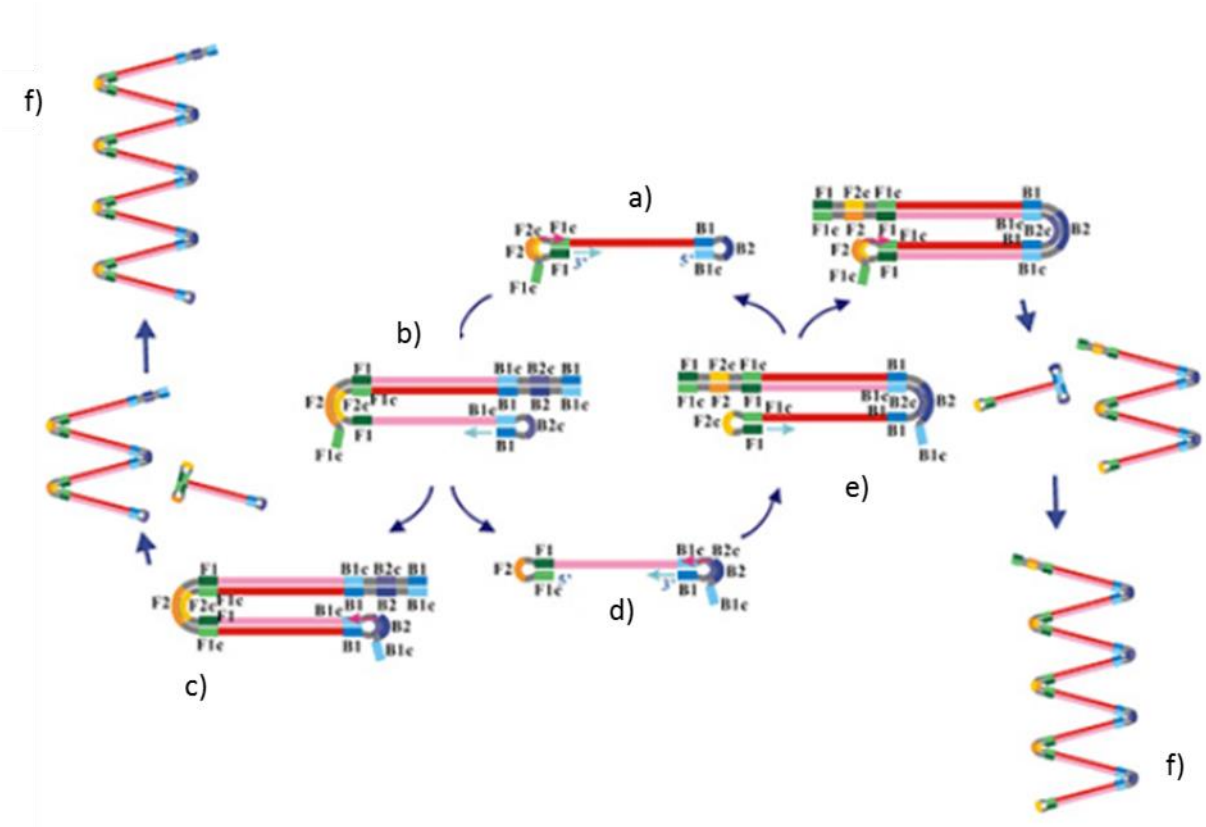

Figure 11. Cycling amplification steps [23].

The particular design of FIP and BIP primers has a pivotal role in the reaction mechanism. The elongation of inner primers permit the self-hybridization of the product and the formation of the dumb-bell structure.

For initial steps of LAMP reaction, all the four primers are used, while later phases requires only the action of FIP and BIP. For this reason, the needed amount of F3 and B3 primers inside the reaction mix is lower in comparison with inner primers concentration.

In the first proposal of the LAMP technique, the first step required was the heat denaturation of the template while the subsequently steps are under isothermal condition. Afterwards, it was observed that at the temperature between 60 and 65 °C and in presence of high concentration of betaine the double-stranded and the single stranded DNA configurations are in a thermodynamic balance [24]. This particular condition favors the annealing of inner primers and permits to avoid the initial heat denaturation step, allowing the entire process to be performed under isothermal condition.

In 2002 Nagamine et al. proposed a pivotal improvement of the LAMP, introducing the Loop primers, that provided a significant acceleration of the reaction time [25].

The Loop primers (either Loop primer B or Loop primer F), contain sequences complementary to the single stranded loop region, and they are designed between the B1c and B2c regions (loop B), or between the F1c and F2c regions (loop F) (Figure 12).

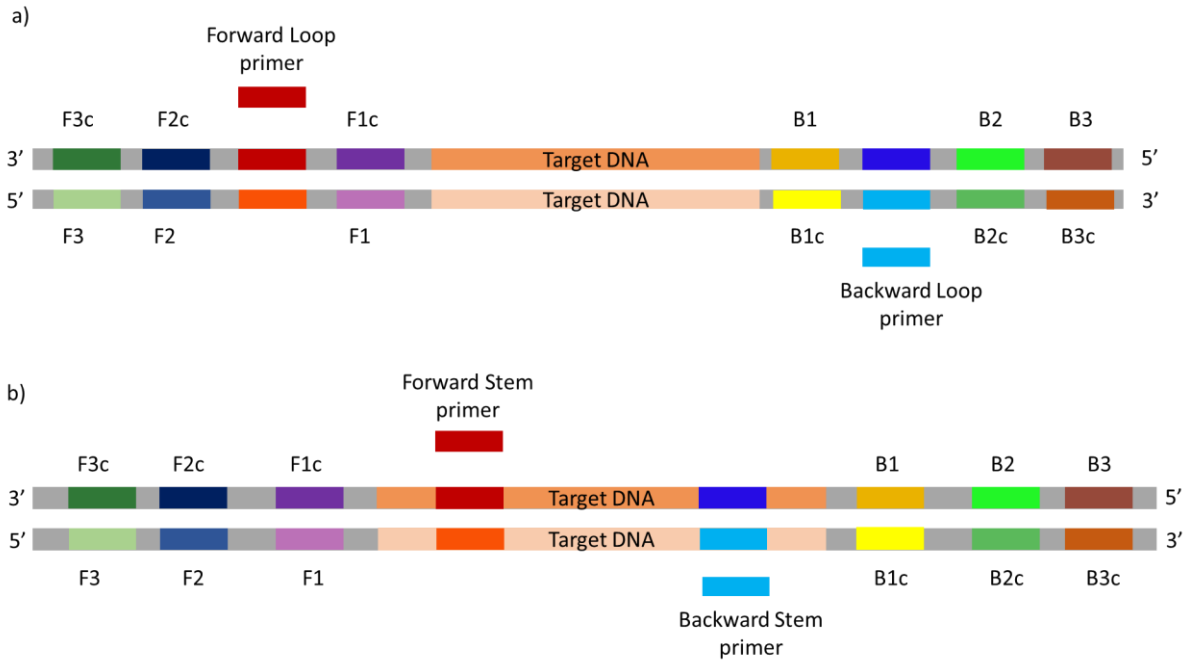

Figure 12. Schematic representation of the six primers involved in LAMP amplification, including loop primers introduced by Nagamine, and their corresponding eight binding regions on the target sequence.

Loop primers anneal to a portion of the dumb-bell structure that is not involved in the annealing with inner primers, providing an increased number of starting points for DNA synthesis for the LAMP method. (Figure 13).

These primers act in the latest phases of the reaction: the elongation and recycling steps. Since loop primers hybridize to the "available" loop regions of the dumbbell, they act together

with FIP and BIP. In this way, all the single strands formed during the amplification can be used as starting point for DNA synthesis.

Therefore, the main effect of the introduction of Loop primers is the increasing of the amplification rate and a significant lowering of the reaction time.

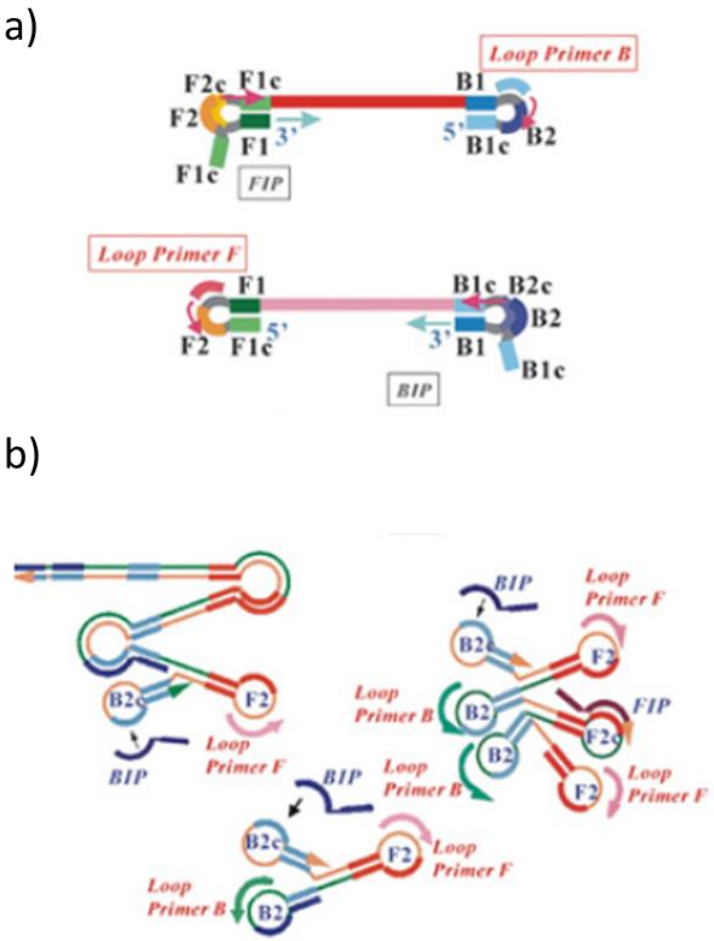

Figure 13. Schematic representation of the loop primers binding sites on the dumb-bell structure (a) and an example of a LAMP product containing multiple sites for Loop primers annealing (b) [23].

Detection of LAMP products can be achieved in Real-Time or through an end-point strategy. Real-Time detection is achieved using fluorescent DNA intercalating agents that are incorporated into the DNA structure [26]. SYBR green I, for example, binds the dsDNA molecules, and the resulting fluorescent emission obtained is directly proportional to the amount of amplicons produced.

The standard end-point detection is the analysis of LAMP products carried out by gel electrophoresis. During the gel preparation an intercalating agent, such as ethidium bromide or its safe alternative dye, the SYBR Safe, is added. The result of gel electrophoresis, visible through the UV light irradiance, is called “smear” because LAMP products migrate differently due to their different length and to their complex structure [27].

A second method is based on the visualization of a white turbidity in the reaction tube. This phenomenon derives from the precipitation of magnesium pyrophosphate formed by the interaction between pyrophosphate, which is formed during amplification, and magnesium ions, which are abundant in the reaction mix [28].

This by-product is produced in proportion to the concentration of amplified products, which is in a greater amount after LAMP amplification. This observation has opened the perspective for the possibility of visualize LAMP products directly in the reaction tubes allowing a naked-eye and instrument-free detection. Moreover, other substances able to improve the rapid visualization of the results can be added, for example fluorescent metal ion indicators, such as calcein, or colorimetric indicators such as Hydroxy naphthol blue (HNB) [29, 30].

A recent strategy for end-point LAMP detection is based on the use of pH-sensitive dye enabling the detection of LAMP products through a color change [31, 32]. During the DNA synthesis, a proton is released for every new phosphodiester bond formed (Figure 14).

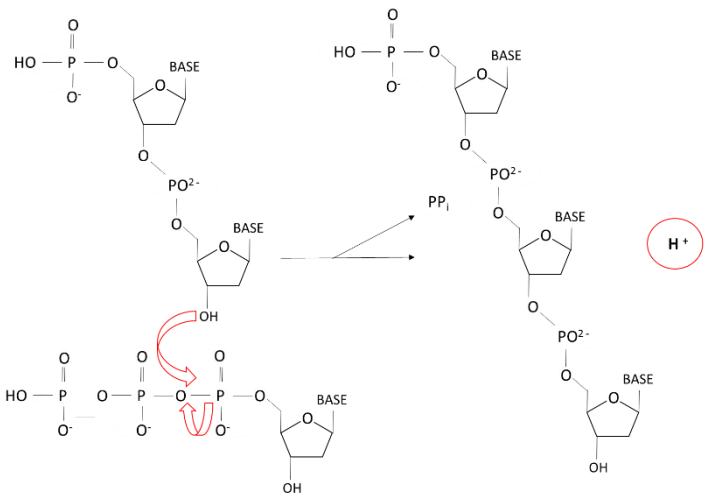

Figure 14. Formation of a phosphodiester bond during DNA synthesis. This biochemical reaction, catalyzed by DNA polymerase, leads to the release of one pyrophosphate molecule and one proton.

Unlike PCR, in the LAMP amplification, the high concentration of long and complex amplification products causes a drop in the solution pH of more or less 2 units, starting from the standard solution pH around 8.5.

This method can be used only for LAMP reactions. In fact, PCR leads to the formation of amplicons with the same length determined by the position of the primers. For this reason, there isn't a sufficient proton's release to achieve such a lowering of the pH. In LAMP, this pH change can be detected with an appropriate organic dye that has different colors for the acid form and the basic one, and that is able to give a visible solution color change in the reaction transition pH range.

PH-sensitive dyes are low-cost, well-functioning and can lead to a clear naked-eye visualization of the result. Moreover, pre-addition of dye inside the reaction mix permits to avoid the opening of the reaction tube that can cause cross-contaminations caused by reaction amplicons, highly probable due to high sensitivity of the LAMP reaction.

Among all these detection methods, the fluorescence-based detection is advantageous because it is fast. On the other hand, colorimetric detection methods make easier the interpretation of the results.

In addition, LAMP has high tolerance to contaminants allowing the employment of simplified nucleic acid extraction protocols into a POC platform [33]. All these features make LAMP a powerful technique ideal to develop a portable test.

1.4 Single Nucleotide Polymorphisms (SNPs)

In 2003, the whole human genome was sequenced by the Human Genome Project, which gave to the scientific community an huge heredity [34, 35, 36]. This project paved the way for characterizing most of the genetic variations distributed through the whole genome. Among these variations, Single Nucleotide Polymorphisms (SNPs) were discovered with a frequency of 1 bases every 1000 in the human genome [37]. This type of polymorphisms are defined as a substitution of a single nucleotide at a specific position in the genome, with a frequency of at least 1% in a random set of individuals in a population. SNPs can be located in non-coding

regions of the genome, without any impact on the phenotype of an individual, in coding regions or in a regulatory part of a gene, altering in these way the function of the encoded protein [38]. Moreover, it is well established that genetic variations are associated to phenotypical features that can affect individual responses to nutrients or drugs. The presence of the polymorphic version of a protein can influence the protein function and, if such a protein is involved in a particular metabolic pathway, it can modulate the response to a certain substance, such as a supplement or a pharmacological treatment [39, 40, 41]. Nowadays, since SNPs are strongly associated with well-known disease or pharmacological side effects, personalized medicine, based on the delineation of the individual and specific genetic profile of patients, is attracting attention.

Moreover, SNPs are the most abundant genetic variation of all the genomes. Hence, they represent biological markers for identification of foods varieties [42, 43]. In this framework, SNPs are exploited as markers for unequivocal authentication in food traceability field [44].

For these reasons, SNP detection provides several advantages for a broad range of applications. Several PCR-based techniques for SNP genotyping have been developed and described in literature [45, 46]. The simplest approach is the hybridization method that enables the allele specific discrimination with two single-stranded DNA molecules employed as probes. Only the perfect matched annealing is stable and provides a specific discrimination. The correct hybridization between the target DNA sample and the probe can be revealed by a fluorescence emission that is evidence of the detection of the correspondent allele type. A particular example are molecular beacon probes that are single-stranded oligonucleotides with a fluorophore and a quencher covalently bonded to the respective ends. In such a case when the probe is annealed to the target sequence the quencher is sufficiently separated to the dye in order to not inhibit the fluorescence emission [47, 48]. Hybridization-based mechanisms are suitable for solid supports, such as microarray technology, enabling simultaneous detection of multiple markers. Another method for allele discrimination is based on primer extension, in fact, only in case of a perfect match between the target sequence and the primer, the synthesis proceeds highlighting the presence of the related allele type [49]. Another method involves DNA ligase enzyme that due to its high specific activity can be used for discrimination. In this approach, two adjacent probes are linked together by the action of ligase enzyme only in case of a correct annealing [50]. Finally, the invasive cleavage

approach relies on the simultaneous annealing of two probes overlapping for a single nucleotide creating a structure that is recognized by an endonuclease. The cleaved product is the result of the correct detection [51].

Overall, the most accurate method for SNP genotyping is the genome sequencing that, albeit with a high cost, enable the reading of the entire target sequence [52]. The first sequencing method was developed by Fred Sanger in 1997. After being developed in 1977, it was first commercialized by Applied Biosystems in 1986. The method relies on the use of chain-terminating dideoxynucleotides [53].

In the last years, next generation sequencing (NGS) or high throughput technologies have been developed, allowing to sequence many fragments in parallel generating a high number of data. Since the introduction of the methodology by Sanger, sequencing technologies has made great strides significant reducing costs. Moreover, NGS technology has influenced the production of dozens of prokaryotic and eukaryotic genome assemblies, including model and crop plants. However, NGS requires very expensive technologies and a big effort for data analysis. Indeed, the main limits that preclude SNP genotyping to routine diagnosis are the high costs and the laboriousness nature of these techniques that can be performed exclusively by skilled users in specialized laboratories. Indeed, easier and portable tools for SNP identification would further widen the possibility of taking advantage of informativity of genetic variations.

1.5 SNPs detection with LAMP technique

LAMP strategies for SNP detection are generally based on the modification of the design of primers. In order to distinguish both the wild-type and the polymorphic alleles, two versions of the same primer are employed. The most common approach described in literature, proposes to design inner primers in order they hybridize to the polymorphic or the wild-type base with the 5' end of both FIP and BIP [54, 55] or with the second base at the 5' end of FIP and BIP [56]. Using the FIP complementary to the wild-type sequence, the dumbbell-like starting structure is formed and the DNA synthesis proceeds, while when the target gene is the mutant allele the DNA synthesis is blocked since no dumb-bell-like structure is generates (Figure 15).

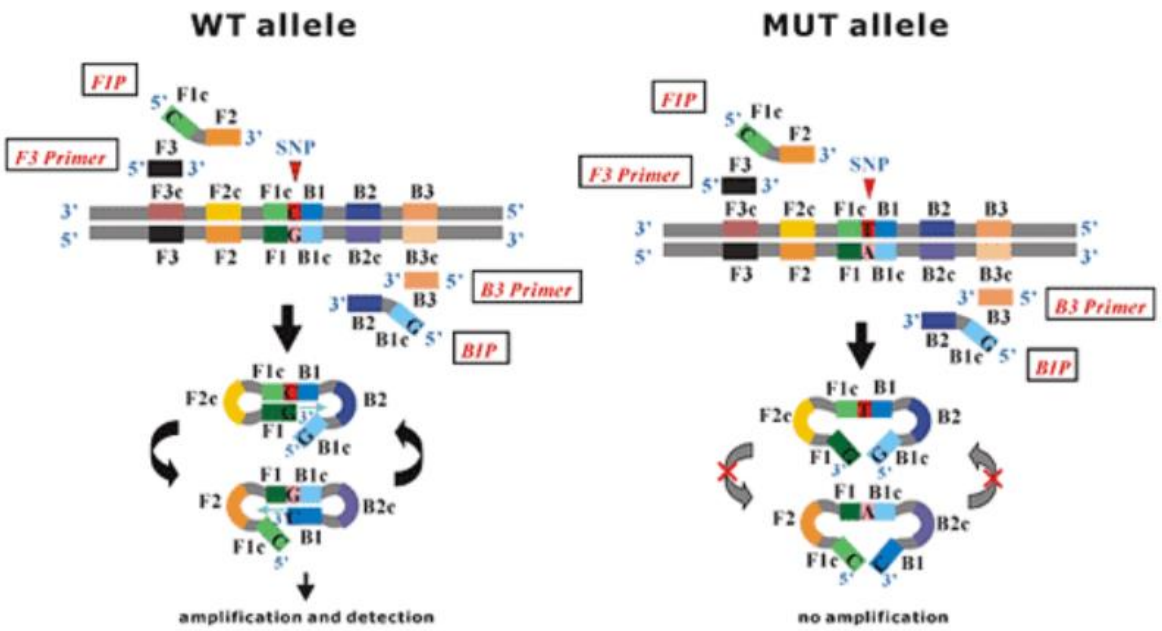

Figure 15. Schematic representation of the most commonly used LAMP-based SNP detection method [23].

Alternatively, the use of outer primers as allele-discriminating primers is proposed, placing the mutation at 3' end of the F3 sequence [57] or employing one inner primer with the mutated base at 5' end [58]. For all these strategies, in presence of the wild-type nucleotide the reaction performed with wild-type primer proceeds, while in presence of the polymorphic one the synthesis performed with wild-type primer is inhibited. Another approach described in literature is Peptide Nucleic Acid-Locked Nucleic Acid (PNA-LNA) mediated LAMP method. PNA probe is complimentary to the wild-type sequence and forms a thermodynamic stable complex with the dumbbell structure preventing the annealing and the extension of the LNA primer [59]. In this case, the amplification in presence of the wild-type sequence is blocked, while it occurs in presence of the mutate genome. Anyway, many of these approaches lead to a high background and a low discrimination efficiency due to the only partial inhibition of the non-specific amplification.

1.7 Bibliography

1. In Vitro Diagnostics Market Size and Share | IVD Industry Growth, 2027. Available at: https://www.fortunebusinessinsights.com/industry-reports/in-vitro-diagnostics-ivd-market-101443. (Accessed: 24th November 2020)

2. Chan, C. P. Y. *et al.* Evidence-based point-of-care diagnostics: Current status and emerging technologies. *Annu. Rev. Anal. Chem.* **6**, 191–211 (2013).

3. Vashist, S. K., Luppa, P. B., Yeo, L. Y., Ozcan, A. & Luong, J. H. T. Emerging Technologies for Next-Generation Point-of-Care Testing. *Trends in Biotechnology* **33**, 692–705 (2015).

4. Coskun, A. F., Nagi, R., Sadeghi, K., Phillips, S. & Ozcan, A. Albumin testing in urine using a smart-phone. *Lab Chip* **13**, 4231–4238 (2013).

5. Wei, Q. *et al.* Detection and spatial mapping of mercury contamination in water samples using a smart-phone. *ACS Nano* **8**, 1121–1129 (2014).

6. Vaitukaitis, J. L. Development of the home pregnancy test. in *Annals of the New York Academy of Sciences* **1038**, 220–222 (New York Academy of Sciences, 2004).

7. Tonyushkina, K. & Nichols, J. H. Glucose meters: A review of technical challenges to obtaining accurate results. *J. Diabetes Sci. Technol.* **3**, 971–980 (2009).

8. Niemz, A., Ferguson, T. M. & Boyle, D. S. Point-of-care nucleic acid testing for infectious diseases. *Trends in Biotechnology* **29**, 240–250 (2011).

9. Saeed, M. *et al.* GeneXpert technology: A breakthrough for the diagnosis of tuberculous pericarditis and pleuritis in less than 2 hours. *Saudi Med. J.* **38**, 699–705 (2017).

10. Tanriverdi, S., Chen, L. & Chen, S. A rapid and automated sample-to-result HIV load test for near-patient application. *J. Infect. Dis.* **201 Suppl 1**, (2010).

11. Kary B. Mullis, Henry A. Erlich, David H. Gelfand, Glenn Horn, R. K. S. Process for amplifying, detecting, and/or cloning nucleic acid sequences using a thermostable enzyme. (1987).

12. Yan, L. *et al.* Isothermal amplified detection of DNA and RNA. *Molecular BioSystems* **10**, 970–1003 (2014).

13. Deiman, B., Van Aarle, P. & Sillekens, P. Characteristics and applications of Nucleic Acid Sequence-Based Amplification (NASBA). *Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology* **20**, 163–179 (2002).

14. Brentano, S. T. & Mcdonough, S. H. Isothermal Amplification of RNA by Transcription-Mediated Amplification (TMA). in *Nonradioactive Analysis of Biomolecules* 374–380 (2000).

15. Hofmann, W. P. *et al.* Comparison of transcription mediated amplification (TMA) and reverse transcription polymerase chain reaction (RT-PCR) for detection of hepatitis C virus RNA in liver tissue. *J. Clin. Virol.* **32**, 289–293 (2005).

16. Hall, M. J., Wharam, S. D., Weston, A., Cardy, D. L. N. & Wilson, W. H. Use of signal-mediated amplification of RNA technology (SMART) to detect marine cyanophage DNA. *Biotechniques* **32**, 604–611 (2002).

17. Walker, G. T., Little, M. C., Nadeau, J. G. & Shank, D. D. Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 392–396 (1992).

18. Walker, G. T. *et al.* Strand displacement amplification - an isothermal, in vitro DNA amplification technique. *Nucleic Acids Res.* **20**, 1691–1696 (1992).

19. Fire, A. & Xu, S. Q. Rolling replication of short DNA circles. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 4641–4645 (1995).

20. Jeong, Y. J., Park, K. & Kim, D. E. Isothermal DNA amplification in vitro: the helicase-dependent amplification system. *Cellular and molecular life sciences : CMLS* **66**, 3325–3336 (2009).

21. Daher, R. K., Stewart, G., Boissinot, M. & Bergeron, M. G. Recombinase polymerase amplification for diagnostic applications. *Clinical Chemistry* **62**, 947–958 (2016).

22. Notomi, T. *et al.* Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res.* **28**, (2000).

23. Eiken GENOME SITE - The principle of LAMP method. Available at: http://loopamp.eiken.co.jp/e/lamp/. (Accessed: 1st December 2020)

24. Nagamine, K., Watanabe, K., Ohtsuka, K., Hase, T. & Notomi, T. Loop-mediated isothermal amplification reaction using a nondenatured template. *Clinical Chemistry* **47**, 1742–1743 (2001).

25. Nagamine, K., Hase, T. & Notomi, T. Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Mol. Cell. Probes* **16**, 223–229 (2002).

26. Nagamine, K., Hase, T. & Notomi, T. Accelerated reaction by loop-mediated isothermal ampli ® cation using loop primers. 223–229 (2002). doi:10.1006/mcpr.2002.0415

27. Notomi, T. *et al.* Loop-mediated isothermal amplification of DNA. **28**, (2000).

28. 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.* **154**, 150–154 (2001).

29. Tomita, N., Mori, Y., Kanda, H. & Notomi, T. Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. *Nat. Protoc.* **3**, 877–882 (2008).

30. Goto, M., Honda, E., Ogura, A., Nomoto, A. & Hanaki, K. I. Colorimetric detection of loop-mediated isothermal amplification reaction by using hydroxy naphthol blue. *Biotechniques* **46**, 167–172 (2009).

31. Jaroenram, W., Cecere, P. & Pompa, P. P. Xylenol orange-based loop-mediated DNA isothermal ampli fi cation for sensitive naked-eye detection of Escherichia coli. *J. Microbiol. Methods* **156**, 9–14 (2019).

32. Nathan A. Tanner. Visual detection of isothermal nucleic acid amplification using pH-sensitive dyes. *Reports* **68**, (2015).

33. Kaneko, H., Kawana, T., Fukushima, E. & Suzutani, T. Tolerance of loop-mediated isothermal amplification to a culture medium and biological substances. *J Biochem Biophys Methods* **70**, 499–501 (2007).

34. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

35. Collins, F. S., Morgan, M. & Patrinos, A. The Human Genome Project: Lessons from large-scale biology. *Science* **300**, 286–290 (2003).

36. Craig Venter, J. *et al.* The sequence of the human genome. *Science (80-.).* **291**, 1304–1351 (2001).

37. Sachidanandam, R. *et al.* A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–934 (2001).

38. Syvänen, A. C. Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nature Reviews Genetics* **2**, 930–942 (2001).

39. Zhang, X., Li, L., Ding, X. & Kaminsky, L. S. Identification of cytochrome P450 oxidoreductase gene variants that are significantly associated with the interindividual variations in warfarin maintenance dose. *Drug Metab. Dispos.* **39**, 1433–1439 (2011).

40. Davignon, J., Bouthillier, D., Nestruck, A. C. & Sing, C. F. Apolipoprotein E polymorphism and atherosclerosis: insight from a study in octogenarians. *Trans. Am. Clin. Climatol. Assoc.* **99**, 100–110 (1987).

41. Bertina, R. M. *et al.* Mutation in blood coagulation factor V associated with resistance to activated protein C. *Nature* **369**, 64–67 (1994).

42. Bali, S. *et al.* Genetic Fingerprinting of Potato Varieties from the Northwest Potato Variety Development Program. *Am. J. Potato Res.* **94**, 54–63 (2017).

43. Kalogianni, D. P. *et al.* Olive oil DNA fingerprinting by multiplex SNP genotyping on fluorescent microspheres. *J. Agric. Food Chem.* **63**, 3121–3128 (2015).

44. Agrimonti, C., Vietina, M., Pafundo, S. & Marmiroli, N. The use of food genomics to ensure the traceability of olive oil. *Trends Food Sci. Technol.* **22**, 237–244 (2011).

45. Kwok, P. Y. Methods for genotyping single nucleotide polymorphisms. *Annual Review of Genomics and Human Genetics* **2**, 235–258 (2001).

46. Sobrino, B., Brión, M. & Carracedo, A. SNPs in forensic genetics: A review on SNP

typing methodologies. *Forensic Science International* **154**, 181–194 (2005).

47. Kostrikis, L. G., Tyagi, S., Mhlanga, M. M., Ho, D. D. & Kramer, F. R. Spectral genotyping of human alleles. *Science* **279**, 1228–1229 (1998).

48. Tyagi, S., Bratu, D. P. & Kramer, F. R. Multicolor molecular beacons for allele discrimination. *Nat. Biotechnol.* **16**, 49–53 (1998).

49. Newton, C. R. *et al.* Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.* **17**, 2503–2516 (1989).

50. Landegren, U., Kaiser, R., Sanders, J. & Hood, L. A ligase-mediated gene detection technique. *Science (80-.).* **241**, 1077–1080 (1988).

51. Lyamichev, V. *et al.* Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.* **17**, 292–296 (1999).

52. Ansorge, W. J. Next-generation DNA sequencing techniques. *New Biotechnology* **25**, 195–203 (2009).

53. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).

54. Yongkiettrakul, S. *et al.* Simple detection of single nucleotide polymorphism in Plasmodium falciparum by SNP-LAMP assay combined with lateral flow dipstick. *Parasitol. Int.* **66**, 964–971 (2017).

55. Badolo, A. *et al.* Detection of G119S ace 1 R mutation in field collected Anopheles gambiae mosquitoes using allele specific loop mediated isothermal amplification (AS-LAMP) method. *Malar. J.* 1–8 (2015).

56. Kwong, K. M. *et al.* Comparison of Single Nucleotide Polymorphism genotyping of CYP2C19 by Loop-mediated isothermal ampli fi cation and real-time PCR melting curve. *Clin. Chim. Acta* **478**, 45–50 (2018).

57. Ferreira, F. *et al.* Allele speci fi c LAMP- gold nanoparticle for characterization of single nucleotide polymorphisms. *Biotechnol. Reports* **16**, 21–25 (2017).

58. Ikeda, S., Takabe, K., Inagaki, M., Funakoshi, N. & Suzuki, K. Detection of gene point mutation in paraffin sections using in situ loop-mediated isothermal amplification. *Pathol. Int.* 594–599 (2007).

59. Itonaga, M., Matsuzaki, I., Warigaya, K., Tamura, T. & Shimizu, Y. Novel Methodology for Rapid Detection of KRAS Mutation Using PNA-LNA Mediated Loop-Mediated Isothermal Amplification. *PLoS One* 1–12 (2016).

Aim of the work

The aim of this thesis work was the development of a competitive tool for SNP genotyping, fine-tuned to be integrated in a Point-of-Care platform.

In particular, this approach expected to be simple, rapid and suitable for on-site and large-scale screenings, with the purpose to propose an alternative solution to the complex and time-consuming traditional SNP-diagnostics. Moreover, the purpose was to develop a highly versatile test workflow, supported by a tailorable design and applicable, in this way, to a broad range of purposes.

For example, in industrial settings SNPs can be exploited as markers to unequivocally track commodities along the production chain. In such cases, managing directly on-site a high number of samples can provide remarkable advantages in standardized controls.

For these purposes, the objective of this thesis work was to propose a reliable tool able to meet the industrial requirements, introducing a minimally instrumented protocol with a rapid and simple procedure, to allow the rapid genotyping of a large number of samples and facilitate quality controls.

In clinical field, SNPs are highly relevant also for nutrigenetics and pharmacogenetics. In particular, several SNPs have been associated to the genetic predisposition to food intolerances, while some others have been linked to the susceptibility to drug treatments. Therefore, the on-site identification of the patient genetic profile can be useful for the Specialist to recommend the appropriate diet or pharmacological treatment. In these contexts, the aim was to offer an easy-to-use assay, based on a minimal instrumentation requirement and basic user-skills, designed to be performed in non-specialized settings, as for example at the Doctor's office or Pharmacy. Beside the test portability, the main goal of the thesis work was to propose a highly acceptable test, introducing a non-invasive sampling step to increase the patient's compliance.

Chapter 2

Model case for the optimization of PE-LAMP-Based assay

2.1 Introduction

Beside all the LAMP-based techniques for SNP genotyping covered above, a recent one proposed by Ding and collaborators in 2019 called probe-enhanced loop-mediated isothermal amplification (PE-LAMP) has revealed very effective.

Loop primers (LPs) have a pivotal role in the Ding's proposed strategy. In fact, in this strategy, LPs act as allele discriminating probes. To perform this role LPs are designed shorter in comparison with classic LAMP protocol and only one LP, forward or backward, is pre-added to each reaction mix.

The principle for SNP genotyping is based on designing two versions of the same LP: one complementary to the wild-type allele, and the other one perfectly matched with the polymorphic site. In particular, for each sample two reaction are performed in parallel, differing only for the type of LP used.

The core of the approach relies on triggering the specific annealing, while disfavoring the mismatched one. In this situation, a single mismatch located on a short primer significantly affects the thermodynamic stabilization of the bond between the probe and the template, impacting on the reaction efficiency. LPs act as accelerating primers, speeding up the onset of the amplification. Indeed, the reaction where the LP is perfectly hybridized will be properly enhanced, while the one with the mismatched annealing between the LP and the target sequence will be delayed. The discrimination between the two possible allele types is based on the time of onset of two concurrent reactions that depends on the stabilization or destabilization of the annealing between the short LP and the target sequence (Figure 16) [1].

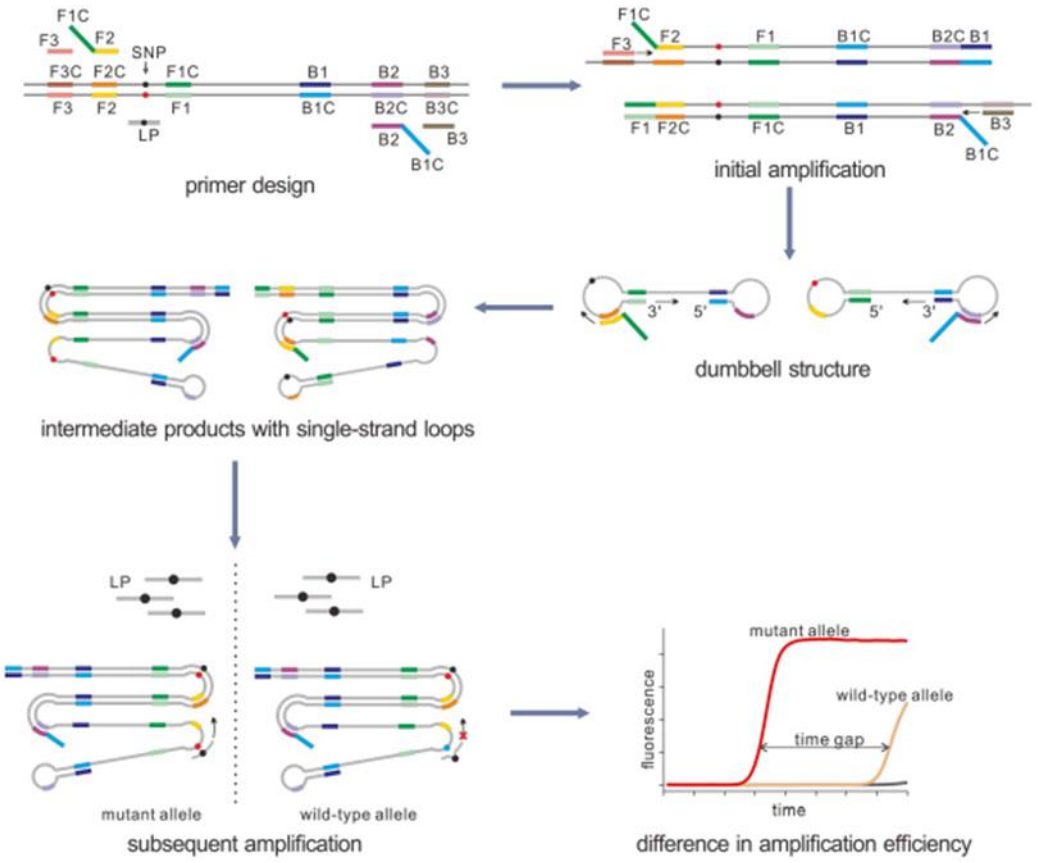

Figure 16. Schematic representation of SNP detection using PE-LAMP. In this strategy the SNP site on the target sequence corresponds to the LP binding site. The amplification proceeds forming a dumb-bell structure with the polymorphic nucleotide located on the loop portion. In this case the LP for the detection of the mutant target is designed and it is employed for the amplification of both mutant and wild-type samples. Indeed, this LP primer will trigger the accelerated amplification of the reaction containing the mutant target. However, it will not be able to proper enhance the reaction containing the wild-type target. Therefore, the two resulting amplification curves will differ for amplification efficiency [1].

Since this strategy resulted effective for SNP genotyping, in this thesis an in-depth optimization was performed.

The proposed thesis work was carried out at NanoBioInteractions & NanoDiagnostics Group at Italian Institute of Technology under the supervision of Dr. Pier Paolo Pompa and Dr.ssa Paola Cecere.

2.2 Results and discussion

In this thesis, starting from the state-of-art PE-LAMP approach [1], an optimized SNP genotyping strategy was developed, investigating one by one all the factors that influence the discrimination efficiency.

The mechanism was studied starting from Ding's outcomes and using the same model case: the polymorphism rs3741219. This particular genetic variation is located on Long Noncoding RNA H19 and in Chinese population is associated with the predisposition of insurgence of breast cancer [2]. Although in Italian population the selected SNP is not associated with any particular phenotype, since it is present in the general population, it was use as model SNP for the here-presented thesis work.

Human genomic DNA target for performing the assay was obtained from salivary samples. In fact, buccal epithelium cells are present in saliva. DNA was extracted from these cells with a simple protocol through a thermochemical lysis.

In rs3741219, the wild type allele has the nucleotide T, while the mutant one presents the nucleotide C.

As reported by Ding and collaborators, the method relies on the single SNP discrimination operated by loop primers (LPs). Accordingly to the principle of the strategy, two versions of the same LP were designed: LP-A, complementary to the wild type site, and LP-G, complementary to the polymorphic base. Two reactions were performed in parallel, one pre-adding LP-A and the other one pre-adding LP-G, to ensure that the same target was analyzed with both probes.

The different behaviors of the reactions, linked to the allele type present in the investigated sample, allowed the SNP genotyping. When the analyzed sample was a heterozygous type for the investigated SNP, then both amplifications were properly enhanced, otherwise, in case of a homozygous sample, the two reactions proceeded differently, in particular, the perfect matched reaction was significantly accelerated in comparison with the mismatched one.

In order to replicate the published results, the best performing pairs of primers that resulted by Ding's study were tested using human DNA of a volunteer resulted heterozygous for the investigated polymorphism. Amplification reactions were performed in Real-Time following

the fluorescent emission of SYBR green I dye, that bounds in a non-specific way the dsDNA, revealing the formation of LAMP products. The starting reaction conditions were identical to those applied by Ding et al. The employed LPs was composed by 11 nucleotides and the SNP site was located at the center of the probe. For an easier comprehension of the work, this couple of primers was named set 1. Although the DNA derived from a heterozygous individual, the reaction performed with LP-A resulted slightly delayed in comparison with the one performed with LP-G, making confusing the interpretation of the results (Figure 1). Moreover, the obtained result was very different from Ding's outcomes that provided a clear discrimination obtaining a time gap between the two reactions of 40 minutes.

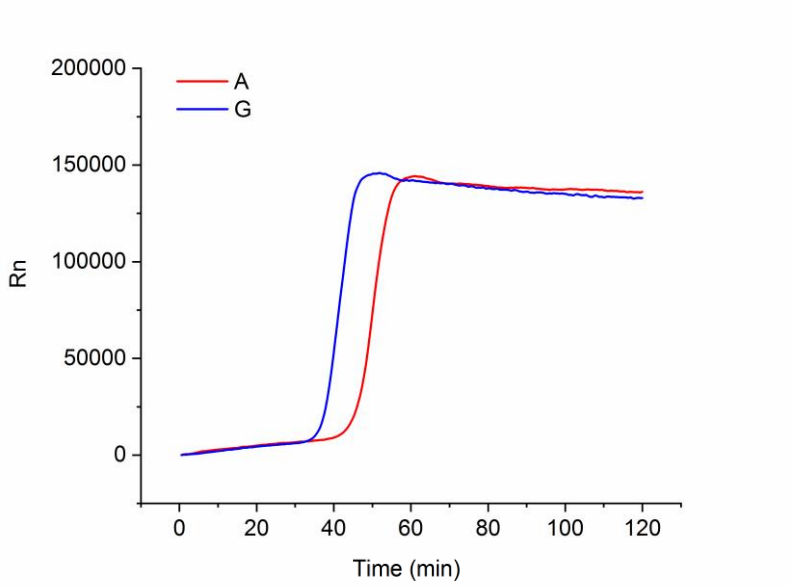

Figure 1. Real Time fluorescence detection of a heterozygosis sample using LP-A (red line) and LP-G (blue line) of LPs set 1.

With the aim to further optimize the approach, LP sequences were modified.

The main focus of this methods relies on the concurrent stabilization of the correctly matched reactions and destabilization of the mismatched ones. To trigger a greater destabilization of the non-specific reaction, two factors were investigated, such as the length of LP sequence and the position of the SNP site on the probes. Hence, two new LP sets were designed as follow: both were composed by 10 nucleotides, one with the SNP site located at the sequence center (set 2) and the other one with the SNP site at the 3' end (set 3). A heterozygous sample was tested using the two different LP sets (set 2 and set 3) with the same reaction conditions.

As reported in Figure 2, the results obtained with set 2 were confounding (Figure 2a), while with set 3 both allele types were correctly identify (Figure 2b).

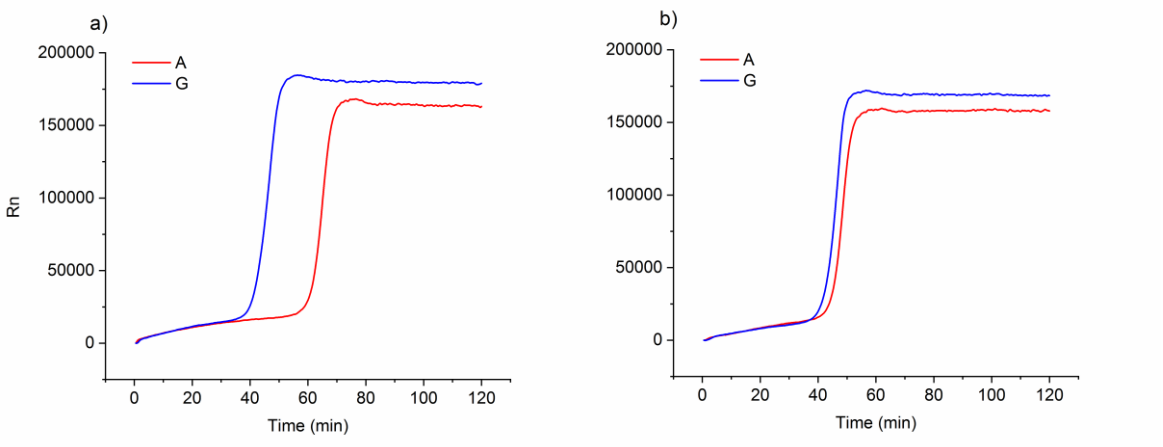

Figure 2. Real Time fluorescence detection of a heterozygosis sample using LP-A (red line) and LP-G (blue line) of two primer sets: set 2 (a) and set 3 (b).

Since the obtained results entirely relied on the LPs employed, all the tested primer sequences were compared: LPs employed in Ding's study (set 1), set 2 and set 3. The main difference between these pairs of probes referred to the melting temperature. In particular, the difference in melting temperature between LP-A and LP-G of each couple of primers was assessed. As reported in Table 1, among LP-A and LP G of set 1 the melting temperature difference was 5.5 °C, while the melting temperature difference between LP-A and LP-G of set 2 was 6.1 °C and between LP-A and LP-G of set 3 was 3 °C. In view of these observations, the results were explained. Each couple of probes had the same length, but a diverse sequence, one presented G and the other one A nucleotides. Consequently, the two probes had a different T melting, since it is well known that Guanine is involved in three hydrogen bounds during the hybridization with Cytosine, while Adenine and Thymine are connected by only two hydrogen bounds.

Therefore, the temperature required for separate the target and the probe sequences is higher for LP-G rather than for LP-A. For this reason, at the same reaction temperature, the percentage of the probe that is annealed to the target sequence is higher when in the sequence is present the Guanine instead of Adenine. Thus, the reaction performed with LP-

G was more stable and, as a result, it was faster rather than the one performed with LP-A. Accordingly with this reading, all the other obtained results revealed clear. In particular, the melting temperature of LP-G of set 1 was higher in comparison with the one of LP-A, therefore, although both allele types were present in the reaction mix, the reaction performed with LP-G occurred faster rather than the one performed with LP-A (Figure 1). Moreover, the melting temperature difference between LP-A and LP-G of set 2 was even higher in comparison with the one of the first pairs of primers tested and biased towards LP-G. As a consequence, the time of occurrence of the two reactions was varying presenting an even bigger time frame (Figure 2a). On the other hand, the melting temperature difference between LP-A and LP-G of set 3 was 3 °C, thereby, even if the reaction performed with LP-G was slightly accelerated, the time gap between the two reaction was relatively narrow (Figure 2b).

For all these reasons, to reduce the unbalance due to the nature of nucleotides interactions in the DNA double helix, another primer set was designed.

The fourth set of primers was designed equalizing the melting temperature of the two probes (A and G) at the cost of achieving two probes with different length. The resulting set 4 presented a difference in melting temperature of 0.4°C. As hypothesized, the perfect overlap of the two curves was obtained, clearly allowing the genotyping of a heterozygous sample (Figure 3). In fact, in case of heterozygous sample even with different sequences both primers should be designed in order they potentially accelerate both reactions with the same rate.

Overall these findings demonstrated that among all the considered factors, such as primer length and SNP position on the sequence, the melting temperature difference between the two probes had the strongest impact on reaction efficiency. Higher was the melting temperature difference between the pairs of probes, higher was the difference in time of occurrence of the two parallel reactions.

Table 1	
Primers sets	**Δ T melt between LP-A and LP-G (°C)**
Set 1	5.5
Set 2	6.1
Set 3	3
Set 4	0.4

Table 1. Melting temperature differential values between the couple of LP-A and LP-G) of the tested sets.

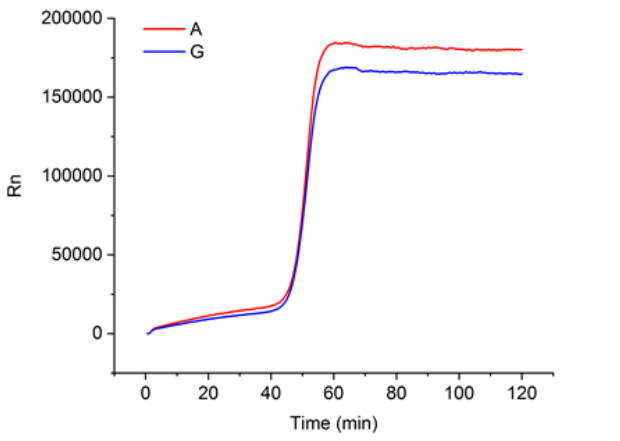

Figure 3. Real Time fluorescence detection of a heterozygosis sample using LP-A (red line) and LP-G (blue line) of primer set 4.

With the aim to simplify all the procedures to make this strategy suitable for a POC diagnostic test, the possibility to skip the dilution of the extracted DNA was considered. Since primers set 4 resulted the most performant, four volounteers were genotyped employing LP-A and LP-G of set 4 without equalizing the template concentrations. As reported in Figure 4, samples 2 and 4 resulted heterozigous, while samples 1 and 3 resulted

homozigous wild-type (AA) and homozigous mutant (GG) for the investigated SNP, respectively. In fact, in sample 1 the reaction performed with LP-A is accellerated rather than the one performed with LP-G suggesting that the investigated genome was characterized by the A variant in homozigosis (AA). On the other hand, in sample 3 the behaviour is inverted indicating the presence of the genotype homozigous GG.

Although the fluorescent discrimination was clear, starting with a different amount of template resulted disadvantageous for the development of a colorimetric version of the assay. In fact, with these reaction conditions, it was impossible to select a precise time point corresponding to the moment when only the perfectly matched reactions occurred, while the non-specific ones not. For this reason and in order to avoid false positive results, for the development of a colorimetric assays is necessarily required that all the specific reactions onset at the same time and visibly separated from all the other reactions.

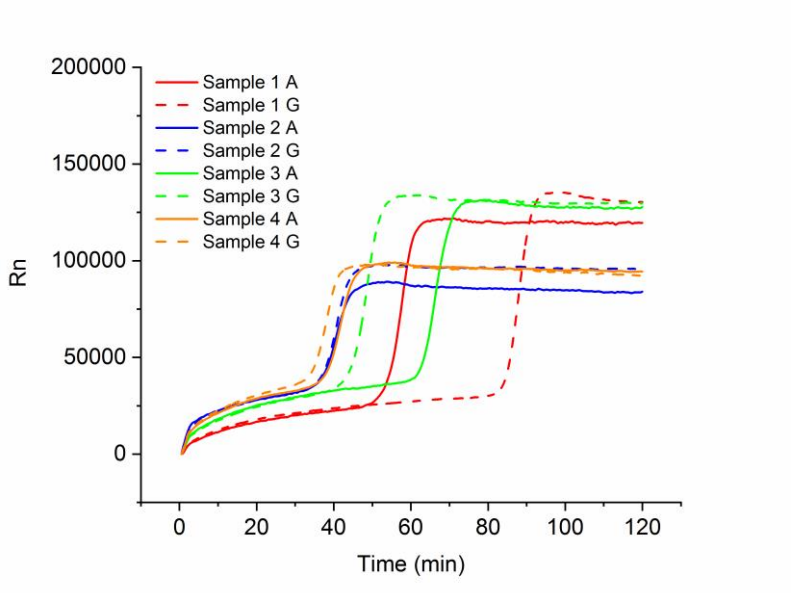

Figure 4. Real Time fluorescence detection of four samples using LP-A (solid line) and LP-G (dashed line) of primer set 4 without diluting the templates.

LAMP is a very sensitive and fast amplification technique. For this reason, LAMP amplification efficiency had to be reduced in order to obtain a single-nucleotide discrimination. In fact, the amount of the reagents usually employed in the reaction mix of classical LAMP protocol are not convenient for this aim. The goal was to sufficiently weak both amplification reactions by modifying reaction conditions, including template and enzyme concentrations. In this way, it was possible to destabilize as much as possible the non-specific amplification, without losing the stabilization of the perfectly matched reaction.

Therefore, with the purpose to obtain a large time frame that could ensure a reliable genotyping also with the colorimetric test, the effect of the template concentration factor on the discrimination efficiency was analyzed.

10’000 and 5’000 copies for reaction as final concentration of the template were employed and the relative results were compared. Three samples, one heterozygous and two homozygous (wild-type and mutant) DNA, were tested to compared each reaction condition.

As showed in Figure 5 and 6, the main effect of the progressive reduction of the template amount was the retardation of non-specific reactions. In particular, it was observed that in both reaction conditions the perfectly matched reactions occurred almost simultaneously after 35/40 minutes. Otherwise, the time of occurrence of the non-specific reactions resulted fluctuating. Overall, employing 10’000 copies for reaction the resulted time gap between the fast and slow reactions was 20/30 minutes (Figure 5) while reducing the template amount the time frame resulted bigger (30/40 minutes) (Figure 6).

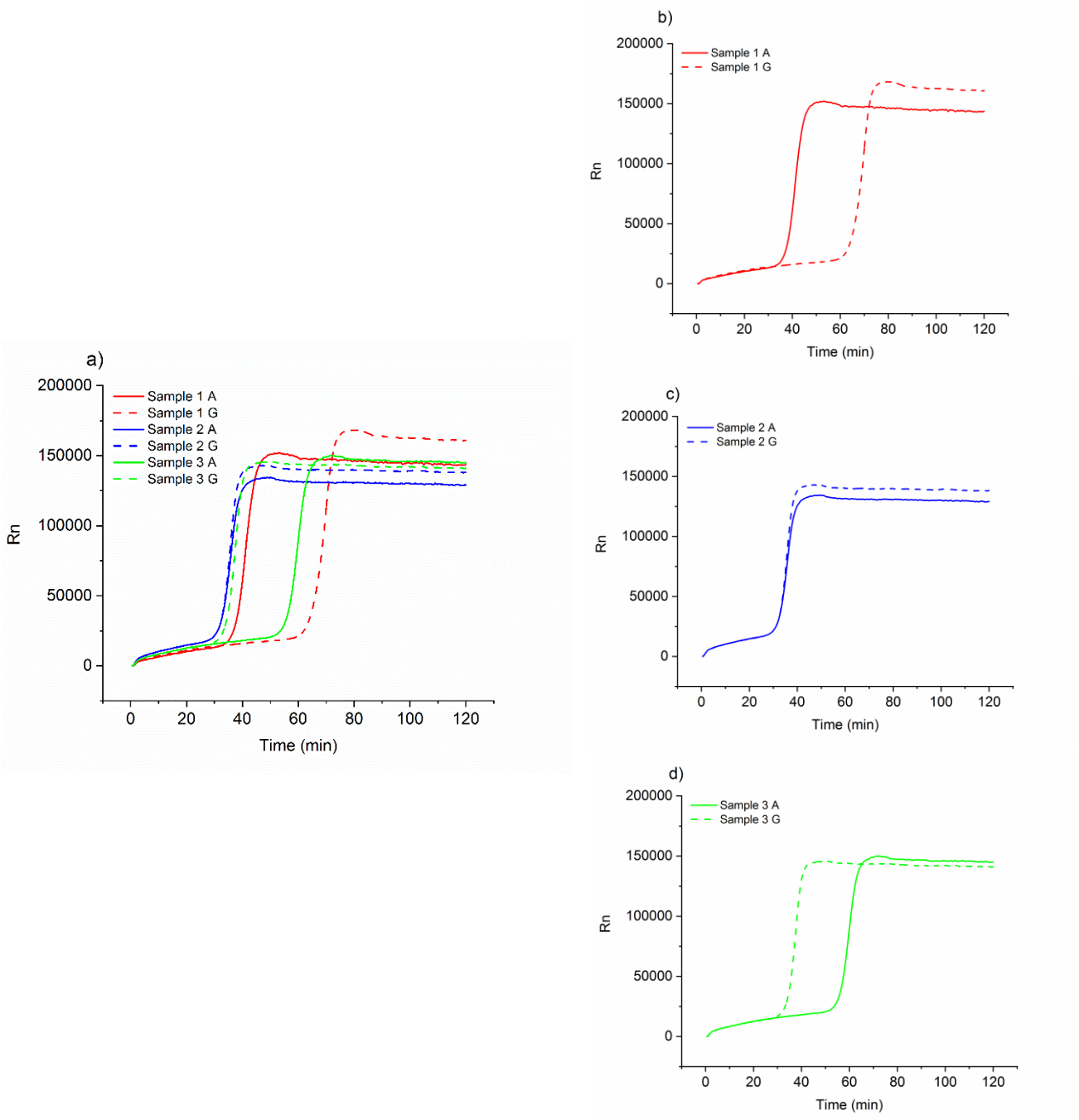

Figure 5. Real Time fluorescence genotyping of three samples using LP-A (solid line) and LP-G (dashed line) of primer set 4 using the template final concentration of 10'000 copies for reactions.

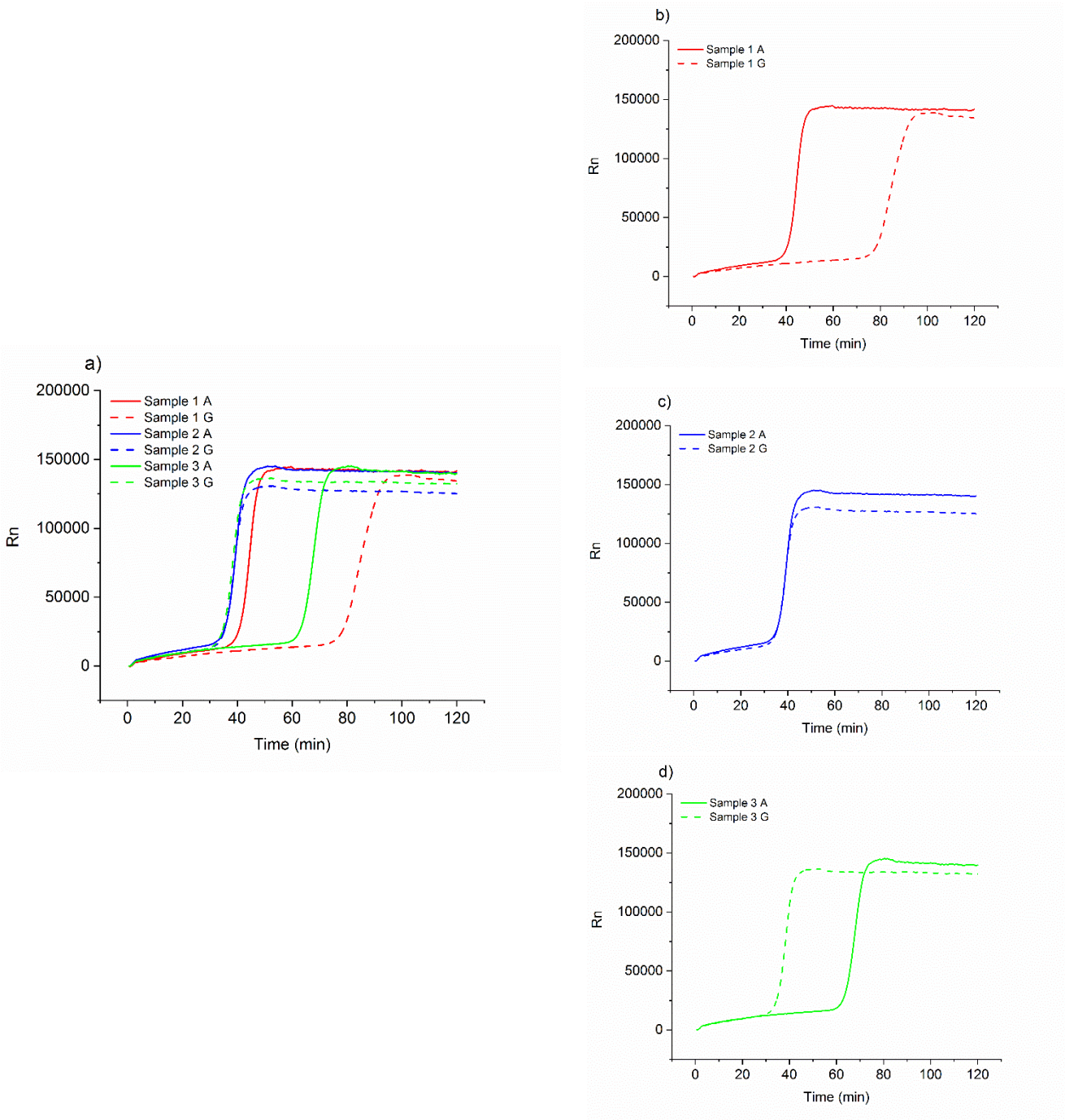

Figure 6. Real Time fluorescence genotyping of three samples using LP-A (solid line) and LP-G (dashed line) of primer set 4 using the template final concentration of 5'000 copies for reactions.

These findings suggested that to enhance the differences between the times of occurrence of the two reactions is important to reduce the reaction efficiency. In particular, the

weakening of the reactions would lead to a greater impairment of the non-specific amplifications.

Therefore, to further disadvantage the mismatch annealing, standard enzyme concentration was reduced of 30% (from 160 U/ml to 112 U/ml). Firstly, 112 U/ml as enzyme final concentration was tested in combination with the final concentration of template of 5’000 copies for reactions. As a result, all the perfectly matched reactions occurred at the same time but with a slight delay (time of occurrence: 40 minutes) and the time gap between the fast and slow reactions did not increased (Figure 7).

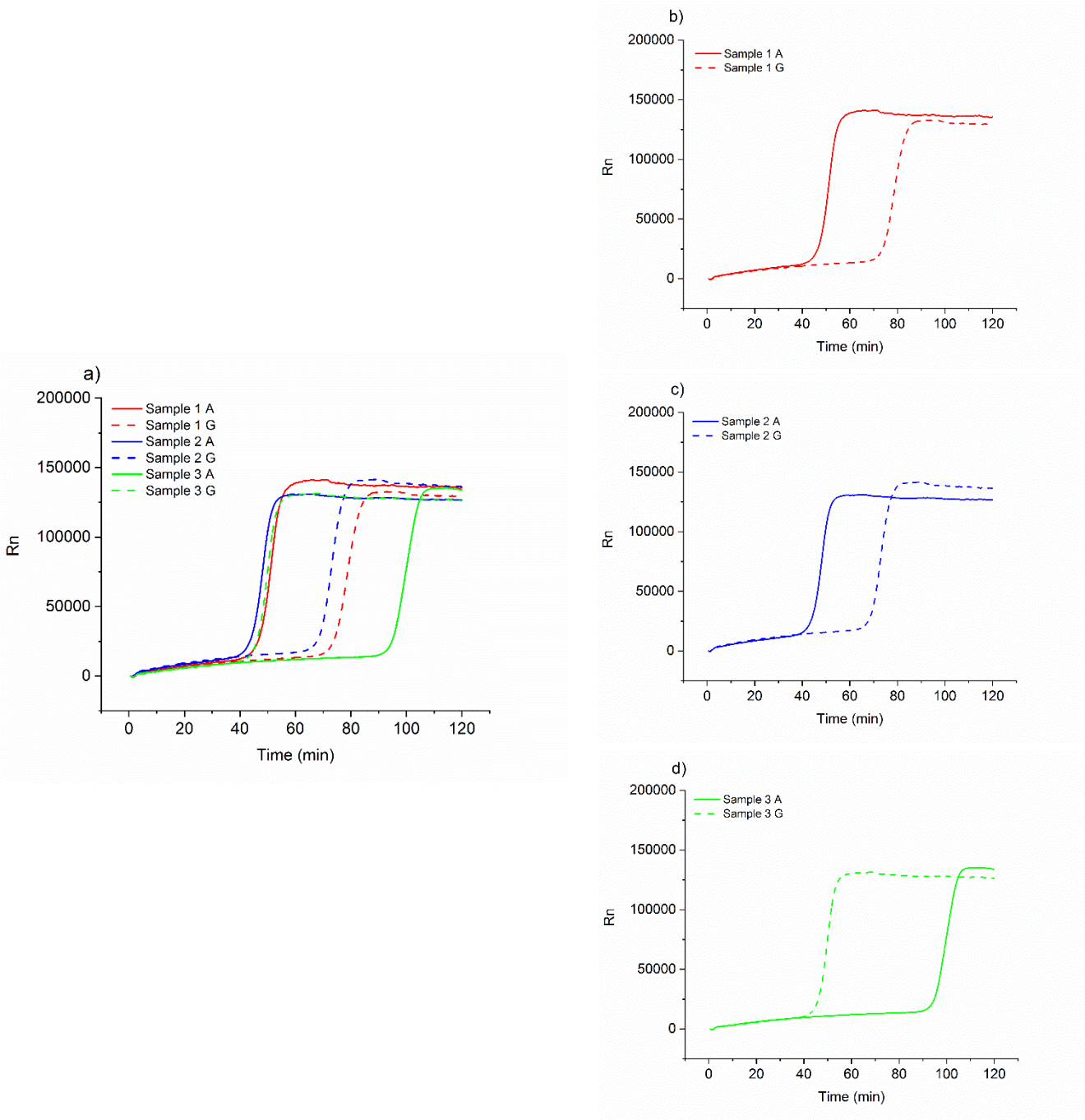

Figure 7. Real Time fluorescence genotyping of three samples using LP-A (solid line) and LP-G (dashed line) of primer set 4 using the template final concentration of 5'000 copies for reactions and a reduction of 30% of the classical enzyme concentration.

Then, a further reduction of the template concentration was tested (1’000 copies for reactions) in combination with the reduced enzyme concentration (112 U/ml). With these conditions, the best results in terms of time frame was observed gaining a difference of 40/45 minutes between the time of insurgence of the fast and slow reactions (Figure 8).

In particular, all the specific reactions occurred with the same rate after about 50 minutes while the non-specific ones occurred after 90 minutes. These findings allowed a clear allele discrimination and guaranteed reliable results. Moreover, they are promising for the development of the colorimetric version of the assay. In fact, with these conditions a precise time point is easily identifiable and, most important, given the intrinsic characteristics of the assay, the not occurrence of false positive results is assured.

Finally, to confirm the result, eight people were genotyped. Interesting, the result was highly reproducible and cumulative resulted time frame was of about 40 minutes (Figure 9).

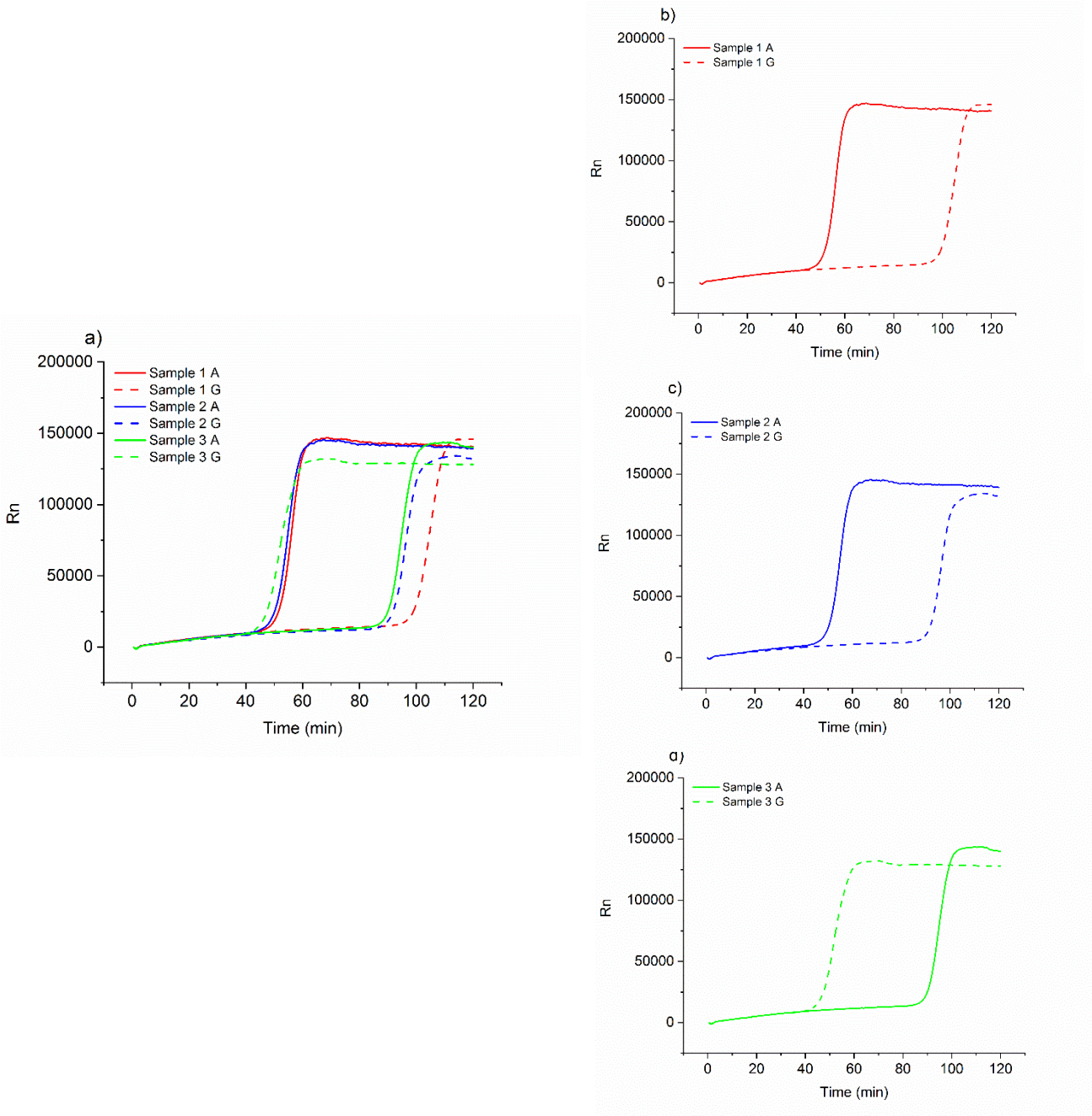

Figure 8. Real Time fluorescence genotyping of three samples using LP-A (solid line) and LP-G (dashed line) of primer set 4 using the template final concentration of 1'000 copies for reactions and reduction of 30% of the classical enzyme concentration.

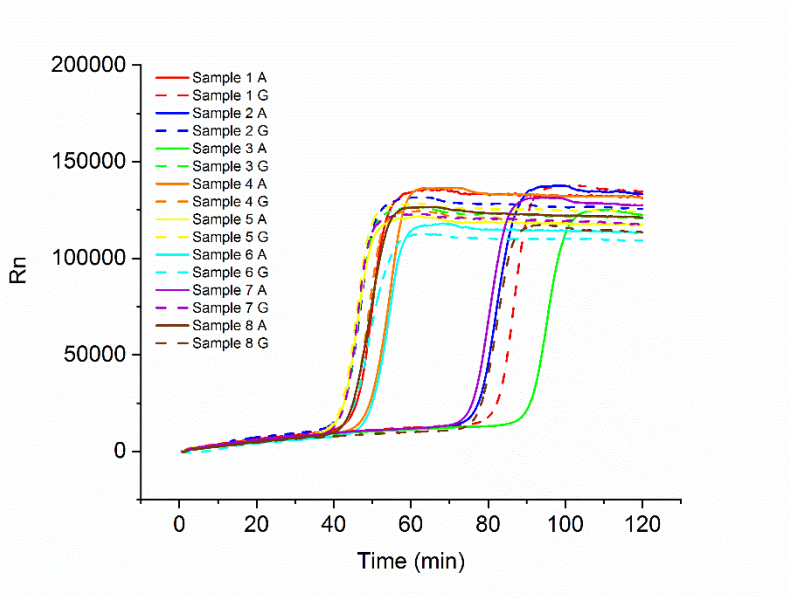

Figure 9. Real Time fluorescence genotyping of eight people using LP-A (solid line) and LP-G (dashed line) of primer set 4 using the template final concentration of 1'000 copies for reactions and reduction of 30% of the classical enzyme concentration.

2.3 Material and methods

DNA Extraction from saliva samples

DNA was extracted from saliva samples using the following protocol. 1mL of saliva was mixed and 4 mL of Phosphate Saline Buffer (PBS) were added. The sample was centrifuged at 1800 rcf for 5 minutes at 4°C and the surnatant was discarded. 1 mL of milliQ water was added to the pellet, the sample was mixed, collected in a safe lock tube and centrifuged at 12000 RPM for 5 minutes at 4°C. The surnatant was discarded and 150 µL of saliva Pretreatment Buffer, composed by PBS and EDTA were added. This buffer allowed to thin the saliva sample and to inactivate nuclease enzymes.

Sample was mixed at 500 RPM for 30 minutes at 56°C. Then, sample was mixed using a Vortex Mixer (VWR International SRL, Milano, Italy). Subsequently the sample was mixed at 500 RPM for 8 minutes at 100°C.

Finally the sample was centrifuged at 12000 RPM for 5 minutes at 4°C. The resulting surnatant contains the extracted ssDNA.

100 µL of the extracted ssDNA was collected and the concentration was measured using Nucleic Acid quantification tool of NanoDrop 2000c UV-visible Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA concentration was converted from ng/µL to copies/µL using DNA copy number calculator tool of Thermo Fisher Scientific website [3]. Finally, samples was diluted in MilliQ water at the desired final concentration.

Primer Design for the optimization of PE-LAMP strategy

Primer sequences for the optimization of PE-LAMP strategy were designed in collaboration with Dr.ssa Paola Cecere of the NanoBioInteractions&NanoDiagnostics Group.

Optimized fluorescent-based LAMP Reaction for PE-LAMP

LAMP reactions were performed in 25 µL of a mixture containing 1.6 µM each of inner primers, 0.4 µM each of outer primers, 0.8 µM of loop primers (Integrated DNA Technologies, Coralville, IA, USA), 1 M of betaine (VWR International SRL, Milano, Italy), 2.5 µL of 10X LAMP buffer (200 mM of Tris-HCl, 100 mM of $(NH_4)2SO_4$, 20 mM of $MgSO_4$, 500 mM of KCl, and 1% v/v Tween 20), 2 mM of $MgSO_4$, 0.8 mM each of dNTPs (Promega, Madison, WI, USA), 0.112 U/µL of Bst 2.0 WarmStart DNA Polymerase (New England

BioLabs, Ipswich, MA, USA), and 5 µL of DNA template at the concentration of 200 copies/µL of the extracted DNA. DNA-free LAMP reactions were included as negative controls. Amplification reactions were performed using forward and backward inner primers (FIP and BIP), forward and backward outer primers (F3 and B3), and one loop primer (LP-A or LP-G). For real-time fluorescent LAMP, 1/50,000 diluted SYBR Green (Thermo Fisher Scientific,Waltham, MA, USA) was pre-added to the reaction mix. Real-time amplifications were performed on an Applied Biosystem real-time instrument (Thermo Fisher Scientific,Waltham, MA, USA) with StepOne Software v2.3 at 63 °C.

2.4 Conclusions

In conclusion, a state-of-art strategy for single-SNP genotyping was optimized combining it with a simple and rapid DNA extraction from saliva samples. The optimized method, exploiting LPs as allele discriminating primers, allowed a precise detection of the allele type present in the sample, resulting in a large difference in terms of times of occurrence.

Moreover, the optimized approach resulted promising for the development of an efficient colorimetric test, ensuring the avoidance of false positive results. The possibility to develop the colorimetric version of the assay coupled with a simple and portable instrumentation that is the only need for the performing of the test are ideal for on-site diagnostics.

The model case studied enabled to acquire a deep knowledge of the variables that mostly influence the resulted discrimination. Such acquired expertise, opened the perspective for the development of tests for SNP genotyping with possible real applications in relevant clinical cases and in quality control fields.

2.5 Bibliography

1. Ding, S. *et al.* One-step colorimetric genotyping of single nucleotide polymorphism using probe-enhanced loop-mediated isothermal amplification (PE-LAMP). *Theranostics* **9**, (2019).

2. Xia, Z. *et al.* Genetic polymorphisms in long noncoding RNA H19 are associated with susceptibility to breast cancer in Chinese population. *Med. (United States)* **95**, e2771 (2016).

3. DNA Copy Number Calculator | Thermo Fisher Scientific - IT. Available at: https://www.thermofisher.com/it/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/dna-copy-number-calculator.html. (Accessed: 1st December 2020)

Chapter 3

A naked-eye assay for varietal traceability of durum wheat

3.1 Introduction

The results obtained in the present chapter have been published in Foods 2020, 9(11), 1691; https://doi.org/10.3390/foods9111691 with the title "A Fast, Naked-Eye Assay for Varietal Traceability in the Durum Wheat Production Chain" authored by Giulia Cibecchini, Paola Cecere, Giorgio Tumino, Caterina Morcia, Roberta Ghizzoni, Paola Carnevali, Valeria Terzi and Pier Paolo Pompa.

Food traceability currently has a key role in the agri-food sector, bringing many benefits to both industries and consumers, such as supply chain optimization, improvement in food safety and quality, and an increase in controls and in the consumers' confidence. Over the years, specific regulations and stringent controls have been introduced in several countries with different level of stringency to ensure food quality and safety and safeguard consumer's health and rights. In particular, food traceability is regulated by several laws that demand to track "one step back" and "one step forward" of the supply chain [1]. Italian pasta market, for example, is regulated by the Law n.580 of 1967 and by the subsequent Decreto del Presidente della Repubblica 41, 5 March 2013 [2, 3]. In such dispositions the use of durum wheat (*Triticum durum*) for pasta production was declared mandatory specifying that a contamination with *Triticum aestivum* species must not exceed the 3% maximum level. In fact, Durum wheat features, such as the gluten composition and the hard caryopsis make it more suitable for pasta production rather than common wheat [4] . Mandatory traceability actions are flanked by voluntary ones, characterized by the identification of standards related, for example, to the valorization of traditional food products and local economy development [5]. At the industrial level, pasta for example is commonly made by mixtures of different durum wheat cultivars of *Triticum durum* wheat. At the same time, there is an increasing presence in the market of high-value mono-varietal pasta from both local producers and industrial ones. In such productions, there is great interest in developing a

voluntary traceability system to identify specific varieties along the production chain. The worldwide-spread pasta producer industry, Barilla, normally stipulates cultivation contracts with a restricted range of varieties, used as mix for pasta production. However, there is a special line, the Aureo wheat that is singularly used in the production. Among all the varieties of durum wheat, Aureo is considered one of the "top quality" since it has high protein and gluten content [6]. Beside the peculiar characteristics of the variety that permits to raise the quality of pasta products, Aureo is cultivated in Italy, boosting this way Italian market and reducing import [7].

Varietal fingerprinting approaches have been developed from mainly to protect intellectual property (IP) of wheat variety to guarantee plant breeders' rights (PBR). Distinctness, uniformity, and stability are among the basic requirements for a variety. The traditional morpho-physiological descriptors, collected during the whole plant life cycle, are classically used for such evaluation purposes. In a further step, morphological descriptors have been flanked by biochemical markers, e.g., the isoforms of seed storage proteins. Such markers can be informative when applied to the grains, but are not effective on transformed products, in which technological processes can impair the proteins' integrity. The subsequent advancement in varietal fingerprinting was based on the introduction of DNA markers that are independent of the environmental influence and available in theoretically unlimited numbers.

Many classes of DNA molecular markers have been developed, ranging from inter simple sequence repeats (ISSRs) to amplified fragment length polymorphism markers (AFLP) and random amplified polymorphic DNA (RAPD), as reviewed by Pasqualone, 2011 [8]. SSRs (simple sequence repeats), considered the second generation of molecular markers and characterized by a simple technique, high polymorphism rate, and robustness, have been widely employed in wheat [9, 10]. SNPs (single-nucleotide polymorphisms) are the third generation of molecular markers, with several advantages, including high frequency across the whole genome, ease of detection, and cost efficiency [11].

Classical techniques for SNP genotyping was described in Chapter 1. In particular, for the wheat genome, SNP arrays are available and permit the identification of 9000 to 90,000 polymorphisms [12, 13, 14]. Moreover, with the development of next-generation sequencing (NGS) technology, increasing numbers of SNPs have been discovered, building the basis for

the development of diagnostic SNP barcodes [15]. Now, the "universalization and minimization of SNP number without compromising identification accuracy is the major challenge in development of varietal profile by rapid genotype assay", as reported by Singh et al., 2019 [16].

However, classical methods are time consuming, expensive, and require costly instrumentations and specialized personnel, limiting their practical applicability for quality control screenings in the food supply chain. A rapid, portable technique for SNP discrimination would thus be of great industrial interest.

In pasta industry, both at industrial and local level, is very important distinguish among breeds trough the study of DNA to ensure a correct labelling of end-products valorizing the production and detecting fraud. In particular a greater advantage would be to be able to estimate the purity levels of the wheat grains delivered to the mill as Aureo variety, and consequently the pureness of the semolina derived by it. Additionally, inside an industrial reality, the possibility to have a tool for rapid screening a large number of samples provide, as advantage, the promptness of the acceptance (or not) of batches of grains at the time of delivery.

Indeed, this thesis work is focused on the necessity, mainly industrial, of having a fast and reliable method for wheat grains and semolina control through the detection of a single SNP. The selected SNP was chosen not only because it is considered a rare variant, but also because it is private of only one variety among the grains covered by the contracts stipulated by Barilla.

Therefore, the strategy optimized in chapter 2 was applied to address this specific requirement of pasta industry. The strategy was proven to be suitable for a portable platform and was demonstrated to be efficient in genotyping wheat cultivars. In particular, was able to detect the presence of Aureo, but also to discover substitution and adulteration of grains made products with different cultivars.

3.2 Results and discussion

The rapid and simple approach described in chapter 2 was applied to address an interesting case, relevant for industrial settings. In particular, the optimized single-SNP discriminating strategy was applied to reveal the presence of Aureo durum wheat in pasta products and to detect the possible adulteration of the declared mono-varietal products with other less precious ingredients.

The results of the following work have been published. The project was performed in collaboration with the research group Dr. Valeria Terzi of Council for Agricultural Research and Economics, Research Centre for Genomics and Bioinformatics, and with Barilla S.p.A Company.

The aim of the work is to provide an efficient tool for Barilla Company quality control area.

This study led to identify a single relevant SNP as marker for the varietal discrimination. For this aim, 29 selected cultivars were analyzed with the DArTseq methodology. The analysis produced a set of 20198 good-quality SNPs (with a percentage of missing values below 25%). Two replicate per variety were performed in order to reduce the risk to select a private SNP allele by a genotyping error. The average genetic distance between two replicates was 98.5% (simple matching), suggesting a genotyping error rate of 1.5% on average. A set of 11 candidate SNPs carrying a private allele of Aureo was identified, with 6 of them presenting no missing values. The SNP marker for discrimination of Aureo was selected among those 6 SNPs with complete genotypes. The selected SNP was proven to be peculiar only for the Aureo genotype of durum wheat.

Starting from the identification of this discriminating SNP, the SNP-sensitive detection method described in the chapter 2 was adapted and extended for the rapid and on-site authentication of Aureo durum wheat in food raw materials.

As explained in chapter 2, the design of loop primers (LPs) has a pivotal role in this strategy. The criterion for SNP discrimination is based on the designing of two different LPs: one perfectly matched with the genome of one cultivar and the other one complementary to the sequence peculiar of all the others cultivars. In this strategy, the length of LPs is shorter than the one used in standard LAMP protocols. In particular, the approach relies on the stabilization of the annealing between short LPs and the perfectly complementary sequence,

while the non-specific one is disfavored. LPs act as accelerating primers, speeding up the onset of the amplification. Hence, the reaction in which the LP is perfectly hybridized will be properly enhanced. On the contrary, due to the short LP length, at the reaction temperature the single mismatch will influence the thermal stability of the primer-target annealing, causing the primer melting and a relevant amplification delay (Figure 1).

The genetic variation G>T is located on chromosome 7A of durum wheat genome. Hence, Aureo presents the polymorphic base T (G>T), while all the other cultivars are characterized by the nucleotide G, located at the same position.

LPs were constructed as the following: LP-T, complementary to the SNP that identifies Aureo variant, and LP-G, perfectly matching the SNP distinctive of all of the other durum wheat varieties. This was sufficient for developing our single-SNP approach, which aimed at the discrimination of a variety within a relatively small panel. Such an approach would be very suitable for specific industry requirements. However, the genotyping effort required to identify a private SNP allele might be higher in situations with a larger panel of varieties or lower genetic diversity. To overcome this issue, a multi-SNP assay could be developed, since a private allele can be tagged more easily using more than one SNP.

Each sample was tested with two reactions performed in parallel. The two reactions differed only for the LP. Since the single mismatch located on a short primer significantly impairs the annealing reducing the amplification rate, the two reaction will behave differently.

The results of the SNP-sensitive LAMP reactions on real wheat samples was performed in Real-Time following the fluorescent emission of SYBR green I dye that intercalating to the dsDNA. The increase of the fluorescent emission is directly proportional to the amount of LAMP product. The amplification curve related to the reaction performed with the Aureo variant had a perfect annealing when performed with LP-T, so in this case, the amplification was properly enhanced. The exponential fluorescent signal related to the accelerated reaction, in fact, emerged after 40 minutes. On the other hand, with the same template, the reaction performed with LP-G was destabilized and significantly delayed (Figure 2a).

The resulting time gap between the two amplification curves was about 40 min, clearly allowing the discrimination.

Conversely, the reactions performed using Marco Aurelio, another genotype of durum wheat taken as example, revealed the inverted behavior, with similar high temporal discrimination between the accelerated/matched (G) and non-matched (T) reactions. This result demonstrated that, in this configuration, a single mismatch can properly modulate the amplification efficiency of LAMP; consequently, it is possible to identify one specific durum wheat cultivar with high discrimination effectiveness (Figure 2b).

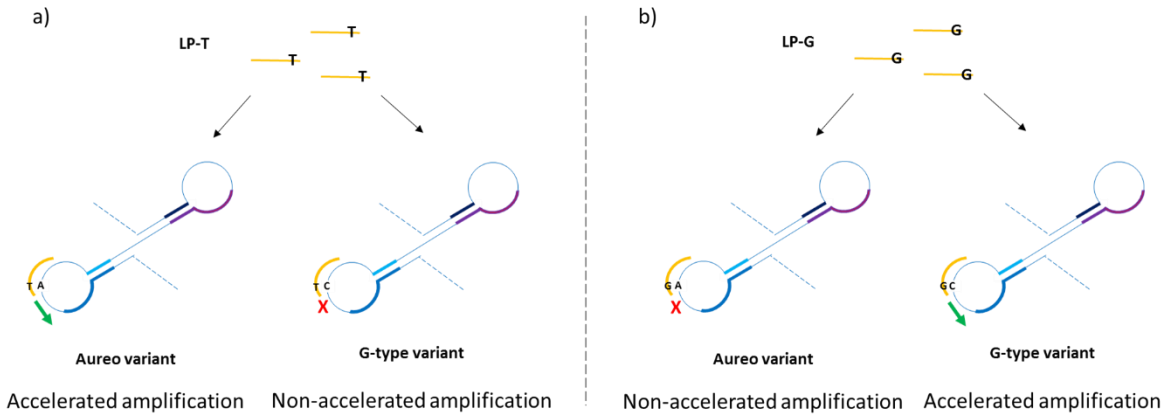

Figure 1. Schematic representation of the hybridization of LPs to amplicon loops from durum wheat cultivars. During the exponential phase of LAMP reaction, the perfect match between LP-T and the complementary region on the Aureo intermediate loops accelerates the amplification, while the presence of the single mismatch between LP-T and G-variant amplicon loops causes the short primer destabilization and a significant reaction delay (a). On the contrary, for LP-G-mediated reactions, the perfect match between LP and the complementary region on the target is achieved in case of the presence of a G-type variant. Consequently, in case of the presence of a G-type variant the reaction is accelerated, while in present of Aureo cultivar the amplification is delayed (b).

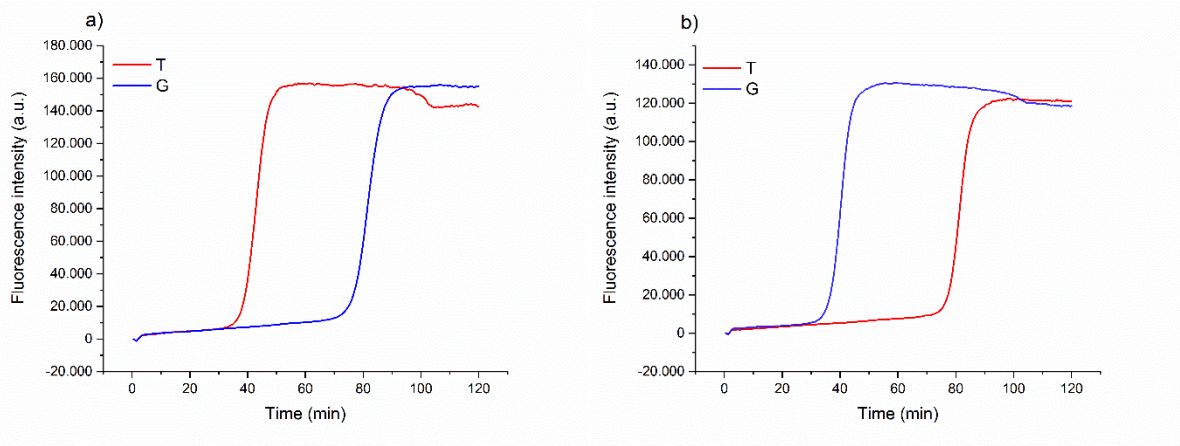

Figure 2. Real-Time fluorescence detection of genotype od durum wheat. LAMP reaction with LP-T (red line) and LP-G (blue line) of Aureo genotype (a) and one G-type variant genome (Marco Aurelio) (b). The time gap between fast and slow reactions is ca. 40 minutes. All the reactions were performed at 63 °C.

To evaluate the reliability of the technique in a real scenario of quality control, 13 durum wheat cultivars were tested. Interestingly, the method was able to efficiently discriminate all of the 13 tested cultivars of wheat versus Aureo with an average time gap of 35-40 minutes (Figure 3). Moreover, it is important to notice that, for all the samples tested, the time gap was nearly constant among all the tested cases and the time of occurrence of the accelerated reactions was constant. Therefore, this method resulted effective in the identification of a substitution of Aureo with another variety, allowing the detection of food fraud in less than 1 hours.

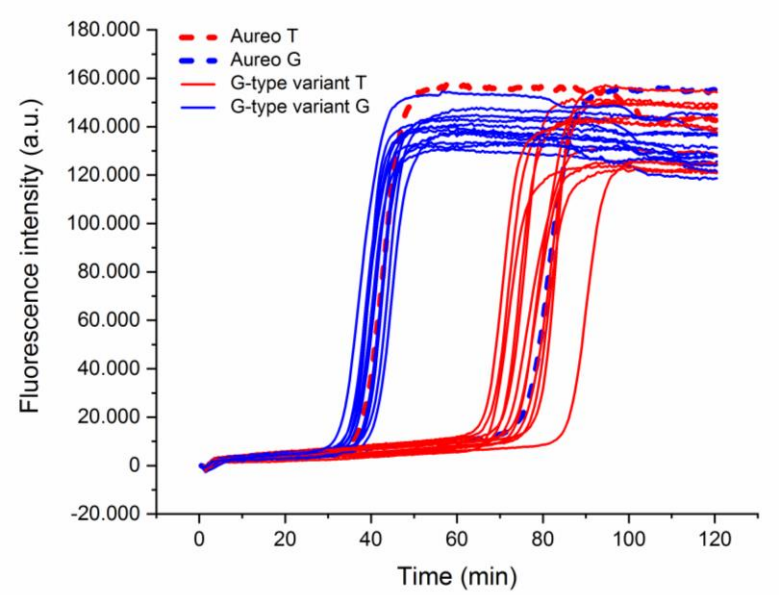

Figure 3. Real-time fluorescence detection of 13 genotype of durum wheat. LAMP reactions with LP-T (red lines) and LP-G (blue) of Aureo and G-type variants (Aureo: dashed lines). The time gap between all the fast and all the slow reactions is 35-40 minutes. All the reactions were performed at 63 °C.

Adulteration of food component with less valuable raw materials is considered a fraud as well as the substitution of food with undeclared ingredients.

Therefore, a more complex scenario of food fraud, in which the Aureo species is adulterated (mixed) with another variety, with different levels of contaminations was investigated. This is important in industrial settings to guarantee the mono-varietal nature of some productions.

To this purpose, different mixtures of Aureo with decreasing contaminations of a G-type variant (from 50 to 10%) were prepared to probe the sensitivity of the strategy to also detect minor adulterations. In this case, Babylone cultivar was used as representative of all the other genotypes of durum wheat.

As reported in Figure 4, the method can effectively discriminate the 50/50 (Aureo/G-type) mixture, with the two reactions displaying similar efficiency. In such a case, the two reactions both behave like accelerated reactions, because the half amount of template of each cultivar is sufficient for the optimized functioning of the amplification (Figure 4c). Overall, the mixture appeared as a heterozygous sample. Comparable results were obtained with a 25% contamination with the G-type variant (Figure 4b). Although the amount of G-

type genome was lower, the reaction with LP-G worked efficiently, exhibiting a negligible delay. Interestingly, our LAMP test was also able to discriminate an adulteration with the G-type variant as low as 10% (Figure 4a). In this latter case, we observed a small delay of the LP-G reaction due to the significantly lower (10%) amount of G-type variant present in the mixture.

However, the test behaved very differently with respect to the result obtained with 100% pure Aureo mixture, overall indicating that the proposed strategy enables accurate discrimination of relevant adulterations, including both substitution and mixing of wheat cultivars with a limit of detection of 10%.

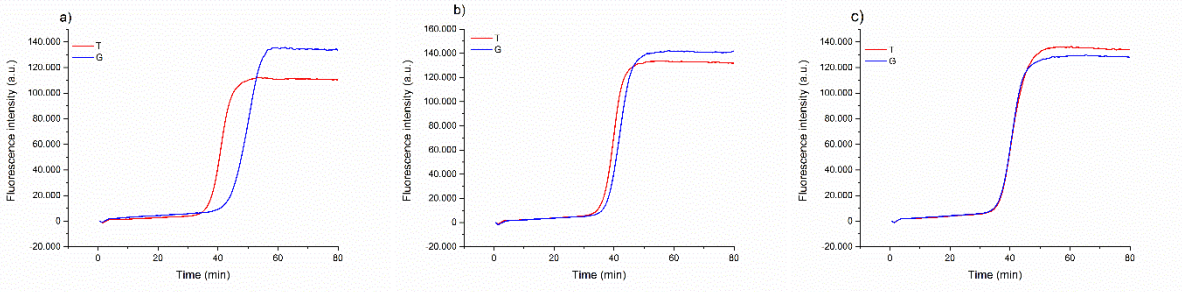

Figure 4. Real-time fluorescence detection of a simulated adulteration, in which a sample of Aureo was mixed with different proportions of a G-type variant (Babylone). (a) 90% of Aureo with 10% of G-type variant; (b) 75% of Aureo with 25% of G-type variant; (c) 50% of Aureo with 50% of G-type variant. The LAMP reactions were performed with LP-T (red line) and LP-G (blue line) at 63 °C.

Both the large time gap, together with the same time of onset of the curves related to the accelerated reactions, were important starting points for the development of the colorimetric configuration of this assay, based on a naked-eye detection of the results. Among the others, colorimetric strategies offer excellent platform to provide fast, on-site and cost-effective diagnostic tests, thanks to the minimal instrumentation required, as a simple portable heater device.

As explained in Chapter 1, incorporation of nucleotides on the newly synthetized DNA strand by DNA polymerase activity leads to the release of hydrogen atoms. Therefore, during the amplification the pH of the reaction environment drops. As already reported in

literature, the decrease of pH value during the amplification can be exploited to obtain a color change. Indeed, a colorimetric interpretation of the results is possible pre-adding a pH-sensitive dye that could provide a clear color change of the reaction tubes upon target amplification [18, 19].

For these reasons, LAMP reaction was performed pre-adding to the reaction mixture the Cresol Red dye, which shifted from purple to yellow while the pH decreases from 8.8 to 7.2. In this way, at the starting pH value, the solution color is purple. In presence of target, the target amplification causes a solution color shift from purple to yellow, indicating that LAMP products have reached a significant concentration, and suggesting that the exponential reaction had started in an early stage. On the contrary, the solution keeps the purple color. Hence, in the SNP-sensitive LAMP assay, stopping the amplification at a specific time point, the positive results (yellow) corresponds to the accelerated reaction, while the negative results (purple) represents the amplification reaction that could not reached a proper stage and the sufficient amount of hydrogen ions, required to trigger the color change. In particular, since the fluorescent-based findings showed that all the fast reactions occurred after 30/40 minutes, while the slow ones arose almost after 60 minutes, it was considered to stop the colorimetric reaction at 1 hour in order to obtain the color change only in case of the accelerated reactions.

In line with previous experiments, for the same target two reactions were carried out: one adding LP-T (tube A) and one adding LP-G (tube B); all the reactions were performed at 63°C for 1 h. As reported in the scheme in Figure 5a, the combination of the yellow/purple colors of the two reactions performed in parallel allowed detecting the possible scenarios, i.e., pure Aureo, the presence of contaminations (cultivar mixing), or full substitution. Colorimetric reactions were performed using as template 100% of Aureo genotype, 100 % of G-type variant and genome mixtures of both. In particular, different proportion of contaminant cultivar were tested, in particular 50%, 25% and 10% of G-type variant (figure 5b). In the case of the pure Aureo cultivar, the reaction performed with LP-T turned yellow (positive result), while the reaction performed with LP-G remained purple (negative result) (Figure 5b, sample 1). On the contrary, the full substitution of Aureo with a G-type variant exhibited the inverted result, with tube A appearing purple and tube B yellow (Figure 5b, sample 5). Interestingly, all of the adulterated samples with different levels of

contaminations (down to 10% adulteration) were clearly detected, with both of the tubes exhibiting yellow color (samples 2–4) due to the amplification of both genotypes of durum wheat. As well as the tests performed in real time, all the mixtures including the adulteration with a G-type variant genome at 10%, were detected with a good sensitivity. Moreover, this strategy was proven to be specific and had enough sensitivity for quality control purposes and the visual inspection of the results being an additional advantage for on-field applications.

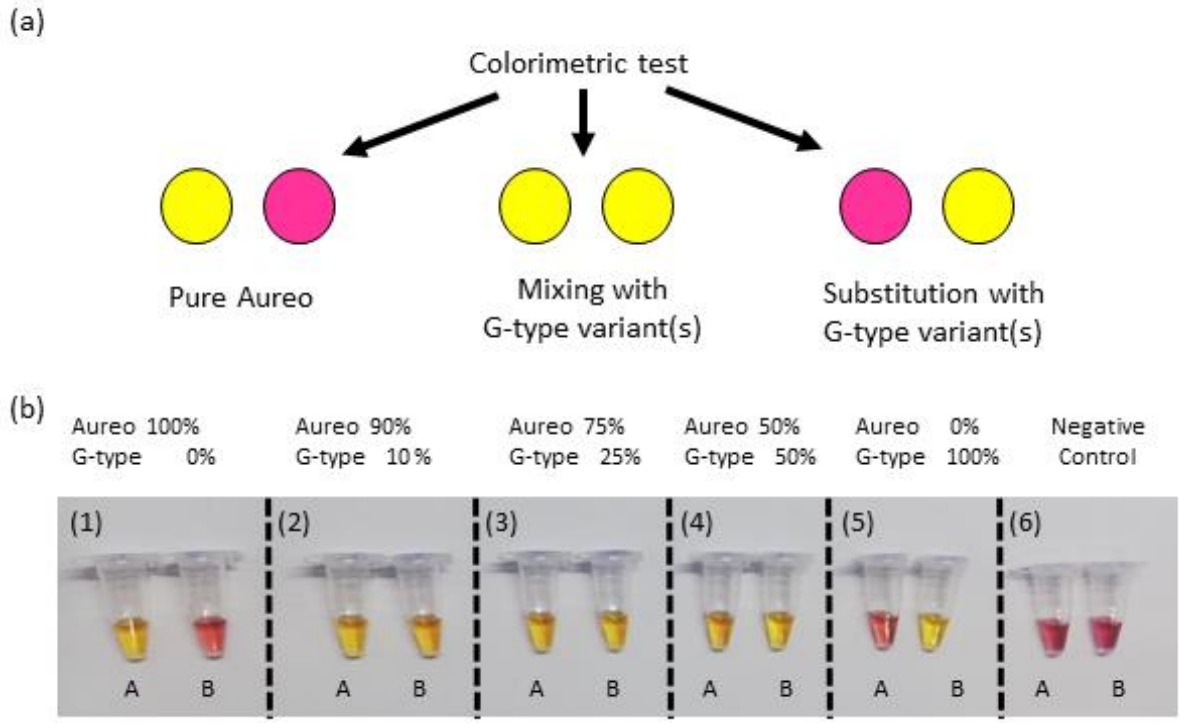

Figure 5. a) Scheme of the colorimetric test displaying the possible yellow/purple color combinations: pure Aureo cultivar (yellow/purple), adulterated mixture with G-type varieties (yellow/yellow), or fully substituted product (purple/yellow). (b) Colorimetric detection of possible adulteration cases of Aureo with another durum wheat cultivar (here Babylon as an example). LAMP reactions with LP-T (tube A) and LP-G (tube B) were all performed at 63 °C for 1 h. From left to right, targets were composed by 100% of Aureo variant of wheat (sample 1), then mixtures of Aureo and G- type variant (from 10 to 100%, samples 2-5), and negative control (sample 6).

To completely validate our system, a more complex case of adulteration was tested, employing a mixture composed by three cultivars. In particular, the mixture was composed

by Aureo and two different G-type varieties of wheat, Babylone and Marco Aurelio representing an adulteration ranging from 50 to 10% of total contaminations.

Remarkably, also in this case, the strategy revealed its high specificity, being able to correctly identify the different genotypes in a complex mixture and even in the situation in which the total amount of the two contaminations was as low as 10% (Figure 6).

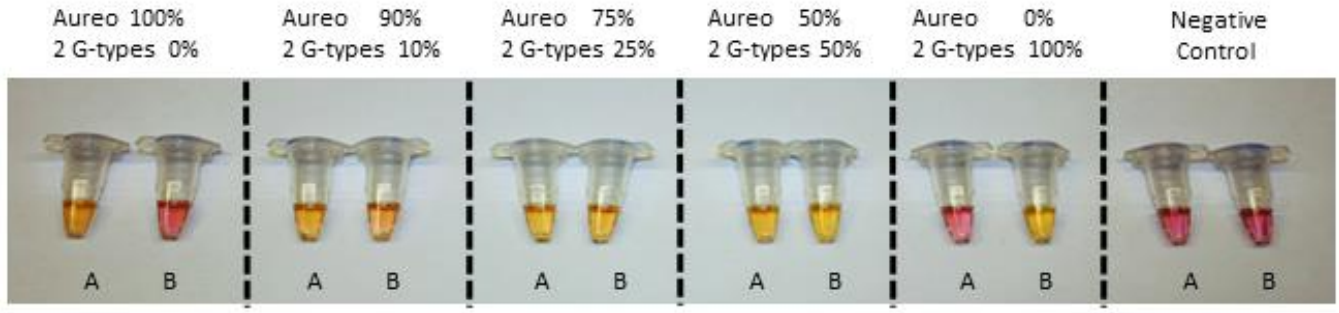

Figure 6. Colorimetric detection of a complex adulteration of Aureo with two G-type wheat variants. LAMP reactions with LP-T (tube A) and LP-G (tube B) were all performed at 63 °C for 1 h. From left to right, targets were composed of 100% Aureo variant, then mixtures of Aureo and two G-type variants (Babylon + Marco Aurelio) with a total contamination from 10 to 100%, and negative control.

3.3 Materials and methods

Plant material and DNA extraction

The following 28 durum wheat cultivars were used: Iride, Rusticano, Saragolla, Odisseo, Maestrale, Bronte, Antalis, Fabulis, Core, Svevo, Orizzonte, Aureo, Achille, Monastir, Claudio, Tirex, Pigreco, Normanno, Marco Aurelio, Relief, Miradoux, Babylone, Simeto, Anvergur, Navigator, Levante, Kyle, and Kronos. Moreover, the tetraploid Triticum turanicum, variety QK-77 (traded as Kamut®), was included in the analysis. All of the cultivars were certified foundation seeds produced directly from the breeder responsible for their maintenance in purity. Ten certified seeds of each variety were sown in duplicates in pots, and the foliar tissues were harvested at the three-leaves stage. DNA samples were isolated and purified using the cetyl trimethyl ammonium bromide method [17]. The quality check, quantification, and concentration adjustment were accomplished with a NanoDrop2000C Spectrophotometer (Thermo Fisher Scientific, Monza, Italy). The concentration of each sample was adjusted to 50 ng/µL. Two biological replicates of DNA extracts were used for each variety.

Genotyping by sequencing and SNP data analysis

Purified DNA samples (1 µg for each sample) were sent to Diversity Arrays Technology Pty Ltd. (http://www.diversityarrays.com/, Canberra, Australia) for sequencing, and SNP marker identification was done by DArTseq genotyping. Sequences of the genomic representations were obtained on a HiSeq2500 instrument.

The set of SNP markers generated by DArTseq was curated, removing markers with more than 25% of missing values (null alleles were considered as missing). SNPs with low frequent alleles were kept in the dataset to allow the research of private alleles of Aureo. This resulted in a dataset containing 20,198 SNPs and 58 samples (two genotyping replicates for each cultivar). After marker curation, all samples had a percentage of missing values below 10% and a percentage of heterozygous calls below 12%. Genetic distances were calculated based on simple matching. A custom R script was used to select SNPs with no missing values and a private allele of both replicates of Aureo. Candidate SNPs were mapped on the durum wheat genome sequence [18] by BLAST (Basic Local Alignment Search Tool). Starting from a

panel of 11 candidates, an SNP mapped on chromosome 7A was selected for further analyses.

Primer design for LAMP strategy

LAMP primers were designed from the 7A chromosome of durum wheat. Primer sequences are shown in Table 1:

Table 1. Names and Primer sequences used in the study.

Name of Primers	**Sequence (5'–3')**
FIP	TGCAGTGACTGATTGTACTGTCCACCACTTCCTCAGGTA
BIP	GTACTGCACTACTGCACCATTCTCGCCTGCAAACACAC
F3	CTGCCGTTGCCAACA
B3	TATCCGCACGCACC
LoopT	TGCAGGCGATG
LoopG	GCAGGCGAGG

Fluorescent-based LAMP reactions

LAMP reactions were performed in 25 µL of a mixture containing 1.6 µM each of inner primers, 0.4 µM each of outer primers, 0.8 µM of loop primers (Integrated DNA Technologies, Coralville, IA, USA), 1 M of betaine (VWR International SRL, Milano, Italy), 2.5 µL of 10X LAMP buffer (200 mM of Tris-HCl, 100 mM of (NH4)2SO4, 20 mM of MgSO4, 500 mM of KCl, and 1% v/v Tween 20), 2 mM of MgSO4, 0.8 mM each of dNTPs (Promega, Madison, WI, USA), 0.112 U/µL of Bst 2.0 WarmStart DNA Polymerase (New England BioLabs, Ipswich, MA, USA), and 5 µL of DNA template at the concentration of 5 ng. DNA-free LAMP reactions were included as negative controls. Amplification reactions were performed using forward and backward inner primers (FIP and BIP), forward and backward outer primers (F3 and B3), and one loop primer (LP-T or LP-G). For real-time fluorescent

LAMP, 1/50,000 diluted SYBR Green (Thermo Fisher Scientific,Waltham, MA, USA) was pre-added to the reaction mix. Real-time amplifications were performed on an Applied Biosystem real-time instrument (Thermo Fisher Scientific,Waltham, MA, USA) with StepOne Software v2.3 at 63 °C.

LAMP reactions for colorimetric assay

LAMP colorimetric reactions were performed as reported above, using 1 M of betaine (VWR International SRL, Milano, Italy), 2.5 µL of 10X LAMP buffer (100 mM of $(NH_4)2SO_4$, 20 mM of $MgSO_4$, 500mMof KCl, and 1% v/v Tween 20), 2mMof $MgSO_4$, 0.8mMeach of dNTPs (Promega, Madison, WI, USA), 0.112 U/µL of Bst 2.0 WarmStart DNA Polymerase (New England BioLabs, Ipswich, MA, USA), and 5 µL of DNA template at the concentration of 5 ng. DNA-free LAMP reactions were included as negative controls. For visualized detection, 0.06 mM of Cresol Red (Sigma-Aldrich, St. Louis, MO, USA) was pre-added to the reaction mix. The amplification efficiency was verified by color changing of the reaction mix. Colorimetric LAMP reactions were performed on a T100 Thermal Cycler (BIO-RAD, Hercules, CA, USA) at 63 °C for 1 h, followed by heat inactivation at 90 °C for 2 min.

3.4 Conclusions

Food identification and traceability are crucial step to ensure the quality of the production. Nowadays, rapid and accurate on-site tests are emerging to flank the commonly used techniques for genetic traceability. In this framework, a single-SNP discriminating strategy was designed to correctly identify mono-varietal products. Moreover, a colorimetric version of the assay, able to discriminate by naked-eye inspection, the purity or the adulteration of Aureo with any G-type variants of durum wheat was developed. The use of isothermal amplification, the LAMP, together with the visualization of the readout by a color change obtained in a short time, make this approach suitable for portable testing. This approach fulfill the need for a rapid and reliable pre-screening of food raw material in pasta industries and was proven to be potentially convenient in industrial applications.

3.5 Bibliography

1. Charlebois, S., Sterling, B., Haratifar, S. & Naing, S. K. Comparison of Global Food Traceability Regulations and Requirements. *Compr. Rev. Food Sci. Food Saf.* **13**, 1104–1123 (2014).

2. LEGGE 4 luglio 1967, n. 580 Disciplina per la lavorazione e commercio dei cereali, degli sfarinati, del pane e delle paste alimentari. Available at: https://www.gazzettaufficiale.it/eli/id/1967/07/29/067U0580/sg. (Accessed: 1st December 2020)

3. Mipaaf - DECRETO DEL PRESIDENTE DELLA REPUBBLICA 5 marzo 2013, n. 41 - Regolamento recante modifiche al DPR 9 febbraio 2001, n. 187, concernente la revisione della normativa sulla produzione e commercializzazione di sfarinati e paste alimentari. Available at: https://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/6230. (Accessed: 1st December 2020)

4. Morcia, C. *et al.* A Chip Digital PCR Assay for Quantification of Common Wheat Contamination in Pasta Production Chain. *Foods* **9**, 911 (2020).

5. Banterle, A. & Stranieri, S. Sustainability Standards and the Reorganization of Private Label Supply Chains: A Transaction Cost Perspective. *Sustainability* **5**, 5272–5288 (2013).

6. AUREO - Varietà Frumento duro | Syngenta Italia. Available at: https://www.syngenta.it/prodotti/sementi/frumento-duro/aureo. (Accessed: 1st December 2020)

7. Barilla acquisterà oltre 140.000 tonnellate di grano italiano dalla Puglia. Available at: https://www.barillagroup.com/it/comunicati-stampa/nuovi-contratti-triennali-di-coltivazione-del-grano-duro-2017-2019-al-sud . (Accessed: 1st December 2020)

8. Pasqualone, A. Authentication of durum wheat-based foods: classical vs. innovative methods.

9. Huang, X. Q., Börner, A., Röder, M. S. & Ganal, M. W. Assessing genetic diversity of wheat (Triticum aestivum L.) germplasm using microsatellite markers. *Theor. Appl.*

Genet. **105**, 699–707 (2002).

10. Laidò, G. *et al.* Genetic Diversity and Population Structure of Tetraploid Wheats (Triticum turgidum L.) Estimated by SSR, DArT and Pedigree Data. *PLoS One* **8**, e67280 (2013).

11. Nadeem, M. A. *et al.* DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. *Biotechnol. Biotechnol. Equip.* **32**, 261–285 (2018).

12. Cavanagh, C. R. *et al.* Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. **110**, (2013).

13. Wang, S. *et al.* Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. 787–796 (2014).

14. Gao, L., Jia, J. & Kong, X. A SNP-Based Molecular Barcode for Characterization of Common Wheat. *PLoS One* **11**, e0150947 (2016).

15. Mangini, G. *et al.* Exploring SNP diversity in wheat landraces germplasm and setting of a molecular barcode for fingerprinting. *Cereal Res. Commun.* **46**, 377–387 (2018).

16. Singh, R. *et al.* Development of model web-server for crop variety identification using throughput SNP genotyping data. *Sci. Rep.* **9**, (2019).

17. Nathan A. Tanner. Visual detection of isothermal nucleic acid amplification using pH-sensitive dyes. *Reports* **68**, (2015).

18. Yongkiettrakul, S. *et al.* Simple detection of single nucleotide polymorphism in Plasmodium falciparum by SNP-LAMP assay combined with lateral flow dipstick. *Parasitol. Int.* **66**, 964–971 (2017).

Chapter 4

A fast, fluorescent-based predictive test for the detection of genetic predisposition to lactose intolerance

4.1 Introduction

Lactose is a disaccharide composed by glucose and galactose. Lactase enzyme, which is located in the small intestine, hydrolyzes lactose, introduced with diet, into its two components.

In case of deficiency of lactase enzyme, as in lactose intolerance, lactose is digested by bacteria inside the colon. Therefore, lactose intolerance symptoms are related to the fermentation of lactose by colonic bacteria that produces gas, causing the consequently intestinal disorders such as nausea, bloated feeling, intestinal cramps, meteorism and diarrhea.

Lactose intolerance can be classified in the following types: Congenital lactase deficiency (CLD), Developmental Lactase Deficiency, Lactase Non-Persistence (NPL) or Primary Hypolactasia and Secondary Lactase Deficiency. CLD is a rare inherited disorder that affect 1 on 60,000 children in the word with a particular prevalence in the Finnish population. Five nonsense mutations have been founded as the cause of this severe gastrointestinal disorder. In particular, these mutations are located on the coding region of the lactase gene impairing irreversibly the final function of the encoded enzyme [1]. Another form of lactose intolerance is the Developmental Lactase Deficiency, which is a condition that occurs in premature babies. In fact, lactase activity increases in fetus during gestation until birth. Therefore, prematurely born infants present a lactose intolerance related to the enzyme deficiency [2]. However, in this case lactose malabsorption symptoms can be avoided by the action of particular species of colonic bacteria [3]. Lactase Non-Persistence (NPL) or Primary Hypolactasia is the most common form of intolerance, and it is related to a downregulation of the enzyme after weaning until the adulthood when the Lactose activity is completely lost [4]. Finally, the lactose intolerance

form called Secondary Lactase Deficiency generally is a reversible condition because it is a consequence of other diseases, such as Chron's disease or infections that cause damages to the intestine [3].

About 70 % of the world population presents Primary Hypolactasia. Its incidence is significant in South America, Africa, Asia and variable in Europe [5, 4, 6]. In Italy, lactose intolerance prevalence is higher in the north and decrease gradually in the center and the south [7].

There are several hypothesis at the basis of the evolutionary reason why this haplotype has been handed down over the centuries. One of these hypothesis is related to vitamin D, which is necessary for calcium absorption. Vitamin D is contained in milk. However, ultraviolet radiation penetrating into the epidermis can trigger vitamin D synthesis. This factor have led North European populations to be more dependent on dietary assumption of milk rather than South Europeans that are more exposed to sunlight, favoring the lactose persistence haplotype in the North [5].

There are several diagnostics tools for identification of Primary Hypolactasia. The most common one is Lactose Hydrogen Breath Test (LHBT) that consists on measuring hydrogen in the patient's breathes at regular intervals after lactose assumption [8]. However, this exam can give false negative results related, for example, to antibiotic treatments or false positive results due to a surplus of intestinal bacterial flora.

Another possibility are genetic tests that identify the genetic predisposition to lactose intolerance [9]. Genetic predispositions rely on genetic factors that increase the chance of developing a particular disease throughout life. In particular, the presence of one or more gene polymorphisms coupled with predisposing environmental conditions and lifestyle increase the risk of the occurrence of a certain phenotype. Unlike monogenic disorder that are caused entirely by genetic factors, such as cystic fibrosis or alfa1-antitripsina deficit, in multifactorial disease, the disorder is the result of gene and environmental interaction. In fact diet, smoke, stress and others environmental factors play an important role in genetic susceptibility.

The presence of a particular SNP upstream the lactase gene 1 locus, -13910C>T, has been identified as an indication of genetic predisposition for Primary Hypolactasia. This SNP is located on minichromosome maintenance complex component 6 (MCM6) gene, which is

adjacent to the lactase (LCT) gene [10]. MCM6 gene regulates the expression of LCT gene and causes the downregulation of lactase enzyme. The presence of the CC genotype in homozygosis is associated to Primary Hypolactasia. The heterozygosis variant CT is not strictly linked to the development of lactase non-persistence trait but it can predispose individuals to lactose malabsorption during periods of stress. Finally, the homozygosis variant TT is associated with Lactose Persistence phenotype [11]. In Italy about 30% of individuals present CC genotype, 60% is heterozygous CT while only 10% has the TT genotype [12]. However, the occurrence of lactose intolerance symptoms depends both on genetic traits and other factors such as, in this case, the dose of lactose assumed and the intestinal flora [5]. As a consequence, CC individuals have to pay attention to manifestation of intolerance related symptoms and adapt their diet to prevent health concerns [13].

Predictive genetic assays are commonly performed by specialized laboratories with a turnaround time for the analysis of about 7-10 days. Samples are collected through a blood draw or a mouth swab and the DNA is extracted from blood cells or from cells of mouth epithelium.

An interesting possibility consist of portable predictive tests that could be potentially available in decentralized settings. Especially, a portable diagnostic tool could be at the disposal of Nutritionists to perform predictive genetic tests directly during specialist medical examinations. This could allow to intervene at an early stage introducing dietary changes in response to the genetic profile.

In this thesis work, the rapid LAMP-based diagnostic strategy illustrated in chapter 2 and 3 was adapted to the detection of -13910C>T SNP. The strategy was designed to be performed in a portable heater device able to read and record a fluorescent signal. This type of instrument is commercially available and it is significantly cheaper in comparison with the Real-Time PCR, which owes its complexity into the ability to carry out thermal cycles (Figure 1). For LAMP applications, in fact, it is only required to reach and maintain one temperature.

The strategy was designed with the aim to develop a non-invasive POC test. In fact, the starting material for performing the assay is a saliva sample, from which cells of the mouth epithelium are collected. Saliva is an easy collectible biological sample and its employment in diagnostic tests permits to increase patient compliance. Moreover, saliva can be collected in

a simple non-treated container. The use of such simpler instrumentation coupled with a non-invasive approach, open the possibility to bring out predictive screenings from specialized settings to decentralized ones.

Figure 1. The ESEquant TS2 instrument for measurement of fluorescence by Qiagen.

4.2 Results and discussion

The optimized SNP-sensitive LAMP-based strategy resulted highly versatile and easily adaptable to other model cases involving a SNP-based discrimination.

Indeed, the single-SNP discriminating method was further applied to a relevant case in the field of nutrigenetics: the diagnosis of the genetic predisposition to Primary Hypolactasia.

Primary Hypolactasia is the form of lactose intolerance caused by a reduction of the production of lactase enzyme after infancy. However, some adults keep the ability to produce the enzyme throughout life. This trait is called Lactase-Persistence and it is associated to the presence of a SNP located on the MCM6 gene, called -13910C>T.

In particular, the allele type (T) is associated with lactase persistence: individuals presenting this genetic variation (T) preserve the ability to digest lactose in the adulthood. On the contrary, the allele type (C) is associated with the intolerant phenotypes that cause health problems related to the lactose malabsorption.

In this project, the genetic test for the predisposition to lactose intolerance was designed combining the SNP-sensitive LAMP-based strategy with a non-invasive sampling, using saliva specimen as clinical sample. In fact, DNA was extracted from buccal epithelium cells that are present in saliva. With the aim to develop a POC diagnostic platform, it was excluded to introduce complex DNA extraction methods (for example with magnetic beads or purification columns) in the test procedure. The extractive procedure had to be simple, rapid, and easy handled to be suitable for a portable approach. For this purpose, the extractive phase described in chapter 2 was optimized as following. The salivary sample was centrifuged using a low-cost benchtop minicentrifuge. The obtained pellet was undergone to a thermochemical lysis, and no further purification steps were performed.

In the experimental investigation of the method efficiency, the extracted DNA was quantified through a UV-visible spectrophotometer and diluted at the final concentration of 200 copies/µL. 5 µL of the diluted DNA were added to the LAMP reaction, without further

purification steps. The use of a raw DNA was feasible thanks to the intrinsic characteristics of LAMP that has a high tolerance to contaminants and matrix components [14].

Overall, the here-used optimized protocol was faster in comparison with the extraction procedure used in the chapter 2, requiring only 15-20 minutes, and it can be performed by non-trained users.

A future perspective for the optimization of the extractive phase of the test is to eliminate the dilution step of the extracted DNA. Optimizing the strategy in order to make it independent from the template concentration would allow to directly add the extracted template in the reaction mix, obtaining a completely portable assay, adaptable to decentralized settings.

The here-proposed SNP-sensitive LAMP strategy required the optimization of primers sequences and reaction conditions. As explained in chapter 1, the reactions were performed employing only one loop primer (LP), forward or backward on the basis of the target sequence around the SNP site. In this case, two primer versions of LP backward were designed accordingly to the SNP.

In this case, two Primers sets (set 1 and 2) were designed to be tested with the strategy. The two sets differed only for the discriminating LPs (LP1-C and LP1-T, LP2-C and LP2-T). For each set, LP-C was designed to detect the base that identify the genotype related to predisposition to lactose intolerance, while the LP-T was designed to be complementary to the polymorphic site. For each saliva sample, two reactions were performed in parallel, one pre-adding LP-C and the other one pre-adding LP-T, to ensure that the same target was examined with both probe types.

In presence of a homozygous CC sample, the amplification reaction operated with LP-C will be accelerated, while the simultaneous reaction performed with LP-T will be delayed. The opposite reactions behavior will occur in presence of a homozygous TT sample. On the

contrary, since heterozygous CT individuals are characterized by the presence of one copy for each allele type, both reaction will be enhanced, generating two overlapped curves.

Since the objective for the presented application is to develop a fluorescent–based detection, in this case amplification reactions were performed in Real-Time following the fluorescent emission of SYBR Green I that occurred as LAMP amplicons were synthetized.

The aim of the strategy optimization phase was to achieve an unequivocal readout of test results, obtaining:

- the maximum time gap possible between the fast and the slow reaction curves in presence of a homozygous sample; for this purpose, the two LPs must be able to accelerate the LAMP amplification only in presence of 100% match, and a relevant delay in presence of a single primer/target mismatch.
- The perfect overlap of the two curves in presence of a heterozygous target; for this purpose, in presence of both allele types, the two LPs must show the same efficiency in amplification acceleration.
- The fastest time of onset of curves, to achieve a rapid detection test.

For these purposes, the 2 primers sets were tested and compared. As shown in Figure 1, the primer set 1 operated a sufficient discrimination for homozygous samples (CC), however is was not able to provide the overlap of the two curves in the heterozygous case (TC), giving misleading results .

Therefore, the primer design was optimized with the set 2, in terms of the SNP binding site on the probe and melting temperature value for both primers, in order to reproduce the same time gap when analyzing a homozygous sample and to achieve a similar amplification rate for the both reactions in case of a heterozygous sample. In particular, primer set 2 was designed as follow: the SNP binding site on the probe was shifted toward the center of the primer sequence and the melting temperature difference between the LP2-C and LP2-T was optimized.

In this way, a greater improvement was obtained in managing heterozygous samples and an adequate time gap was achieved in case of a homozygous sample (Figure 2). Moreover, in

case of homozygous sample, the onset of the accelerated amplifications was more reproducible in comparison with the results obtained with primer set 1.

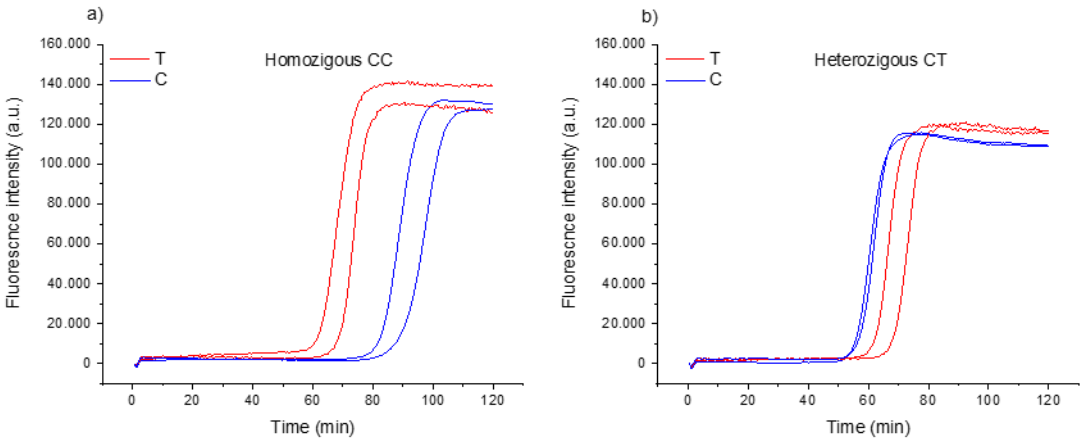

Figure 1. Real-time fluorescence detection of homozygous CC (a) and heterozygous CT samples (b). LAMP reactions were performed with primer set 1. The same sample was tested twice and for each sample two parallel reactions using LP1-C and LP1-T were performed. The homozygous sample is characterized by the acceleration of the reactions performed with LP1-C, while the heterozygous sample, since both allele copies are present, is distinguished by the enhancement of both reactions.

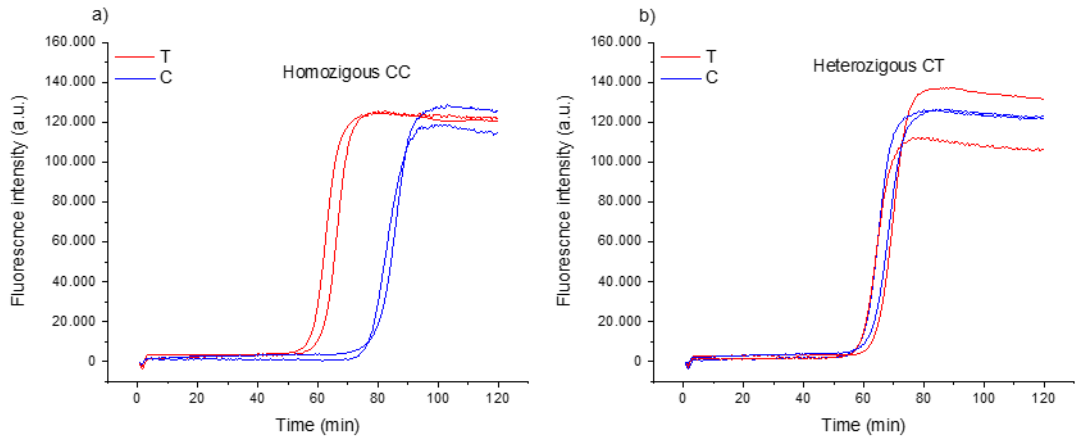

Figure 2. Real-time fluorescence detection of homozygous CC (a) and heterozygous CT samples (b). LAMP reactions were performed with primer set 2. The same sample was tested twice and for each sample two parallel reactions using LP2-C and LP2-T were performed. The homozygous sample is characterized by the acceleration of the reactions performed with LP2-C, while the heterozygous sample, since both allele copies are present, is distinguished by the enhancement of both reactions.

Therefore, the primer set 2 was selected and, giving the results of an high numbers of tested samples, further optimization to reaction conditions were introduced.

The objective was to obtain reliable and reproducible results that were independent from samples' variability. In particular, reaction conditions were slightly modified, modulating the $MgSO_4$ concentration and the enzyme concentration.

Finally, the best result in terms of discrimination efficiency was obtained (Figure 3).

In this case, a heterozygous and a homozygous CC samples were amplified with the primer set 2 with the optimized reaction conditions and at 61°C of reaction temperature. A control sample without the pre-addition of any LP was introduced. As reported in Figure 3a, in case of a heterozygous sample, both reactions were correctly enhanced with the same rate, indicating the presence of both allele type in the target sample. In case of a homozygous individual, the reaction where the LP perfectly matched to the nucleotide associated to lactose intolerance is significantly accelerated, indicating the genotype of the investigated sample (Figure 3b). Interesting the behavior of the reaction performed without including any LP is comparable to the one of the slow reaction. These results underlined the substantial effect played by LPs. In fact, a single mismatch located on a short probe can nullify the LP action (Figure 3).

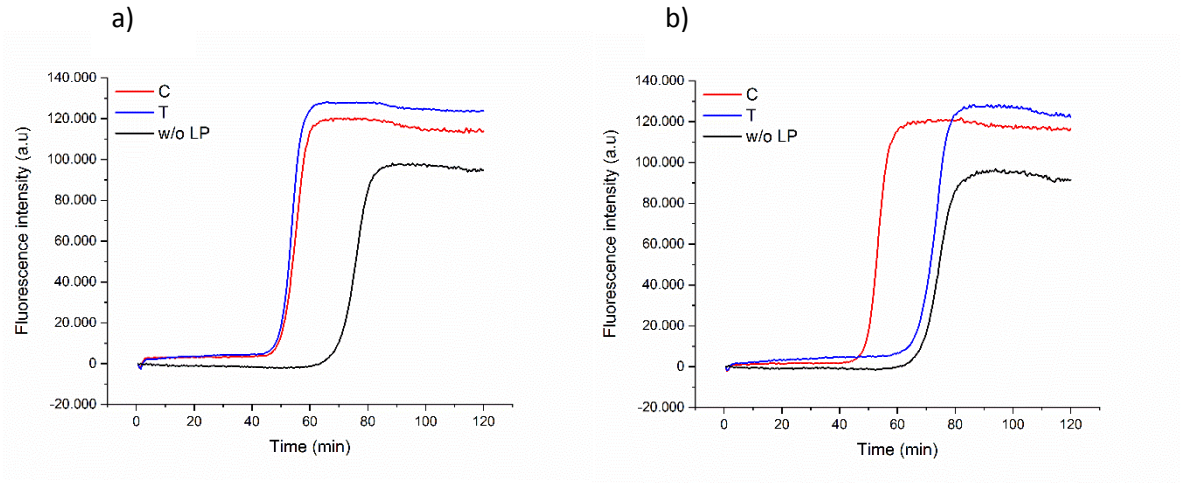

Figure 3. Real-time fluorescence detection of heterozygous CT (a) and homozygous CC samples (b). LAMP reactions were performed with primer set 2 composed by LP2-C (red line) and LP2-T (blue line). Amplification reaction of both heterozygous and homozygous samples without the pre-addition of any LP was performed (black line). The time gap between fast and slow reactions was about 15 minutes.

To confirm these results, 10 people were tested using the here-presented SNP genotyping assay. The 10 people to test were chosen randomly and their characteristics (e.g. sex, age, apparent lactose intolerance) were listed in Table 1. Observing the fluorescent based detection outcomes, the results were clear and can easily be interpreted. In case of a homozygous sample the two curves are sufficiently separated with a time gap ranging from 12.5 and 15 minutes, while in case of heterozygous samples the two curves perfectly overlapped. Moreover, with this approach a reliable SNP detection is achieved within 1 hour, providing a significant advantage in comparison with the standard techniques (Figure 4).

Patient	**Sex**	**Age**	**Lactose intolerance symptoms**
1	Female	28	No
2	Female	36	No
3	Female	33	No
4	Male	36	Yes
5	Female	36	No
6	Female	29	No
7	Female	26	Yes
8	Female	31	No
9	Female	35	No
10	Male	30	Yes

Table 1. In this table are given more details about the 10 people genotyped for the -13910C>T SNP.

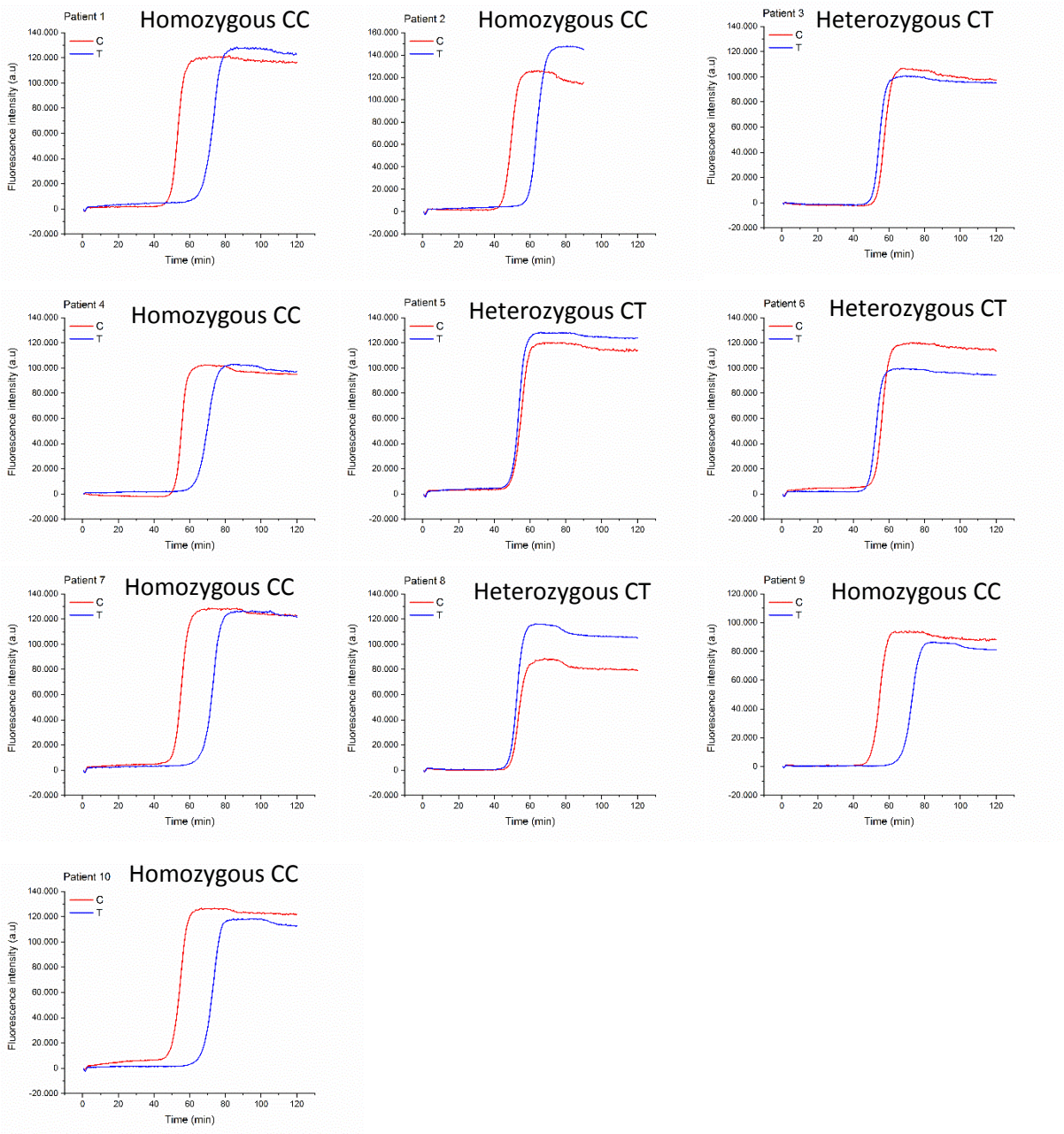

Figure 4. Real-time fluorescence genotyping of lactose intolerance predisposition for 10 people. LAMP reactions were performed with primer set 2 composed by LP2-C (red line) and LP2-T (blue line). Heterozygous CT samples were detected by the correct enhancement of both reactions, while homozygous CC samples were identified by the acceleration of the reaction performed with LP2-C.

To validate the efficiency of the method, the test results were compared with the ones provided by a certified testing laboratory, Lifebrain.

To perform the Lifebrain test, the test must be purchased online; a buccal swab is delivered at patient home. The patient self-collects the sample and sends back it to the specific laboratory where the genetic test is performed. The test results are provided with a turnaround time of 7 days.

Interestingly, the results obtained with the optimized SNP assay, compared with results listed in the report from the testing laboratory, were validated with 100% of concordance (Figure 5). Since homozygous TT individuals are poorly represented in the general population, among 10 patients tested, no one showed this particular polymorphic genotype (Figure 4 and 5).

	life brain Eccellenza per la tua salute	**SNP-sensitive LAMP-based test**
Patient 1	Detected alleles: C	Detected alleles: C
Patient 2	Detected alleles: C	Detected alleles: C
Patient 3	Detected alleles: C and T	Detected alleles: C and T
Patient 4	Detected alleles: C	Detected alleles: C
Patient 5	Detected alleles: C and T	Detected alleles: C and T
Patient 6	Detected alleles: C and T	Detected alleles: C and T
Patient 7	Detected alleles: C	Detected alleles: C
Patient 8	Detected alleles: C and T	Detected alleles: C and T
Patient 9	Detected alleles: C	Detected alleles: C
Patient 10	Detected alleles: C	Detected alleles: C

Figure 5. Comparison of Genotyping of 10 individuals for -13910C>T genetic variant with a laboratory analysis test performed by Lifebrain and with the SNP-sensitive LAMP-based test. The results showed 100% of concordance.

4.3 Materials and methods

Optimized protocol for DNA extraction from saliva samples

DNA was extracted from saliva samples using the following protocol. 500 µL of saliva were collected in a safe lock Eppendorf tube and mixed with 1 mL of Phosphate Saline Buffer (PBS). Samples were then centrifuged at 6000 RPM for 10 seconds using Mini Microcentrifuge (Labnet International, Edison, New Jersey, USA) and the resulting supernatant was discarded. The pellet was suspended in 200 µL of Pretreatment Buffer and was mixed at 500 RPM for 10 minutes at 95°C using a ThermoMixer (Eppendorf, Milano, Italy). After this step, samples were mixed with using a Vortex Mixer (VWR International SRL, Milano, Italy) and then centrifuged for 1 minutes at 6000 RPM using Mini Microcentrifuge (Labnet International, Edison, New Jersey, USA). After centrifuge a pellet, composed by cells debris and proteins, resulted separated from the surnatant that contains the extracted ssDNA. Finally, sample concentration was assessed as described in Chapter 2 and it was diluted in MilliQ water at the final concentration of 200 copies/µL.

Primer design for LAMP strategy

Primer sequences for the predictive test for lactose intolerance were designed by Dr.ssa Paola Cecere of NanoBioInteractions&NanoDiagnostics Group starting from the MCM6 human gene.

Fluorescent-based LAMP reactions for predictive test for the detection of genetic predisposition to lactose intolerance

LAMP reactions were performed in 25 µL of a mixture containing 1.6 µM each of inner primers, 0.4 µM each of outer primers, 0.8 µM of loop primers (Integrated DNA Technologies, Coralville, IA, USA), 1 M of betaine (VWR International SRL, Milano, Italy), 2.5 µL of 10X LAMP buffer (200 mM of Tris-HCl, 100 mM of $(NH_4)2SO_4$, 20 mM of $MgSO_4$, 500 mM of KCl, and 1% v/v Tween 20), 2 mM of $MgSO_4$, 0.8 mM each of dNTPs (Promega, Madison, WI, USA), 0.80 U/µL of Bst 2.0 WarmStart DNA Polymerase (New England BioLabs, Ipswich, MA, USA), and 5 µL of DNA template at the concentration of 200 copies/µL of the extracted DNA. DNA-free LAMP reactions were included as negative controls. Amplification reactions were performed using forward and backward inner primers (FIP and BIP), forward

and backward outer primers (F3 and B3), and one loop primer (LP-C or LP-T). For real-time fluorescent LAMP, 1/50,000 diluted SYBR Green (Thermo Fisher Scientific,Waltham, MA, USA) was pre-added to the reaction mix. Real-time amplifications were performed on an Applied Biosystem real-time instrument (Thermo Fisher Scientific,Waltham, MA, USA) with StepOne Software v2.3 at 61 °C.

4.4 Conclusions

An effective fluorescent-based SNP discriminating tool was developed and applied to an example of a clinically relevant nutrigenetic case, in particular, the genetic predisposition to lactose intolerance. The assay resulted effective in the discrimination of both homozygous and heterozygous patients starting from simple saliva samples. The test results were confirmed by the current reference assay performed by a specialized laboratory. In addition, a significant improvement of the turnaround time of the analysis was achieved. In fact, with the described optimized strategy a reliable results was obtained within 1 hour, while the ordinary lab test requires at least 7 days.

Overall these findings opened the perspective for the development of a rapid and portable tool for identification of genetic-based predisposition to certain traits that could be associated to the intolerance to particular nutrients. This fast, low cost and easy-to-use assay, would turn genetic screening into analysis accessible to everyone.

Moreover, such predictive tests would allow a rapid screening of patient's genome, furnishing to the Nutritionist the information to recommend the appropriate diet to the patients.

On the contrary, in case of a patient presenting gastrointestinal disorders and a genetic profile that does not suggest lactose malabsorption problems, Primary Hypolactasia could be fast excluded among the symptoms' causes.

4.5 Bibliography

1. Kuokkanen, M. *et al.* Mutations in the translated region of the lactase gene (LCT) underlie congenital lactase deficiency. *Am. J. Hum. Genet.* **78**, 339–344 (2006).

2. Antonowicz, I. & Lebenthal, E. Developmental Pattern of Small Intestinal Enterokinase and Disaccharidase Activities in the Human Fetus. *Gastroenterology* **72**, 1299–1303 (1977).

3. Heyman, M. B. Lactose intolerance in infants, children, and adolescents. *Pediatrics* **118**, 1279–1286 (2006).

4. Vesa, T. H., Korpela, R. & Marteau, P. Lactose Intolerance. *J. Am. Coll. Nutr.* **19**, 165S-175S (2000).

5. Ugidos-Rodríguez, S., Matallana-González, M. C. & Sánchez-Mata, M. C. Lactose malabsorption and intolerance: a review. *Food and Function* **9**, 4056–4068 (2018).

6. Sahi, T. Genetics and epidemiology of adult-type hypolactasia. *Scand. J. Gastroenterol.* **29**, 7–20 (1994).

7. Cavalli-Sforza, L. T., Strata, A., Barone, A. & Cucurachi, L. Primary adult lactose malabsorption in Italy: Regional differences in prevalence and relationship to lactose intolerance and milk consumption. *Am. J. Clin. Nutr.* **45**, 748–754 (1987).

8. Argnani, F. *et al.* Hydrogen breath test for the diagnosis of lactose intolerance, is the routine sugar load the best one? *World J. Gastroenterol.* **14**, 6204–6207 (2008).

9. Coluccia, E. *et al.* Congruency of genetic predisposition to lactase persistence and lactose breath test. *Nutrients* **11**, (2019).

10. Enattah, N. S. *et al.* Identification of a variant associated with adult-type hypolactasia. *Nat. Genet.* **30**, 233–237 (2002).

11. Kuokkanen, M. *et al.* Transcriptional regulation of the lactase-phlorizin hydrolase gene by polymorphisms associated with adult-type hypolactasia. *Gut* **52**, 647–652 (2003).

12. L'intolleranza al lattosio - CDI Centro Diagnostico Italiano. Available at:

https://www.cdi.it/approfondimento/lintolleranza-al-lattosio/. (Accessed: 2nd December 2020)

13. Obermayer-Pietsch, B. M. *et al.* Genetic predisposition for adult lactose intolerance and relation to diet, bone density, and bone fractures. *J. Bone Miner. Res.* **19**, 42–47 (2004).

14. Kaneko, H., Kawana, T., Fukushima, E. & Suzutani, T. Tolerance of loop-mediated isothermal amplification to a culture medium and biological substances. *J Biochem Biophys Methods* **70**, 499–501 (2007).

Chapter 5

A fast, fluorescent-based pharmacogenetic assay for the detection of the 677C>T polymorphism on MTHFR gene

5.1 Introduction

Folate is the natural form of vitamin B9. As it is an essential nutrient, folate is not synthetized by human organism and have to be taken with diet. Several dietary sources are rich in folate, such as green leafy vegetables, beans, grains, eggs, pork and beef meats [1]. Natural dietary folates are water-soluble and thermolabile.

Folic acid is the synthetic form of vitamin B9, used as supplements and present in fortified foods. Folic acid is stable, more resistant to cooking procedures rather than folates and cheaper to produce (Figure 1).

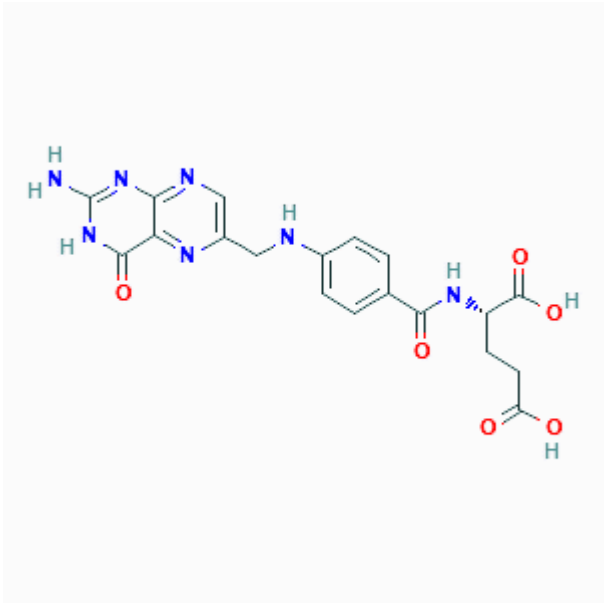

Figure 1. Chemical structure of Folic Acid [2].

Folic acid undergoes to two modifications catalyzed by the Dihydrofolate Reductase (DHFR) enzyme prior to enter into the folate cycle. Finally, it has to be transformed into its active form, the 5-methyl-THF. 5-methyl-THF is then directed towards the biosynthesis of methionine, purine and pyrimidine (Figure 2) [3, 4].

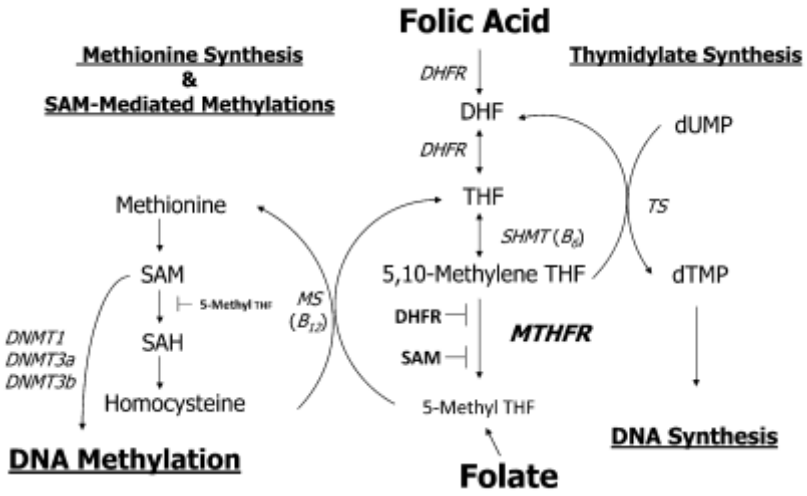

Figure 2. Folic acid metabolic pathway [4].

Since folate is required for the biosynthesis of nucleotides and proteins and for the methylation process, it is essential in all the phases of fetal growth, such as implantation and organ development. Hence, folate supplementation during pregnancy is strongly recommended to reduce the risk of congenital abnormalities. In fact, the process for formation of central nervous system during embryonic development, the neurulation, can be impaired by a folic acid deficit that can lead to fetus malformation, such as Neural Tube Defects (NTD) [1]. Moreover, the folate deficit can cause a high level of Homocysteine (Hcy) in the bloodstream, that is as well a risk factor for occurrence of NTD in embryos, recurrence of pregnancy lost and other pregnancy complications [5, 6]. In fact, Hcy acts with a negative feedback mechanism limiting the physiological cells methylation process. In particular, DNA

methylation process is at the base of epigenetic modification of the DNA, influencing gene expression, which has a pivotal role in embryogenesis [7, 4].

Foods fortified with folic acid has been introduced in several countries with the purpose to reduce the birth defects. In United States and Canada, for example, the fortification of food with folic acid is mandatory, while in Italy it is strongly recommended. As demonstration of the importance of folic acid supplementation, the consume of enriched products in United Stated and Canada has led to a significant decrease of the number of newborns with neural tube defects [8].

The methylenetetrahydrofolate reductase (MTHFR) enzyme is involved in the folate metabolic pathway. In particular, MTHFR enzyme catalyzes the reduction of 5,10-methylenetetrahydrofolate (5,10-methylene-THF) to 5-methyltetrahydrofolate (5-methyl-THF). MTHFR straddles between two metabolic pathways: folate cycle, which is committed to purine and pyrimidine synthesis, and methionine cycle, that is aimed at protein production (Figure 3). 5-methyl-THF and Hcy, through the action of the enzyme Methionine Synthase and using cobalamin (vitamin B12) as a cofactor, are transformed into Methionine, that is the starting step for protein synthesis. Methionine through S-adenosyl Homocysteine (SAM) is involved in methylation processes that are fundamental for several cell physiology processes, especially, for epigenesist and imprinting stages of embryo development (Figure 2 and 3).

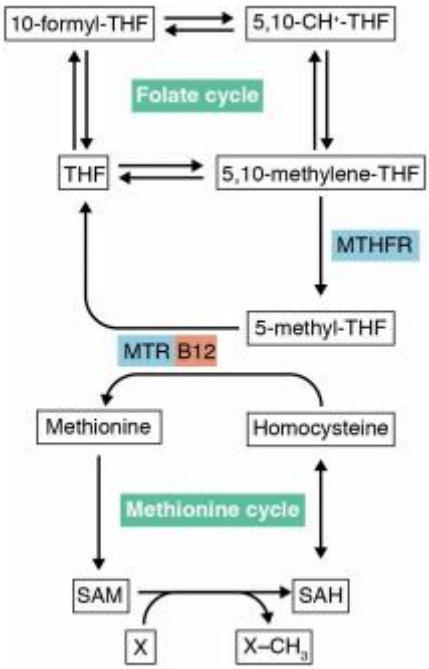

Figure 3. The Folate cycle and the Methionine cycle [3] .

Recently, a particular SNP located on MTHFR gene, the 677C>T genetic variation, has been identified. The most frequent allele presents the nucleotide C at the 677 position in the MTHFR gene, while the polymorphic allele at the same position has the nucleotide T [7].

This gene variant encodes for a thermolabile variant of the enzyme characterized by a reduced enzymatic activity [9]. In particular, in comparison with the wild type form CC, which is perfectly functional, the polymorphism in homozygosis TT leads to the production of an enzyme with a remaining activity of about 30% and the heterozygosis haplotype CT generates an enzyme with an activity of 65% [10].

The presence of 677C>T genetic variant of MTHFR enzyme has a crucial importance for pregnancy concerns and congenital defects. Indeed, malfunctioning of MTHFR enzyme causes a reduction of the availability of the folate active compound, causing high plasma Hcy and a reduction of methionine synthesis, which are risk factors for pregnancy complications (Figure 3) [7, 11, 12, 13].

Several studies reported that the folate concentration in the serum of women that present 677C>T variant in homozygosis is lower when compared with women with the TC and CC genotypes even after folic acid supplementation [14]. On the contrary, it has been reported in literature that homozygous CC subjects have a greater bloodstream concentration of 5-methyl-THF, the active form of folic acid, in comparison with homozygous TT ones [15].

The prevalence of 677C>T polymorphism in homozygosis ranges in Europe from the 8% in Germany to the 21% in Hispanic population. In Italy the homozygosis TT genotype frequency ranges from 18 to 25 % [11, 16, 17].

The Recommended Dietary Allowance (RDA) of folate for lactating and pregnant women is 600 µg/d [18], while women with MTHFR polymorphism are advised to assume up to 800 µg/d [19].

Folic acid, as well as all the other vitamins, should be taken in the right balance in order to avoid deficit and surplus that could be both harmful. In fact, excess of folate intake can lead to the development of the unmetabolized folic acid (UMFA) syndrome. At present, this issue is being debated, however, several health risks are associate to UMFA. In particular, it has been reported that UMFA can cause dysfunction of immune system. Additionally, it seems to play a dual role in carcinogenesis reducing cancer risk and promoting the growth of existing neoplastic lesions [20, 21, 22]. In literature it is also reported that UMFA can lead to a pseudo MTHFR deficiency and to the downregulation of folate transporters presents in human intestinal cells. [23, 24] . Finally extra folic acid supplementation may also mask vitamin B12 deficiency. People with a deficit of vitamin B12 and a surplus of folic acid never develop megaloblastic anemia, which is the main symptom of the lack of vitamin B12. Folic acid supplementation, in details, prevents the insurgence of vitamin B9 deficiency related symptoms.

Due to all of these reasons, even if supplementation with folic acid in women with the 677C>T polymorphism is not sufficient to increase the plasma level of the metabolic active compound, an overdose of unmetabolized folic acid could be damaging.

Thereby, the alternative supplementation consists of the reduced and methylated form of folates, 5-methyl-THF, that permits to bypass the MTHRF catalyzed step. This form can replace the classical folic acid supplementation in women with the polymorphism. Supplementation

with 5-methyl-THF provides several advantages for women who carry the genetic variation and it has been demonstrated that the administration of the active compound revealed efficient in increasing the plasma concentration of 5-methyl-THF [25, 26].

Overall these findings seriously highlight the pharmacogenetic issue related to folic acid assumption. For these reasons, supplementation with a standard dose which is independent from patient genetic profile must be avoided. Information about the mother genotype should became a routine tests in order to select the right folate supplementation in accordance to the level of activity of the MTHFR enzyme.

The genetic identification of this polymorphism can be actually performed in a specialized laboratory through a blood test. DNA is extracted from blood cells, and the identification of the genetic variant is performed with standard techniques as the genome sequencing. The result is provided with a turnaround time of 7-10 days. Due to the complexity of the test it is not part of the routine analysis. A valid alternative could be a rapid and easy-to-use POC test providing genetic pre-screening directly during the Gynecologist's visit. In this way the traditional administration of a high dose of folic acid can be routinely avoided benefiting the alternative supplement compound.

In this thesis a Real-Time strategy for 677C>T polymorphism on MTHFR gene suitable for portable testing was developed. In contrast to the currently available analysis, this application was designed to be performed in a non-invasive way starting from saliva samples. The strategy, as well as for the application described in chapter 4, is suitable to be performed in a portable heater fluorimeter. This type of approach would make possible the choice of the correct supplementation in response to the rapid result of the genetic test.

5.2 Results and discussion

The optimized single-SNP genotyping LAMP-based strategy, thanks to its highly versatility was applied to a third model case addressing a pharmacogenetic issue. In this case, the strategy was declined to obtain the detection of the presence of 677C>T polymorphism on MTHFR gene. The presence of this SNP is associated with a reduced activity of the MTHFR enzyme that is involved in folic acid metabolism. The polymorphism in homozygosis TT generates a form of MTHFR enzyme with a reduced activity. In particular, the homozygosis TT genotype leads to the production of a form with the 30% of the full enzymatic activity, while the heterozygosis CT genotype encodes for an enzyme with a remaining activity of 65%. Genotyping in a simple and rapid way this genetic variation has various advantages in clinical practice. In particular, it would be useful for choosing the right folic acid supplementation for women during pregnancy. In presence of the genetic variation (T) the supplementation with the active form of folic acid could be suggested directly by the doctor after a rapid analysis of the patient's genome.

The purpose of this project was to develop a fluorescent-based strategy suitable for a POC approach. Thus, in order to introduce a non-invasive sampling, DNA was extracted from saliva samples with the easy-handle protocol described in chapter 4. Loop primers (LPs) for this application were designed in order to accomplish the identification of 677C>T SNP. Herein, the polymorphic base was detected by the LP-T, while the most frequent allele is identified by LP-C. Two reactions were performed in parallel to ensure the correct genotyping of both the allele types.

For homozygous CC or TT samples, only one amplification reaction will be accelerated depending on the allele type present in the sample, while the concurrent reaction performed with the other loop primer will be delayed. For example, in case of a homozygous CC, the accelerated reaction will be the one performed with LP-C, the opposite behavior will occur in presence of homozygous TT individuals. In presence of heterozygous CT individuals both reaction will be properly enhanced, generating two overlapped curves gave by the same amplification efficiency.

To achieve a fluorescent detection it is important that the amplification curves obtained from the analysis of a heterozygous samples were perfectly overlapped and the ones related to the

identification of a homozygous sample were sufficiently separated (at least 10 minutes) in order to obtain unequivocal results.

For these purposes, 2 primer sets were tested (LP1-C and LP1-T, LP2-C and LP2-T).

LP hybridizes to the portion of the target sequence where the SNP is located. The primer set 1 was proven to operate a sufficient discrimination for homozygous samples. However, the curves resulted from the amplification of a heterozygous sample had different amplification rates, providing an uncertain interpretation of the test outcomes (Figure 1).

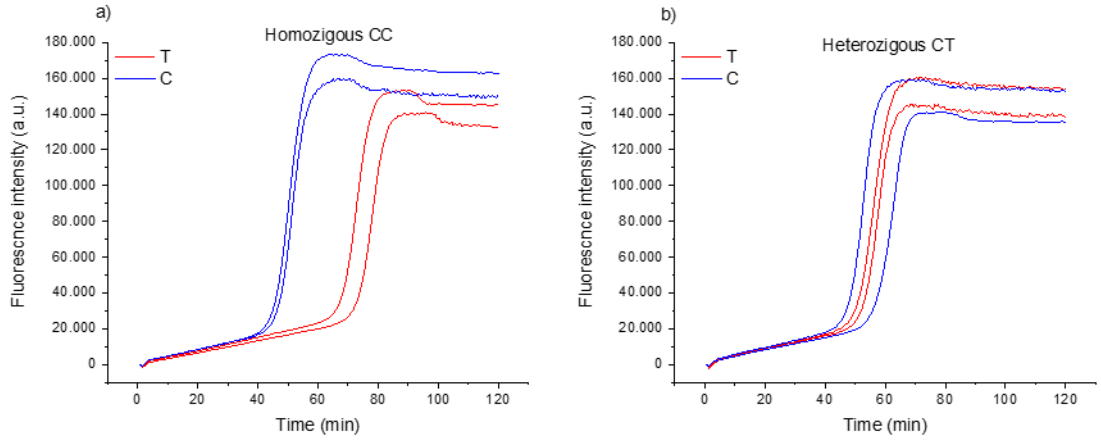

Figure 1. Real-time fluorescence detection of homozygous CC (a) and heterozygous CT samples (b). LAMP reactions were performed with primer set 1 at the reaction temperature of 63°C. The same sample was tested twice and for each sample two parallel reactions using LP1-C and LP1-T were performed. The accelerated reactions in case of a homozygous sample were performed with LP1-C, while in case of the heterozygous sample, both reactions were accelerated.

This result was explained by the fact that LP of set 1 was able to form heterodimers with the sequence of others primers involved in the reaction. This unintended bond led to reduce the amount of LP in the reaction mix, reducing the availability of the discriminating probes. Unfortunately, since in this kind of strategy the primer design is constrained by the SNP site, it was not possible to operate a substantial modification on the sequence of the primers. Therefore, the subsequent optimization was consisted on raising the reaction temperature of 2 °C. With a higher temperature the primer annealing is disfavored. This led to a further disfavoring of non-specific reactions resulting in a larger time gap between the two

amplification curves in homozygous samples. However, on the other side, at a higher reaction temperature, the specific annealing is more random. As a result, the time of onset of all the specific reactions, resulting from both homozygous and heterozygous samples, was less reproducible and slightly delayed (Figure 2).

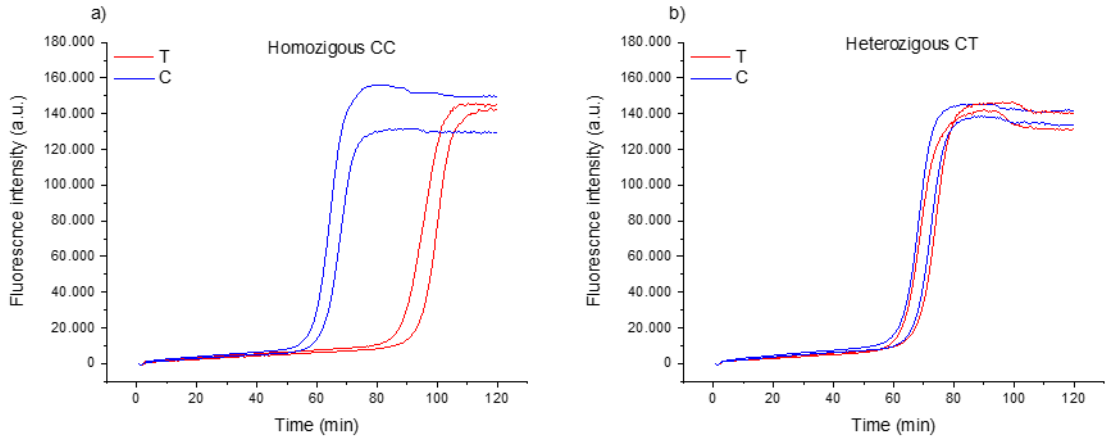

Figure 2. Real-time fluorescence detection of homozygous CC (a) and heterozygous CT samples (b). LAMP reactions were performed with primer set 1 at the reaction temperature of 65°C. The same sample was tested twice and for each sample two parallel reactions using LP1-C and LP1-T were performed. The accelerated reactions in case of a homozygous sample were performed with LP1-C, while in case of the heterozygous sample, both reactions were accelerated.

To increase the result reproducibility, another primer set, set 2, was designed, increasing the melting temperature of both primers (LP2-C and LP2-T). Primer set 2 was tested at 65°C and the best result was obtained. This optimization led to achieve a compromise between the amplification efficiency of 100% matched reactions and the annealing destabilization of the single nucleotide mismatched ones (Figure 3).

All the possible cases were tested, introducing a control sample performed without the pre-addition on any LP type. As showed in Figure 3a, in case of a wild-type homozygous sample CC, the accelerated reaction resulted the one performed with LP2-C, on the contrary with the same target the reaction performed with LP2-T was sufficiently delayed. Then, in case of a mutant homozygous sample TT, reactions behavior was inverted: the LP-T mediated reaction resulted accelerated, while the one mediated by LP-C was delayed (Figure 3b). In both cases, the time gap between the two curves distinctly allowed the discrimination.

Finally, in case of a heterozygous sample, both reactions were enhanced with the same speed (Figure 3c). In all cases, for all the tested cases, reactions performed without LPs did not occurred within 2 hours.

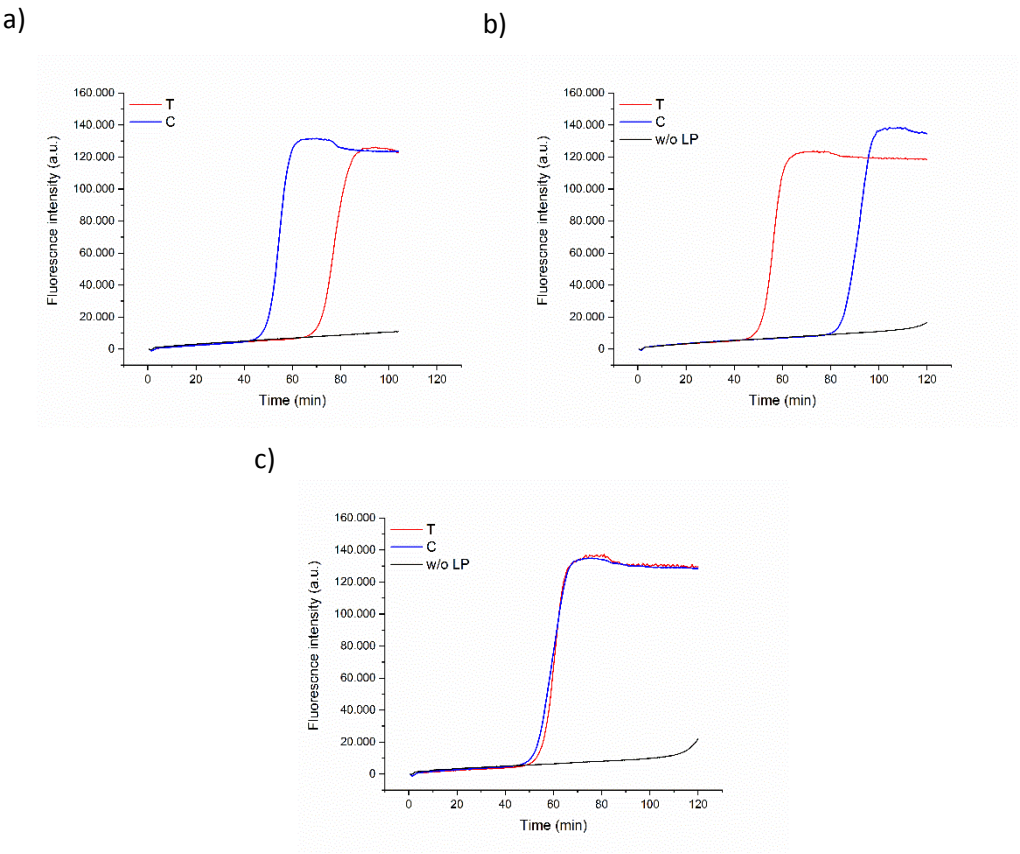

Figure 3. Real-time fluorescence detection of a homozygous CC (a), a homozygous TT (b) and a heterozygous CT (c) samples. LAMP reactions were performed with primer set 2 composed by LP2-C (red line) and LP2-T (blue line). Amplification reaction of both heterozygous and homozygous samples without the pre-addition of any LP was performed (black line). The time gap between the accelerated and the delayed reactions was 20 minutes in case of homozygous CC sample and 40 minutes in case of homozygous TT sample.

To confirm these results, 10 people were tested using the optimized SNP genotyping assay. The 10 people were the same volunteers tested for the -13910C>T SNP, except for patient 4 and 8. They were all chosen randomly and their characteristics (e.g. sex and age) were listed in Table 1.

Patient	Sex	Age
1	Female	28
2	Female	36
3	Female	33
4	Male	43
5	Female	36
6	Female	29
7	Female	26
8	Female	27
9	Female	35
10	Male	30

Table 1. In this table are given more details about the 10 people genotyped for the 677C>T SNP.

Observing the Real-Time results, in all of these cases, the two amplification curves were separated enough to distinguish the allele types present in the samples. In case of a homozygous sample the LP employed in the reaction that onset faster, indicated the sample genotype. In case of heterozygous samples the two curves amplified at the same time, indicating the present of both allele types in the sample. Also for this application the fluorescent-based detection generated a fast and unequivocal result for each of the tested individuals (Figure 4).

These outcomes were validated comparing the results obtained with the optimized SNP-sensitive strategy with the results provided by a certified testing laboratory, Lifebrain. Currently, Lifebrain offers the analysis of this SNP through a blood test. Since this analysis can't be performed through buccal swab there isn't the possibility to receive and perform the test at home. In fact, a blood draw has to be performed by a healthcare professional or a nurse and the sample is analyzed by the specialized laboratory. As well as for the predictive tests (chapter 4), the test results are provided with a turnaround time of 7 days. In this case,

both the collection of the sample and the genetic test have to be performed exclusively in a specialized laboratory.

Remarkably, comparing the results obtained with the optimized SNP assay with the results given by Lifebrain test, all the outcomes were validated with 100% of concordance (figure 4 and 5).

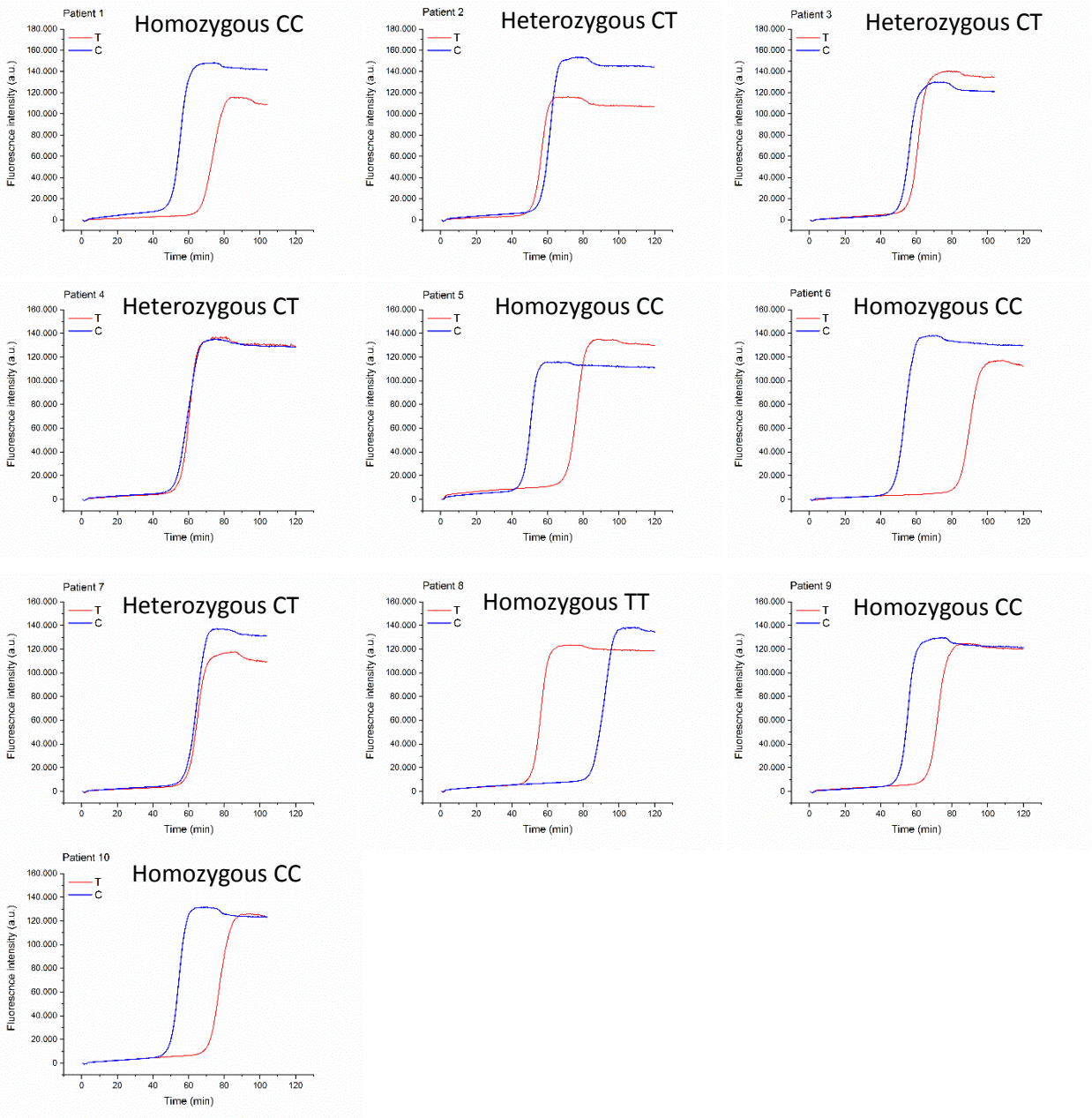

Figure 4. Real-time fluorescence genotyping of 10 people. LAMP reactions were performed with primer set 2 composed by LP2-C (red line) and LP2-T (blue line). Heterozygous samples were detected by the correct enhancement of both reactions, while homozygous CC or TT samples were identified by the acceleration of the reaction performed with LP corresponding to the allele type.

		SNP-sensitive LAMP-based test
Patient 1	Detected alleles: C	Detected alleles: C
Patient 2	Detected alleles: C and T	Detected alleles: C and T
Patient 3	Detected alleles: C and T	Detected alleles: C and T
Patient 4	Detected alleles: C and T	Detected alleles: C and T
Patient 5	Detected alleles: C	Detected alleles: C
Patient 6	Detected alleles: C	Detected alleles: C
Patient 7	Detected alleles: C and T	Detected alleles: C and T
Patient 8	Detected alleles: T	Detected alleles: T
Patient 9	Detected alleles: C	Detected alleles: C
Patient 10	Detected alleles: C	Detected alleles: C

Figure 5. Genotyping of 10 individuals for 677C>T genetic variant with a standard test performed by Lifebrain and with the optimized strategy based on the enhancement of the perfectly matched reaction. The results showed 100% of concordance.

5.3 Materials and methods

Optimized protocol for DNA Extraction from saliva samples

DNA was extracted from saliva samples using the following protocol. 500 µL of saliva were collected in a safe lock Eppendorf tube and mixed with 1 mL of Phosphate Saline Buffer (PBS). Samples were then centrifuged at 6000 RPM for 10 seconds using Mini Microcentrifuge (Labnet International, Edison, New Jersey, USA) and the resulting supernatant was discarded. The pellet was suspended in 200 µL of Saliva Pre-treatment Buffer and was mixed at 500 RPM for 10 minutes at 95°C using a ThermoMixer (Eppendorf, Milano, Italy). After this step, samples were mixed with using a Vortex Mixer (VWR International SRL, Milano, Italy) and then centrifuged for 1 minutes at 6000 RPM using Mini Microcentrifuge (Labnet International, Edison, New Jersey, USA). After centrifuge a pellet, composed by cells debris and proteins, resulted separated from the surnatant that contains the extracted ssDNA. Finally, sample concentration was assessed as described in Chapter 2 and it was diluted in MilliQ water at the final concentration of 200 copies/µL.

Primer Design for LAMP strategy

Primer sequences for the Folic Acid pharmacogenetic assay were designed by Dr.ssa Paola Cecere of NanoBioInteractions&NanoDiagnostics Group starting from the MTHFR human gene.

Fluorescent-based LAMP Reactions for folic acid pharmacogenetic assay

LAMP reactions were performed in 25 µL of a mixture containing 1.6 µM each of inner primers, 0.4 µM each of outer primers, 0.8 µM of loop primers (Integrated DNA Technologies, Coralville, IA, USA), 1 M of betaine (VWR International SRL, Milano, Italy), 2.5 µL of 10X LAMP buffer (200 mM of Tris-HCl, 100 mM of $(NH_4)2SO_4$, 20 mM of $MgSO_4$, 500 mM of KCl, and 1% v/v Tween 20), 2 mM of $MgSO_4$, 0.8 mM each of dNTPs (Promega, Madison, WI, USA), 0.112 U/µL of Bst 2.0 WarmStart DNA Polymerase (New England BioLabs, Ipswich, MA, USA), and 5 µL of DNA template at the concentration of 200 copies/µL of the extracted DNA. DNA-free LAMP reactions were included as negative controls. Amplification reactions were performed using forward and backward inner primers (FIP and BIP), forward and backward outer primers (F3 and B3), and one loop primer (LP-C or LP-T).

For real-time fluorescent LAMP, 1/50,000 diluted SYBR Green (Thermo Fisher Scientific,Waltham, MA, USA) was pre-added to the reaction mix. Real-time amplifications were performed on an Applied Biosystem real-time instrument (Thermo Fisher Scientific,Waltham, MA, USA) with StepOne Software v2.3 at 65 °C.

5.4 Conclusions

A reliable fluorescent-based approach for SNP genotyping was applied to a particular pharmacogenetic issue, in particular the detection of the 677C>T polymorphism on MTHFR gene. This genetic variation causes the production of an enzyme with a reduced activity, which entails a reduced metabolization of folic acid introduced as supplementation.

The detection resulted accurate and effective in distinguish all the possible genotypes. A non-invasive and easy handled protocol, using saliva samples, was developed and coupled with the fluorescent-based approach enabling to genotyping patients with a turnaround time of a 2 hours. Results were validated using the gold standard technique performed by a specialized laboratory and were confirmed with 100% of concordance. The employment of such assay will provide a greater advantage in comparison with the currently available techniques. Moreover, this approach would be more affordable and would enable the suggestion of a personalized therapy, based on the genetic profile. Overall these findings demonstrated that, by adjusting the primer design in accordance with an optimization of reaction conditions, every case of a single SNP associated with a specific trait of interest is potentially detectable.

5.5 Bibliography

1. Van Der Put, N. M. J., Van Straaten, H. W. M., Trijbels, F. J. M. & Blom, H. J. Folate, Homocysteine and Neural Tube Defects: An Overview. *Exp. Biol. Med.* **226**, 243–270 (2001).

2. Folic acid | C19H19N7O6 - PubChem. Available at: https://pubchem.ncbi.nlm.nih.gov/compound/Folic-acid. (Accessed: 12th December 2020)

3. Zheng, Y. & Cantley, L. C. Toward a better understanding of folate metabolism in health and disease. *Journal of Experimental Medicine* **216**, 253–266 (2019).

4. Crider, K. S., Yang, T. P., Berry, R. J. & Bailey, L. B. Folate and DNA methylation: A review of molecular mechanisms and the evidence for Folate's role. *Advances in Nutrition* **3**, 21–38 (2012).

5. Finnell, R. H. *et al.* Neural tube and craniofacial defects with special emphasis on folate pathway genes. *Critical Reviews in Oral Biology and Medicine* **9**, 38–53 (1998).

6. Ray, J. G. & Laskin, C. A. Folic acid and homocyst(e)ine metabolic defects and the risk of placental abruption, pre-eclampsia and spontaneous pregnancy loss: A systematic review. *Placenta* **20**, 519–529 (1999).

7. Ueland, P. M., Hustad, S., Schneede, J., Refsum, H. & Vollset, S. E. Biological and clinical implications of the MTHFR C677T polymorphism. *Trends in Pharmacological Sciences* **22**, 195–201 (2001).

8. Berry, R. J., Bailey, L., Mulinare, J. & Bower, C. Fortification of flour with folic acid. *Food Nutr. Bull.* **31**, (2010).

9. Kang, S. S., Passen, E. L., Ruggie, N., Wong, P. W. & Sora, H. Thermolabile defect of methylenetetrahydrofolate reductase in coronary artery disease. *Circulation* **88**, 1463–1469 (1993).

10. Frosst, P. *et al.* A candidate genetic risk factor for vascular disease: A common mutation in methylenetetrahydrofolate reductase. *Nat. Genet.* **10**, 111–113 (1995).

11. Botto, L. D. & Yang, Q. 5,10-Methylenetetrahydrofolate reductase gene variants and

congenital anomalies: A huge review. *American Journal of Epidemiology* **151**, 862–877 (2000).

12. Lori Lathrop Stern, Joel B. Mason, J. S. and S.-W. C. Genomic DNA Hypomethylation, a Characteristic of Most Cancers, Is Present in Peripheral Leukocytes of Individuals Who Are Homozygous for the C677T Polymorphism in the Methylenetetrahydrofolate Reductase Gene. *Cancer Epidemiol. Biomarkers Prev.* **9**, 849–53 (2000).

13. Kumar, K. S. D., Govindaiah, V., Naushad, S. E., Devi, R. R. & Jyothy, A. Plasma homocysteine levels correlated to interactions between folate status and methylene tetrahydrofolate reductase gene mutation in women with unexplained recurrent pregnancy loss. *J. Obstet. Gynaecol. (Lahore).* **23**, 55–58 (2003).

14. Nelen, W. L. D. M. *et al.* Methylenetetrahydrofolate reductase polymorphism affects the change in homocysteine and folate concentrations resulting from low dose folic acid supplementation in women with unexplained recurrent miscarriages. *J. Nutr.* **128**, 1336–1341 (1998).

15. Bagley, P. J. & Selhub, J. A common mutation in the methylenetetrahydrofolate reductase gene is associated with an accumulation of formylated tetrahydrofolates in red blood cells. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13217–13220 (1998).

16. Zappacosta, B. *et al.* Genotype Prevalence and Allele Frequencies of 5,10-Methylenetetrahydrofolate Reductase (MTHFR) C677T and A1298C Polymorphisms in Italian Newborns. *Lab. Med.* **40**, 732–736 (2009).

17. Zappacosta, B. *et al.* 5,10-Methylenetetrahydrofolate reductase (MTHFR) C677T and A1298C polymorphisms: Genotype frequency and association with homocysteine and folate levels in middle-southern Italian adults. *Cell Biochem. Funct.* **32**, 1–4 (2014).

18. Bailey, L. B. & Gregory, J. F. Folate metabolism and requirements. *J. Nutr.* **129**, 779–782 (1999).

19. Servy, E. & Menezo, Y. The Methylene Tetrahydrofolate Reductase (MTHFR) isoform challenge. High doses of folic acid are not a suitable option compared to 5 Methyltetrahydrofolate treatment. *Clin. Obstet. Gynecol. Reprod. Med.* **3**, (2017).

20. Ulrich, C. M. & Potter, J. D. Folate Supplementation: Too Much of a Good Thing? *Cancer Epidemiol. Prev. Biomarkers* **15**, 189–193 (2006).

21. Servy, E. J., Jacquesson-Fournols, L., Cohen, M. & Menezo, Y. J. R. MTHFR isoform carriers. 5-MTHF (5-methyl tetrahydrofolate) vs folic acid: a key to pregnancy outcome: a case series. *J. Assist. Reprod. Genet.* **35**, 1431–1435 (2018).

22. Lucock, M. & Yates, Z. Folic acid fortification: A double-edged sword. *Current Opinion in Clinical Nutrition and Metabolic Care* **12**, 555–564 (2009).

23. Christensen, K. E. *et al.* High folic acid consumption leads to pseudo-MTHFR deficiency, altered lipid metabolism, and liver injury in mice. *Am. J. Clin. Nutr.* **101**, 646–658 (2015).

24. Ashokkumar, B., Mohammed, Z. M., Vaziri, N. D. & Said, H. M. Effect of folate oversupplementation on folate uptake by human intestinal and renal epithelial cells. *Am. J. Clin. Nutr.* **86**, 159–166 (2007).

25. Seremak-Mrozikiewicz, A. Metafolina - alternatywa dla suplementacji niedoboru folianów u kobiet ciezarnych. *Ginekol. Pol.* **84**, 641–646 (2013).

26. Prinz-Langenohl, R. *et al.* [6S]-5-methyltetrahydrofolate increases plasma folate more effectively than folic acid in women with the homozygous or wild-type 677C→T polymorphism of methylenetetrahydrofolate reductase. *Br. J. Pharmacol.* **158**, 2014–2021 (2009).

Conclusions and future perspectives

In this thesis work, a single-SNP discriminating strategy based on Loop-Mediated Isothermal Amplification (LAMP) was presented. The proposed approach was optimized with the aim to move SNP diagnostic technologies from specialized laboratories to decentralized settings.

SNPs are genetic variations represented by the substitution of a single nucleotide at a specific position in the genome with a prevalence greater than 1% in the general population. SNPs reflect the natural variability of living organisms, as humans, animals and plants. For this reason, SNPs are exploited as markers for genetic testing in several areas. In particular, in food quality control, SNPs genotyping provides an unequivocal identification of ingredients and varietal discrimination of food products, playing a key role in Identity Preservation along all the agri-food value chain. On the other side, in clinical practice, SNPs are associated to specific phenotypes that influence the individual responses to certain substances, such as nutrients, supplements or drugs. In this way, the genotyping of this kind of SNPs can give strategic information on food intolerances predisposition or susceptibility to drug treatments.

Due to the complexity in terms of costs, laboriousness and required instrumentations, the currently SNP genotyping techniques are not adaptable to decentralized settings and on-site diagnostics. Therefore, in this thesis work, a single-SNP discriminating strategy suitable for Point-of-Care testing was optimized starting from a state-of-art approach, and it was applied to three different model cases. In particular, this approach was applied to an industrial relevant case addressing a specific requirement for food traceability, the varietal discrimination of durum wheat. Additionally, it was applied to two clinical cases, for the detection of the genetic predisposition to lactose intolerance and for the detection of a SNP which affects folic acid metabolism.

For the quality control application, the optimized colorimetric strategy was proven to be effective in the identification of both substitution and adulteration of mono-varietal grains, providing an unequivocal interpretation of the results based on a naked-eye instrument-free test readout.

For clinical fields, the optimized strategy was coupled with a non-invasive sampling, using saliva as clinical sample. Additionally, a fast and easy handled protocol for DNA extraction was introduced, using a simple and low-cost instrumentation. However, further optimization are demanded in order to use directly the sample without the dilution step, developing a completely portable strategy requiring only a small benchtop centrifuge and a simple heater device for the extraction.

For both clinically relevant applications, the single-SNP discriminating strategy resulted effective in providing a clear SNP genotyping, by detecting both heterozygous and homozygous samples. Despite the current available tests that require complex instrumentation, the employment of skilled users and necessitate a long turnaround time, the proposed assay, was demonstrated to provide SNP genotyping within 2 hours.

Overall, the strategy resulted versatile and potentially applicable to every case where a single SNP is involved. Moreover, the method resulted adaptable to both Real-Time fluorescent and colorimetric detections by interpreting the test result observing the increase of the fluorescence intensity signals or the color change of the reaction mixtures.

The acquired expertise coupled with the interest in providing new on-site diagnostics tools opens the perspective for future application of the single-SNP discriminating strategy to other real relevant cases in multiple fields.

Appendix

Anti-angiogenic effect of graphene oxide in primary human endothelial cells

Alongside the main project, I was involved in another project of the NanoBioInteractions & NanoDiagnostics Group regarding the investigation of the possible toxicological effects of Graphene Oxide on the angiogenesis process of human endothelial cells.

The here presented project gave significantly results that are described in the present appendix. The results obtained have been published in ACS Appl. Mater. Interfaces 2020, 12, 20, 22507–22518 doi.org/10.1021/acsami.0c03404 with the title "Antiangiogenic Effect of Graphene Oxide in Primary Human Endothelial Cells authored by Giulia Cibecchini, Marina Veronesi, Tiziano Catelani, Tiziano Bandiera, Daniela Guarnieri, and Pier Paolo Pompa.

1.1 Introduction

Endothelium is the inner layer of blood vessels and it is formed by endothelial cells. These cells are in direct contact with the bloodstream and act as a physical barrier that regulates the passage of endogenous and exogenous substances between the blood and the tissues, thus representing one of the primary defense for our organism [1]. In addition, endothelial cells can promote and sustain the formation process of new blood vessels from preexisting vasculature. This phenomenon, called angiogenesis, occurs in physiological conditions to maintain the body homeostasis and can be altered in several pathological states such as cancer, ischemic and inflammatory diseases [2]. Hence, understanding if exogenous agents, such as xenobiotics, nanomaterials, etc., may modulate the angiogenesis process is pivotal to avoid imbalance in the body homeostasis, and, at the same time, to develop new therapeutic strategies in the treatment of the above-mentioned pathologies by exploiting the pro- or anti-angiogenic effects of these agents.

Graphene is a single layer of carbon packed in a honeycomb structure [3]. Due to its unique physical-chemical properties, such as room-temperature electron mobility, mechanical properties, optical absorption, impermeability to gases, ability to sustain extremely high densities of electric current and the possibility to be chemically functionalized, graphene and

its related materials (GRMs) have attracted tremendous attention, opening a wide range of potential applications in several fields [4, 3]. Among the others, the possible use of GRMs in biology and medicine, such as in bio-imaging, cancer theranostics, gene and drug delivery, tissue engineering, and biosensing, has been recently reported [5, 6, 4]. In particular, graphene oxide (GO), the oxidized form of graphene achieved by chemical exfoliation of graphite, is considered a promising material for biological applications due to its excellent aqueous solubility [4]. However, the growing interest on the potential use of GO and the development of its large-scale production raise concerns about the nanosafety of this new material, to determine any potential risk for humans [7]. Therefore, *in vitro* cytotoxicological investigations and systematic evaluations of its biocompatibility are essential to develop GO-based biomedical materials before their application *in vivo*.

In the last decade, numerous studies have investigated the toxicology of GO on a broad spectrum of cell types. However, how GO elicits its biological effects is still not clear. In particular, there is a lack of systematic studies on the interaction of GO with biological systems to analyze the mechanisms and the molecular pathways involved at the cellular and subcellular level. In the assessment of GO safety, it is also fundamental to consider the route of exposure that determines the physiological and cellular environment that such materials meet when entering the human body, determining their biological effects. Considering the application potential of GO as drug or gene carriers, the presence of GO-based materials in the bloodstream and the interaction with the blood vessel walls is plausible. Various studies reported in the literature have investigated the GO biocompatibility and the possibility to use GO to affect angiogenesis. Indeed, many conflicting findings on the effects of the interaction of GRMs with the process of formation of new vasculature have been reported to date [8, 9, 10, 11, 12]. To shed light on this scenario, systems biology approaches based on integrated omics and bioinformatics analyses could be used as powerful tools to explore the interactions between nanomaterials and biosystems [13]. Among the others, metabolomics aims at the comprehensive characterization of all metabolites and, since metabolites are the products of a wide array of biochemical processes, information gained from metabolomics effectively complements other cell analyses.

In this work, we assessed the response of primary human umbilical vein endothelial cells (Huvec), chosen as an *in vitro* model of endothelium [14, 15, 9], after their interaction with GO, by taking advantages of an integrated experimental approach based on the combination of

cytotoxicity, angiogenesis and metabolomics analyses. In particular, we analyzed in depth parameters such as cell morphology, viability, migration, oxidative stress, as well as the internalization routes and ability of GO to affect angiogenesis. The results obtained by these analyses were correlated to preliminary metabolomics characterization to find the relationship between cytotoxicity, angiogenic potential alteration, and cell metabolism mediated by GO.

The results of the following work have been published [16]. The project was performed at NanoBioInteractions & NanoDiagnostics Group at Italian Institute of Technology in collaboration with the research group of Dr. Tiziano Bandiera of D3-Pharma Chemistry, Italian Institute of Technology, the Electron Microscopy Facility, Italian Institute of Technology, and under the supervision of Dr. Pier Paolo Pompa and Dr.ssa Daniela Guarnieri.

1.2 Results and discussion

Effect of GO exposure on cell viability and cell membrane integrity

GO flakes used in this work were obtained through oxidation of carbon fibers (GANF Helical-Ribbon Carbon Nanofibers, GANF). The initial GO suspensions were washed with MilliQ water by centrifugation, to remove the presence of acids, until a pH value of ≈5 was observed, as previously reported [17]. The mean GO flake length, determined by transmission electron microscopy (TEM) analysis, was around 400 nm, and the surface charge was negative, around −35 mV, at pH 7 [17].

Previous biological experiments carried out with the same GO material on different cell systems showed very different outcomes. More precisely, no cytotoxic effect on human intestinal epithelial Caco2 cells was observed up to 9 days of chronic exposure mimicking dietary ingestion [17]. Conversely, primary neurons and astrocytes were significantly affected by GO acute exposure [18, 19, 20]. In this work, we carefully examined the interactions of GO nanoflakes with primary human endothelial Huvec cells. After 24 h incubation with increasing concentrations of GO (ranging from 0 to 50 μg/mL), optical microscopy analysis revealed a huge amount of GO associated to the cells. In fact, GO appeared as dark areas organized in large aggregates or single spots mainly distributed in the cytoplasm and in the perinuclear region (Figure 1a). Moreover, at the highest concentrations, some GO aggregates were visible outside the cells interacting with the external part of the cell membrane. An even higher uptake was observed in HeLa cell line (Figure 1b), treated at the same conditions and used for comparison. Such large uptake in both cell types could be due in part to the good dispersion of GO nanoflakes at physiological pH as previously reported [17]. Because of the significant cellular internalization, the effect of GO on cell viability was assessed. To this aim, 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assay was performed after treatment of both Huvec and HeLa cells with different concentrations of GO (from 0 to 50 μg/mL) for 24 and 48 h. Despite microscopy observations indicated a similar association/internalization of GO in both cell types, the effect on cell viability was significantly different. In particular, a dose-dependent cytotoxicity was observed in Huvec cells: about 20% and 40% decrease of cell viability was observed after 48 h incubation with 25 and 50 μg/mL of GO, respectively (Figure 1c). Conversely, no cell death was detected in HeLa cells in the same experimental conditions (Figure 1d) even though the

amount of internalized GO was qualitatively comparable/superior to Huvec cells. These results further demonstrated the cell-dependent response to the same material, in agreement with previous works reporting on the variable biological reactions to GO treatment depending on cell type [21, 22, 23, 17]. For instance, it was showed that GO was not internalized by A549 cell lines and, as a consequence, did not lead to any decrease in cell viability [22]. Moreover, in Caco2 cells GO was poorly internalized and also in this case it did not induce evident cytotoxicity [17]. It was further demonstrated that incubating A549 and RAW 264.7 cells with GO for longer exposure times (4 days), a large amount of the nanomaterial was internalized and confined inside vesicular structures, with a resulting time- and dose-dependent cytotoxicity [23].

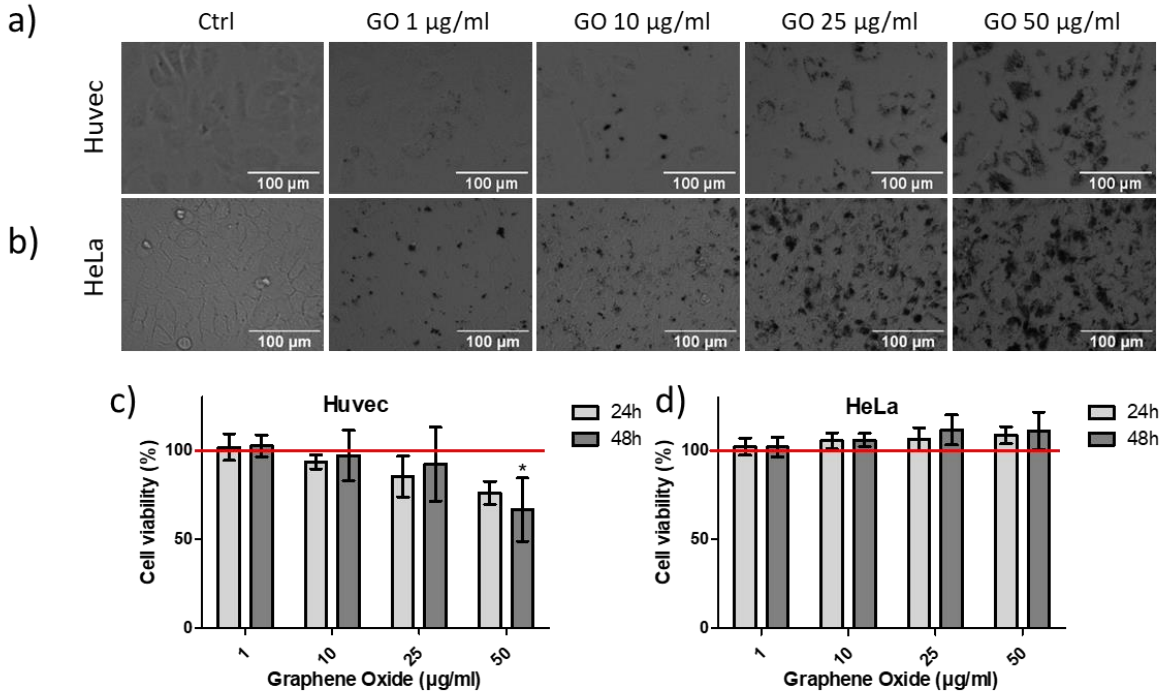

*Figure 1. Optical microscope images of Huvec (a) and HeLa cells (b) after 24 h incubation with increasing concentrations of GO, ranging from 0 to 50 µg/mL. Cell viability of Huvec (c) and HeLa (d) assessed with MTS assay after 24 and 48 h of incubation with GO. Data are normalized to the values of the untreated samples, set to 100% (red line), and represent the average of six different experiments performed in triplicate. The error bars represent the standard deviation. One-way ANOVA was performed to determine statistical significance between treatments and control (*p<0.05).*

To elucidate the mechanism underlying the viability decrease in Huvec cells by GO, we first looked at the plasma membrane integrity, assessing the extracellular release of lactate dehydrogenase (LDH) upon GO exposure. However, as shown in Figure 2, no significant variations in released LDH levels were found after 24 and 48 h of incubation with GO at all the concentrations tested. Therefore, the reported toxicity of GO was not dependent on the breakage of the cell membrane, in spite of its high intracellular accumulation. Consistently, other works showed that GO did not induce any damage at the cell membrane of Caco2 cells [24, 17].

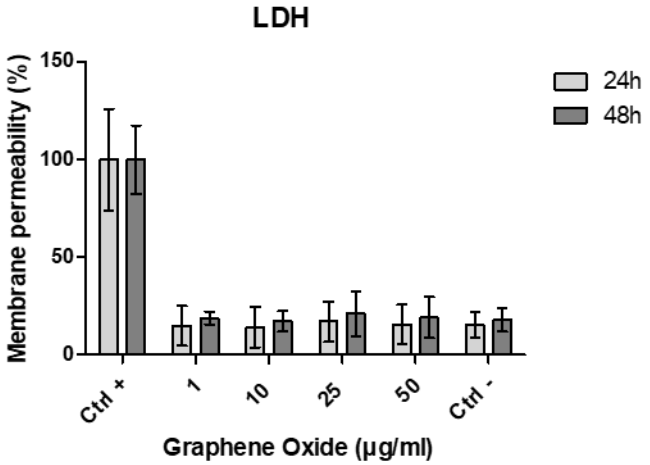

Figure 2. Membrane integrity assay on Huvec cell after GO incubation for 24 and 48 h. Positive controls were obtained with lysis solution provided by the kit. Data are normalized on positive control samples expressed as 100% and reported as the average ± standard deviation. The graph reports the average of three independent experiments. One-way ANOVA was performed to determine statistical significance between treatments and negative control (Ctrl-).

Cell uptake and intracellular localization of GO

Cell uptake and intracellular fate of GO are crucial issues to determine whether GO cytotoxicity is related to its interference with intracellular structures and/or pathways. Therefore, we analyzed GO intracellular localization by confocal microscopy, by exploiting its intrinsic property of efficiently scattering excitation light [25]. Numerous spots of GO, heterogeneous in size and shape, were visible in the cytosol and around the nucleus of Huvec cells (Figure 3), confirming the preliminary observations with the optical microscope. Such distribution suggested an endocytosis-mediated uptake of GO nanoflakes, as already demonstrated for other cell types [17, 18]. In fact, most of nanomaterials enter cells via endocytosis [26, 27], are confined in vesicles and, usually, end up into late endosome/lysosome

structures [28, 29, 30]. To address this issue, Huvec cells were stained with Lysotracker after 24 h of incubation with GO at the concentration of 0, 1, and 10 µg/mL. Lysotracker is a probe commonly used to study the intracellular trafficking of nanomaterials; it is selectively accumulated in lysosomes due to their acidic environment [31, 32]. However, as shown in Figure 4, no apparent colocalization of GO spots with Lysotracker was observed. Moreover, a significant decrease of Lysotracker signal was detected with the increase of internalized GO. We ascribed the quenching of Lysotracker fluorescence to the co-presence of large GO aggregates into lysosomes. In fact, several papers reported on the capability of GO to quench the fluorescence emission of other dyes [33, 34, 24], a property that is widely exploited for the development of graphene-based nanosensors [35, 36]. Therefore, the quenching of Lysotracker fluorescence can be considered an indirect evidence of the endo-lysosomal confinement of GO after its cellular internalization.

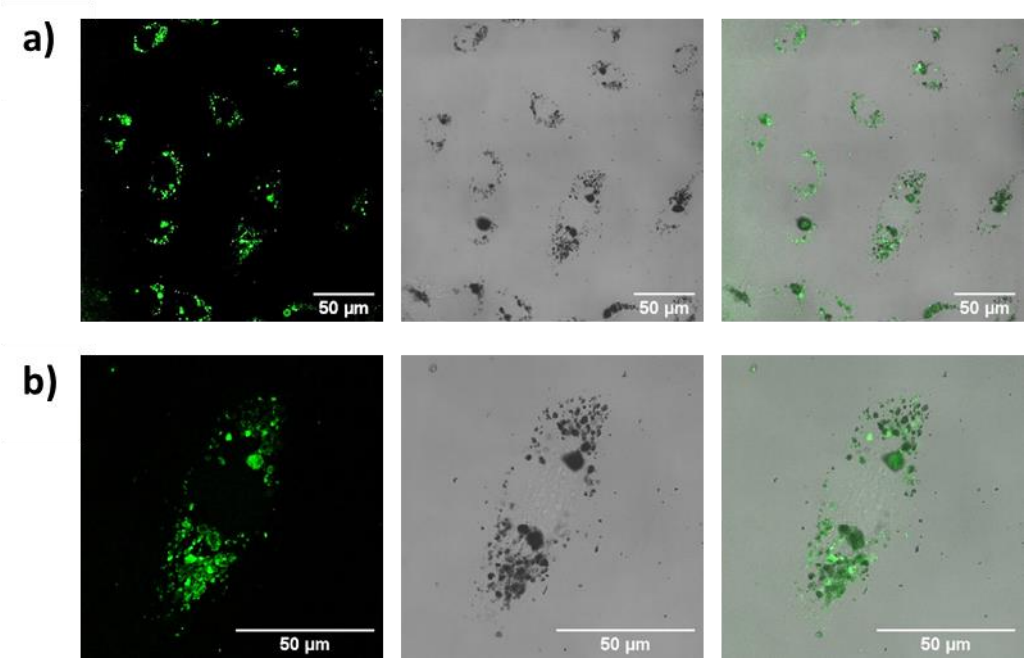

Figure 3. Confocal microscopy images of Huvec cells after 6 h of incubation with GO at the concentration of 25 µg/mL. Green spots represent the GO scattering of the laser beam (left panels); brightfield images showing GO as dark spots (middle panels); merge (right panels).

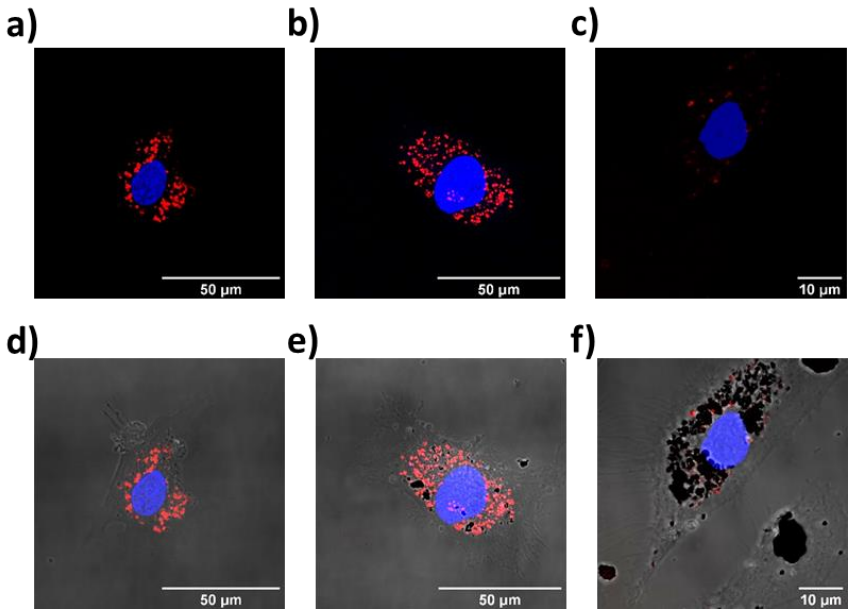

Figure 4. Confocal microscopy analysis of Huvec cells incubated 24 h with GO at 0 µg/mL (a and d), 1 µg/mL (b and e) and 10 µg/mL (c and f) and stained with Lysotracker (red) and Hoechst 33342 (blue) to highlight lysosomes and nuclei, respectively. Indirect co-localization of GO with lysosomes shown as red fluorescence quenching by GO. Top panels: fluorescence images; bottom panels: merge of fluorescence and brightfield images. GO appears as dark spots in the brightfield images.

To confirm this observation, transmission electron microscopy (TEM) analysis was performed (Figure 5), demonstrating the presence of GO nanoflakes inside the cells. GO was found both as single lamellar flakes and as big aggregates localized inside vesicular structures, prevalently distributed in the perinuclear region. The vesicular structures were likely endo-lysosomal compartments, thus indicating that GO material was internalized by endocytosis. Furthermore, the morphology of the cells as well as the plasma membrane resulted not to be affected by GO uptake, compared to the non-treated control cells (Figure 5 a and d), in accordance with what we observed through LDH assay.

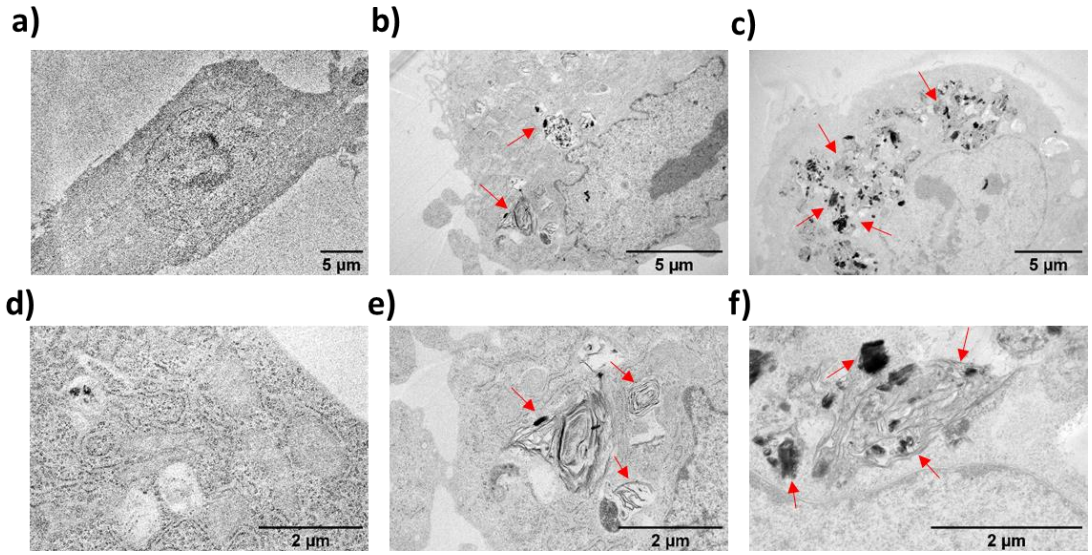

Figure 5. TEM images of GO flakes internalized in Huvec cells after 24 h of incubation with GO at the concentration of 0 µg/mL (a, d), 1 µg/mL (b, e) and 10 µg/mL (c, f). Red arrows indicate GO nanoflakes.

Effect of GO exposure on angiogenesis

Huvec cells are able to form capillary-like structures after specific stimulation *in vitro* [37, 14, 15, 38]. Here, these cells were used to investigate whether their interaction with GO could promote or inhibit this process by performing a Matrigel assay. As reported in Figure 6a, the exposure of cells to GO for 6 h reduced their capability to form capillary-like structures in a dose-dependent manner. In particular, we observed a reduction (around 40%) in the number of meshes formed by cells treated with the highest GO concentration (50 µg/mL) compared to the positive control. In fact, optical microscopy images showed areas with less interconnections among the cells forming the capillary network after the treatment with GO than in the positive control (Figure 6b).

Interference of GO with angiogenesis process is a critical issue in the literature. It was investigated using various assays and *in vitro* or *in vivo* models leading to different and conflicting findings. For instance, Mukherjee et al. demonstrated by egg yolk angiogenesis assay (chick embryo angiogenesis, CEA) in ECV-304 cells that, depending on the GO concentrations and the amount of intracellular ROS formation, this nanomaterial could be proangiogenic or anti-angiogenic [8]. On the contrary, Pei-Xin Lai et al., using both tube

formation assay and chick chorioallantoic membrane (CAM) assay, reported that bovine serum albumin-capped GO (BSA-GO) possessed binding affinity towards VEGF-A165, an angiogenic factor, inhibiting angiogenesis in Huvec cells by sequestering VEGF-A165 [9]. Moreover, Jeong et al. demonstrated that both GO and PEG-coated GO caused various angiogenic patterning defects, such as incomplete vessels, in zebrafish embryos suggesting that GO affected the normal growth and branching of vessels by regulating VEGF and Notch signaling pathways essential for angiogenesis [10]. In agreement with some of these studies, our observations showed a negative effect of GO exposure on the angiogenic potential of Huvec cells with a dose dependent inhibition of capillary-like structure formation.

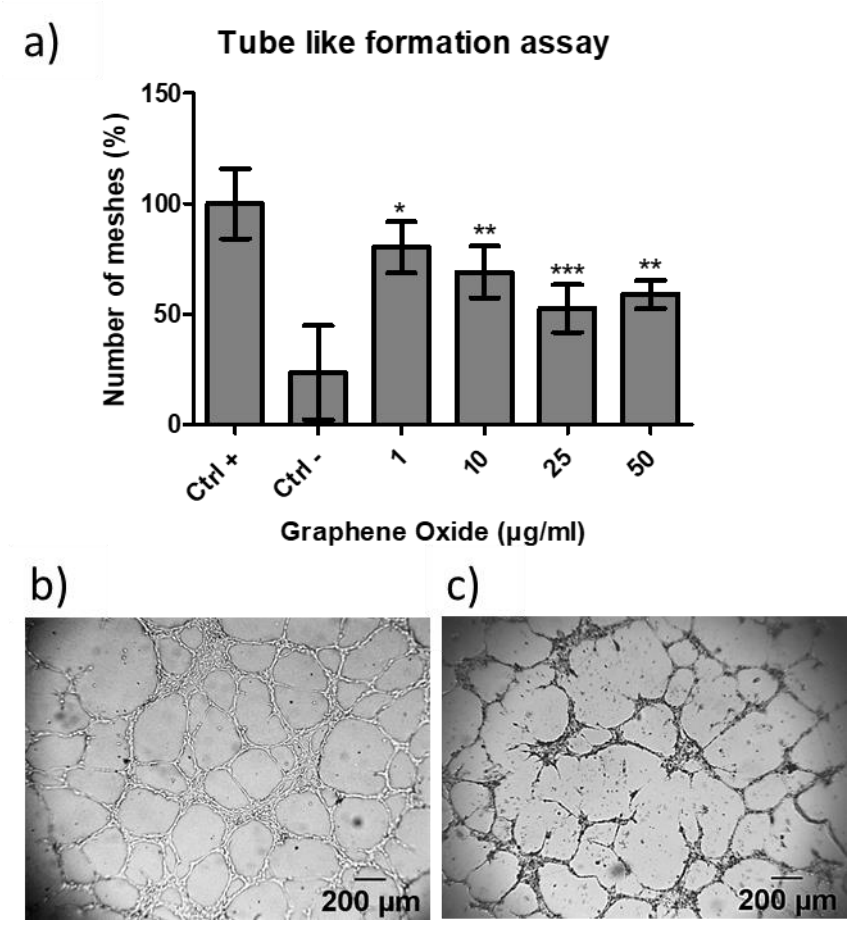

*Figure 6. Capillary tube like-formation assay on Huvec cells after 6 h incubation of GO. Data are normalized to positive control samples expressed as 100% and reported as the average ± standard deviation. The graph reports the average of six independent experiments (a). Representative images of positive control (b) and 50 µg/mL GO treated samples after 6 h of incubation (c). One-way ANOVA was performed to determine statistical significance between treatments and positive control (Ctrl+) (*** p<0.001, ** p<0.01, * p<0.05).*

Cell migration ability after GO exposure

Angiogenesis is a complex process that requires migration of endothelial cells [39]. Basically, endothelial cells activate migratory pathways in response to growth factors gradients, integrins bindings, and mechanical stress induced by blood flow. These stimuli are able to drive the remodeling of the actin cytoskeleton into plasma membrane projections and actin protrusions, which are involved into the extension at the leading edge, the attachment to extracellular matrix and the contraction of the cell body [39, 40].

We investigated if the intracellular accumulation of GO could interfere with the cell motility as a possible explanation of its anti-angiogenic effect. The assessment of the migration ability of Huvec cells upon the incubation with GO was performed by a wound healing assay. We observed that the interaction with GO affected the ability of cells to heal the wound over time and depending on the GO concentration (Figure 7). At the highest concentration (25 μg/mL) only ~60% of the wound was repaired compared to the positive control (100%) and cells treated with 1 μg/mL of GO (~80%) (Figure 7a). Representative images of Huvec cells acquired after 24 h incubation with GO at the concentration of 0, 1 and 25 μg/mL were showed in Figure 7(b-d). Similar results were obtained by Zhou et al. who reported that graphene and GO both inhibited migration and invasion of cancer cells by suppressing the activity of the electron transfer chain in mitochondria, leading to a decrease in ATP production fundamental for actin polymerization [41]. In contrast, Hussein et al. demonstrated that GO subjected to ultra-sonication for two days was biocompatible at low concentrations and promoted cell migration [42]. Moreover, the use of graphene in wound healing was reported by Lu et al. who prepared chitosan–PVA nanofibers containing graphene and by Fan et al. who designed an Ag/Graphene polymer hydrogel; both of these composite materials combined wound repair with antibacterial intrinsic activity of graphene [43, 44]. However, in our experimental conditions, GO was abundantly internalized and accumulated inside the cells within endo-lysosomal vesicles. Therefore, the close interaction of this nanomaterial with the intracellular machinery could elicit different cellular response by affecting specific cell functions and structures. Among the others, the rearrangement of cytoskeleton is important for cell migration [45]. Analyzing the assembly of actin microfilaments, we observed some differences between cells treated with GO and non-treated cells (Figure 8). In particular, confocal microscope images of Huvec cells treated with GO for 24 h at the concentration of 1 and 10 μg/mL and labelled with phalloidin indicated that actin microfilaments appeared less organized in proximity of the GO

intracellular aggregates. On the other hand, actin fibers of untreated cells were orderly assembled and oriented inside the cytoplasm. It is known that cell migration and endocytosis use the same cellular structures such as cytoskeleton to work. Previously, Frontiñan-Rubio et al., observed cytoskeleton alterations in keratinocytes treated with the same GO for 7 days [46]. These observations in association with our results on Huvec migration ability demonstrated that GO, internalized and confined inside lysosomes, interfered with the arrangement of cell body required for migration process of these cells and consequently with their ability to form loop structures resembling a capillary network.

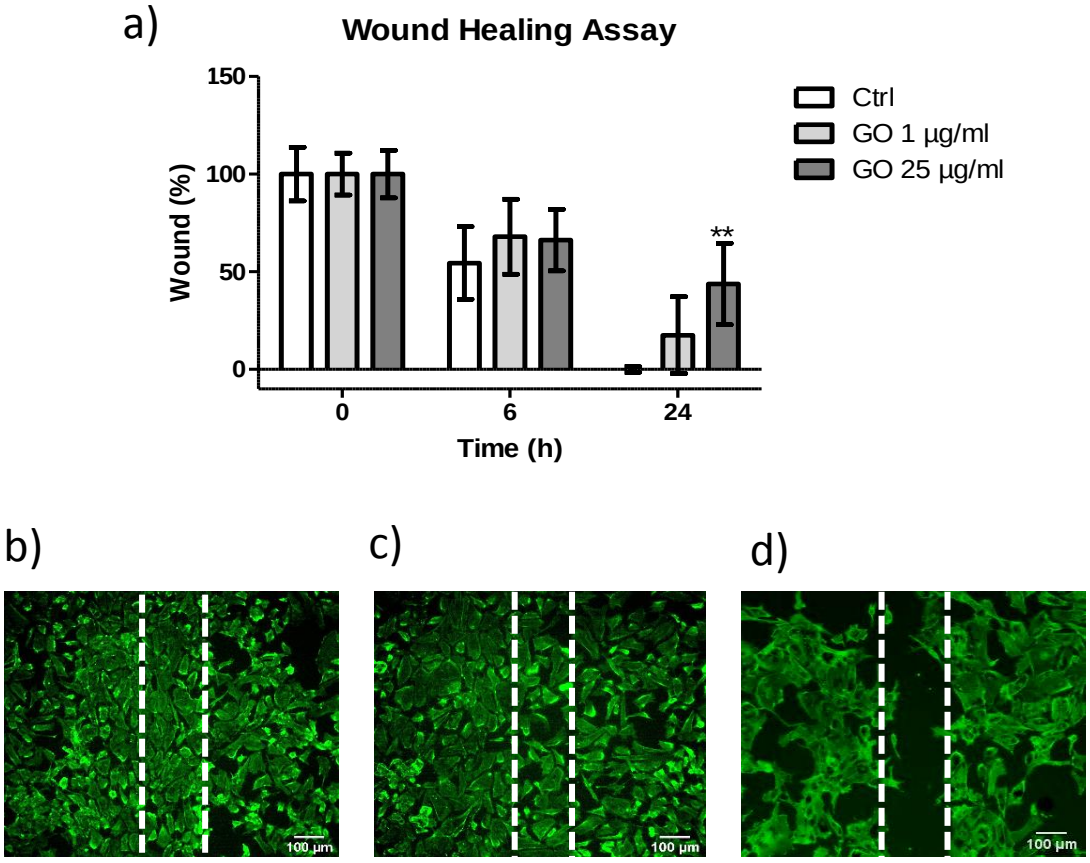

*Figure 7. (a) Wound healing assay on Huvec cells after 0, 6 and 24 h of incubation with GO. Data were normalized on the starting time point expressed as 100% (reported as mean ± standard deviation of three independent experiments) (b-d). Representative images of Huvec cells labelled with Alexa-Fluor 488-phalloidin for F-actin (green) at 24 h time point. Cells were treated with GO at the concentration of 0 µg/mL (b), 1 µg/mL (c) and 25 µg/mL (d). Scale bar 100 µm. One-way ANOVA was performed to determine statistical significance of treatments (GO 1 and 25 µg/ml) respect to non-treated cells (Ctrl) at comparable time-points (** p<0.01).*

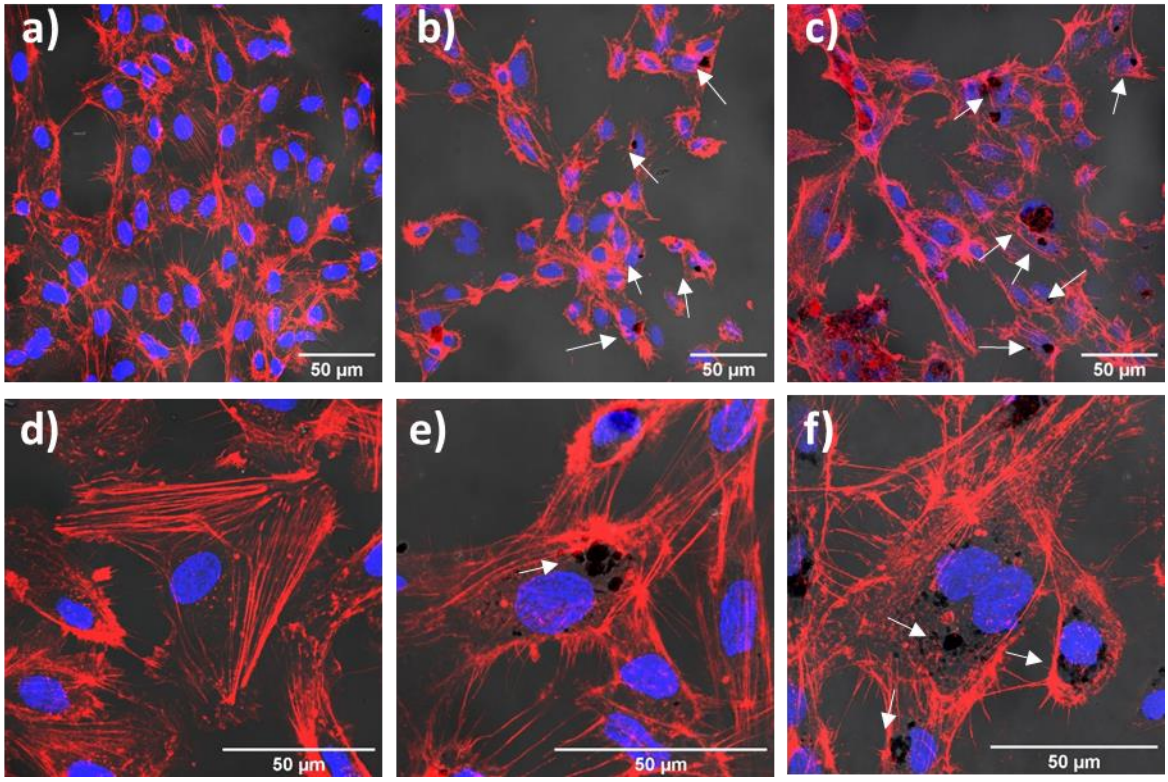

Figure 8. Confocal images of Huvec cells after 24 h of incubation with GO at the concentration of 0 µg/mL (a, d), 1 µg/mL (b, e) and 10 µg/mL (c, f). Cells were stained with Phalloidin (red) and Hoechst 33342 (blue) to highlight actin filaments and nuclei, respectively. GO was acquired in brightfield mode (dark spots). White arrows highlight GO flakes.

Effect of GO in Reactive Oxygen Species (ROS) production

The overproduction of ROS is related to cellular stress conditions. To further elucidate the effects of GO on Huvec cells, we assessed the generation of ROS after the exposure to different concentrations of this material for 24 and 48 h. ROS were detected using 2',7'-dichlorodihydrofluorescein diacetate (H_2DCFDA) that is a cell permeant molecule, which can be oxidized by intracellular ROS forming 2',7'-dichlorofluorescein (DCF), a green fluorescent product, commonly used for the assessment of oxidative stress [23, 22, 47]. As reported in Figure 9a, a significant increase of ROS levels was observed after 24 and 48 h of incubation with GO at higher concentrations (50 µg/mL) in comparison with untreated cells. A similar dose dependent trend resulted from the viability assay, suggesting that the oxidative stress strongly contributes to the toxicity of the GO nanomaterial, in agreement with previous studies. In fact, exposure to GO was reported to induce the formation of intracellular ROS in several cell types including Huvec [23, 22, 24].

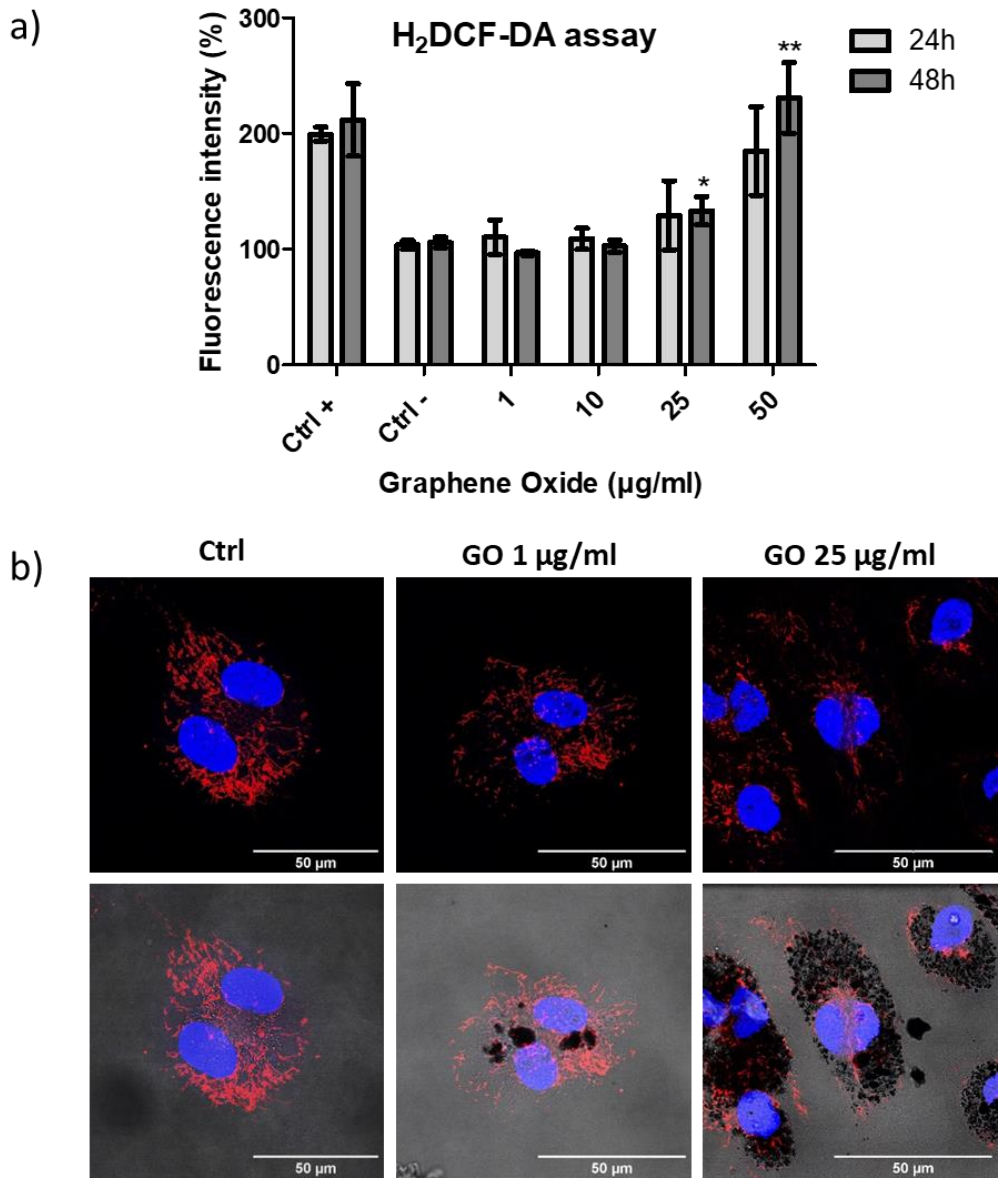

*Figure 9. (a) ROS levels in Huvec cells untreated or exposed to GO ranging from 1 to 50 µg/mL for 24 and 48 h. Positive control (Ctrl +) was performed treating cells with 1 mM H2O2 for 15 minutes. Data are normalized to untreated cells (Ctrl -) and reported as average ± standard deviation of three independent experiments. One-way ANOVA was performed to determine statistical significance between treatments and negative control (** p<0.01, * p<0.05). (b) Confocal microscopy images of Huvec cells after 24 h incubation with GO. Mitochondria were stained with Rhodamine 123 (red) and nuclei with Hoechst 33342 (blue) (top and bottom panels). GO was acquired in brightfield mode (dark spots) as reported in the bottom panels that show the merge of fluorescence and brightfield images.*

In oxidative stress conditions, most of the ROS are produced by mitochondria. These organelles play a vital role in energy production and are involved in programmed cell death [48, 49, 50, 51]. Studying GO related cytotoxicity, Zhou et al. found a correlation between the size

of GO flakes and their ability to accept electrons from the electron chain transport in mitochondria [41]. Therefore, in our study, we looked at mitochondria to verify any possible interaction between these intracellular organelles and the internalized GO nanomaterial. TEM analysis did not show GO inside or associated to the mitochondria (Figure 5). Moreover, their morphology was regular and not affected by the treatment of the cells with GO. These observations suggested that there was no direct interaction between the internalized GO and mitochondria. However, confocal microscopy analysis of Huvec cells treated with 0, 10 and 25 µg/mL of GO for 24 h and labelled with Rhodamine 123 dye, which is widely used to mark mitochondrial organelles [41, 52, 53], showed differences in their intracellular distribution (Figure 9 b). We hypothesized that the presence of large GO aggregates confined in endo-lysosomal structures could change the local distribution of mitochondria: i) directly, due to physical hindrance within the cytoplasm as evidenced in Figure 9b, and ii) indirectly, because of the incorrect assembly of microfilaments (Figure 8) that are involved in the intracellular trafficking of mitochondria [54]. Notably, it has been recently demonstrated that defects in the interaction of mitochondria with actin microfilaments compromise mitochondria functionalities and may affect cell fate and apoptosis in retinal pigment epithelial cells [55].

Metabolomics analysis by Nuclear Magnetic Resonance (NMR)

The metabolism of endothelial cells has been recognized as a driving force of angiogenesis [56]. Changes in endothelial metabolism are able to induce or inhibit different steps of the angiogenic process [57]. Consequently, external agents capable to reprogram cell metabolism may influence their angiogenic potential. To corroborate all the results obtained so far and to further clarify how GO interferes with cellular functions, we performed an exploratory metabolic analysis of the media conditioned by Huvec cells cultured in presence and in absence of 25 µg/mL GO. Nuclear Magnetic Resonance (NMR) is a well-recognized technique for the identification and quantification of metabolites in growth media, biofluids, and intact tissues [58, 59, 60]. The technique is nondestructive, noninvasive, very robust, and shows high reproducibility. Moreover, due to its physical characteristic, NMR represents an intrinsically quantitative method and allows the simultaneous quantification of different molecules, providing a precise picture of the metabolic profile in a precise moment [58]. As a preliminary study, we focused on the analysis of growth media of GO-treated and untreated cells, which were harvested and analyzed after 24 h of incubation. Both unsupervised Principal

Component Analysis (PCA) analysis and supervised multivariate analysis (PLS-DA) segregated very well the two groups (controls and GO-treated, see Figure 10a and 10b, respectively), indicating a different metabolite profile and a consequent alteration of the cell metabolism by the graphene-oxide. The quantitative NMR analysis showed a different consumption of several metabolites in the samples treated with GO compared to non-treated controls, revealing a substantial effect of GO exposure in the metabolic profile of Huvec cells. In particular, we found statistically significant differences in consumption levels of niacinamide and several amino acids, namely phenylalanine, arginine, isoleucine, proline, tyrosine, leucine and glutamine (Figure 10c). Conversely, no differences in consumption of lactate and glucose were observed (Figure 10d), likely suggesting no direct interference of GO with glycolysis.

Niacinamide is a precursor of nicotinamide adenine dinucleotide (NAD^+): it is present in the extracellular environment, can be transported into cells and used for the synthesis of NAD^+ [61]. NAD^+ and its reduced form, NADH, play an important role in the maintenance of the cellular redox balance. NADH, in fact, participates in several intracellular reactions as electron donor, maintaining cellular redox homeostasis [62]. In turn, the cellular redox balance is crucial to the survival of cells because an increase in oxidative species may trigger apoptotic pathways [62]. Therefore, an increase in the consumed levels of niacinamide suggests an alteration of NAD^+/NADH balance and, hence, oxidative stress that we actually observed upon GO incubation (Figure 9).

The observed variation in amino acids consumption further indicates an effect of GO on endothelial cell activities. Beside glycolysis and fatty acid oxidation, in recent years, it has been shown that amino acid metabolism also plays an important role in regulating and maintaining vascular functions [63]. These functions include vascular tone, coagulation and fibrinolysis, cell growth and differentiation, redox homeostasis, plus immune and inflammatory responses. Notably, our results indicated a higher consumption of some essential (isoleucine, leucine, and phenylalanine) and non-essential (proline, tyrosine, glutamine, and arginine) amino acids. The first ones must be provided by the diet, while the second ones can be synthetized *de novo* by the cells. It has been reported that the increase of oxidative stress can reprogram the metabolism and affect *de novo* amino acid synthesis in endothelial cells [63]. In general, amino acids are fundamental for the regulation of energy metabolism and some of them are linked to the intracellular oxidative stress level in endothelial cells. For these reasons, a modification of their consumption levels could be a

benchmark of cellular stress [64]. In particular, recent studies have demonstrated that glutamine catabolism is essential for endothelial proliferation through its role in the replenishment of the Krebs cycle [65]. Concomitantly, depriving endothelial cells of glutamine caused vessel sprouting defects due to impaired proliferation and migration [65] . Furthermore, arginine plays a crucial role in regulating the angiogenic process through different products of its catabolism. Indeed, arginine is not only involved in protein synthesis, it is also a single substrate for endothelial nitric oxide synthase (eNOS) to produce the important vasoprotective molecule nitric oxide (NO) [66]. Regarding angiogenesis, conflicting data suggested that NO can inhibit or stimulate this process depending upon the NO level [67]. Additionally, proline is a major amino acid for the synthesis of collagen and the production of extracellular matrix; therefore, it may play an important role in vascular remodeling [68].

Although our results were preliminary and other investigations are necessary, these findings indicated that GO interfered with biochemical processes of Huvec cells involved in the redox balance, cell proliferation and migration, thus perturbing cell homeostasis and angiogenesis.

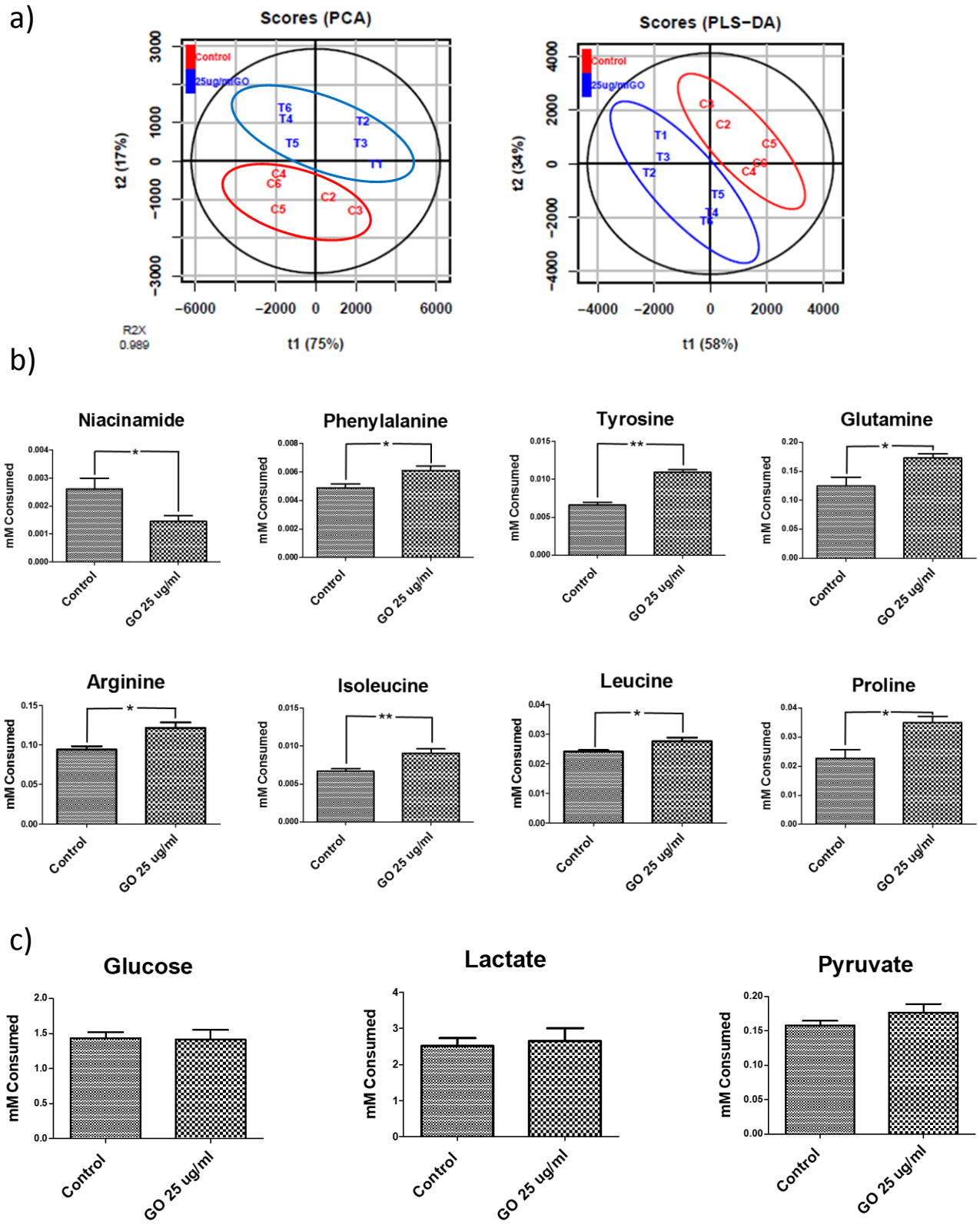

Figure 10. Principal Component Analysis (PCA) (a) and Partial Least Squares for Discriminant Analysis (PLS- DA) (b) score plots of Huvec cells treated with 25 µg/mL Graphene oxide (T) and without GO (C) for 24 h. Boxplots of representative metabolites identified from univariate analysis as significantly different between Huvec cells treated with and without GO (T and C, respectively) (c). Boxplots of Glucose, Lactate and Pyruvate metabolites that have shown not significative difference between control and treatment (d) ($p < 0.05$; ** $p < 0.005$).*

1.3 Materials and methods

Preparation of GO suspension

Graphene oxide (GO) was provided by Grupo Antolin (Burgos, Spain). The stock solution was prepared by suspending GO nanoflakes in sterile MilliQ water at the final concentration of 0.5 mg/mL and it was sonicated 5 minutes in an ultrasonic bath (Ultrasonic Cleaner - VWR) before each use. GO batches were tested for the presence of endotoxin by Limulus Amebocyte Lysate (LAL) assay (Pierce, Thermo Scientific). Negligible levels of endotoxin were found according to US Food and Drug Administration (FDA) guidelines.

Cell culture

Human umbilical vein endothelial cells (Huvec) (Gibco) were cultured in Medium 200 (Gibco) supplemented with Low Serum Growth Supplement Kit (LSGS Kit, Gibco) according to the manufacturer procedure. The final concentrations of the components in the supplemented medium were: 2% v/v fetal bovine serum (FBS), 2% v/v; hydrocortisone, 1 µg/mL; human epidermal growth factor, 10 ng/mL; basic fibroblast growth factor, 3 ng/mL; and heparin, 10 µg/mL. HUVECs were used for experiments between passages 2 and 6. Human cervix epithelioid carcinoma cell lines (HeLa) (ATCC) were cultured in Dulbecco's Modified Eagles Medium (DMEM) (Sigma-Aldrich) supplemented with 10% (v/v) FBS (Sigma-Aldrich), 100 U/mL penicillin and 100 mg/mL streptomycin (Sigma-Aldrich). Cells were maintained in an incubator at 37 °C under a humidified atmosphere and 5% CO_2.

MTS assay

Cell viability after exposure to GO was assessed by measurement of the metabolic activity via MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-cyrboxymethoxy-phenyl)-2-(4-sulfophenyl)-2H-tetrazolium, inner salt) assay (Cell- Titer96 Aqueous One Solution, Promega). The assay was performed following the manufacturer's protocol. Huvec and HeLa cells were seeded in a 96-well plate at the density of 1 x 10^4 cells per well and incubated overnight at standard culture conditions. The medium was removed, and cells were incubated with different concentrations of GO (0, 1, 10, 25, 50 µg/mL), previously dispersed in supplemented Medium 200 for Huvec cells and complete DMEM for Hela cells. Cells were incubated for 24 and 48 h at 37 °C and 5% CO_2. Non-treated cells were used as a control. After the incubation, the cell

culture medium was removed, replaced by 120 µL of MTS working solution (20 µL MTS reagent plus 100 µL of cell culture medium) and incubated with cells for 2 h at 37 °C and 5% CO_2. The absorbance at 490 nm wavelength of each sample was measured using a plate reader (Biotek Synergy HT plate reader). The absorbance value of starting MTS working solution diluted in cell culture medium was taken as blank. Cell viability was calculated as (Absorbance samples – Absorbance blank). Data were normalized to the control samples (non-treated cells) and reported as mean ± standard deviation of four replicates of six different experiments. One-way Anova was performed to determine statistical significance.

Membrane integrity assay

The plasma membrane integrity after exposure to GO was assessed as a function of lactate dehydrogenase (LDH) extracellular release by CytoTox-ONE™ Homogeneous Membrane Integrity Assay (Promega) following the manufacturer's protocol. Briefly, Huvec cells were seeded in a 96-well plate at the density of 1 x 10^4 cells per well and were grown overnight at standard cell culture conditions. The medium was removed, and cells were incubated with different concentrations of GO (0, 1, 10, 25, 50 µg/mL), previously dispersed in cell culture medium, for 24 and 48 h at 37 °C. Non-treated cells were used as a control. Assay plate was removed from 37 °C incubator and equilibrate to 22 °C approximately 20 minutes before the test. Then, 100 µL of CytoTox-ONE™ mix solution were added directly into each well. Positive control samples were performed adding 2 µL of lysis buffer solution, provided by the kit, and blank samples were prepared mixing 100 µL of CytoTox-ONE™ mix with cell-free culture medium. The plate was incubated at 22 °C for 10 minutes. 50 µL of Stop solution were added to each well and the absorbance at 560 - 590 nm wavelength was measured using a plate reader (Biotek Synergy HT plate reader). The absorbance of blank was subtracted to the absorbance of the samples. Data were normalized to negative control samples and reported as the average ± standard deviation. The graph reports the average of three independent experiments. One-way Anova was performed to determine statistical significance.

Fluorescence cell staining and confocal microscopy analysis

The intracellular localization of GO was studied by staining lysosomes with LysoTracker dye (Life Technologies) at the final concentration 1.54 x 10^{-7} M. Briefly, Huvec cells were seeded

on a glass bottom dish (WillCo Wells) grown overnight at standard cell culture conditions and then incubated with GO at the concentration of 0,1 μg/mL and 25 μg/mL. After 24 h of incubation, the cells were washed twice with PBS with Ca^{2+} and Mg^{2+} and incubated with LysoTracker in PBS for 15 minutes at 37 °C and 5% CO_2 and Hoechst 33342 (Thermo Fisher Scientific) at a concentration of 5 μg/mL for 5 minutes at room temperature. After the staining, the cells were washed twice with PBS with Ca^{2+} and Mg^{2+} and Live Cell Imaging Medium (Gibco) was added to each dish. Images were acquired in live staining by a confocal microscope (Leica SP5 Inverted) with a 63× oil immersion objectives, 405 and 561 nm excitation laser wavelengths and a resolution 1024 × 1024 pixels. GO was excited with 488 nm laser beam and light scattering of the laser was captured localizing the emission window at 505-530 nm. To observe cell morphology upon GO incubation, cells were washed twice with PBS, fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton X100 in PBS and stained with Alexa-Fluor 488- phalloidin or Alexa-Fluor 546- phalloidin. Cells were then analysed by the confocal microscope as described above.

Transmission Electron Microscopy (TEM) analysis

To perform TEM analysis, the cells were seeded on a 6 mm glass coverslip, grown overnight at standard cell culture conditions and incubated with GO at 0, 1 and 10 μg/mL concentration in cell culture medium for 24 h. At the end of incubation, cells were washed with PBS, fixed for 2 h in 1.5% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.4), post fixed in 1% osmium tetroxide in the same buffer and stained overnight with 1% uranyl acetate aqueous solution. Cells were then dehydrated in a graded ethanol series, infiltrated with propylene oxide and embedded in epoxy resin (Epon 812, TAAB). Semithin and thin sections of the embedded cell monolayer were cut with an ultramicrotome (UC6, Leica) equipped with a diamond knife (Diatome). Images were collected with a Jeol JEM 1011 (Jeol, Japan) electron microscope, operating at an acceleration voltage of 100 kV, and recorded with a 11 Mp fiber optical charge-coupled device (CCD) camera (Gatan Orius SC-1000).

Capillary-like tube formation assay

Capillary-like tube formation assay was performed by slightly modifying the procedure described elsewhere [16]. Matrigel Basement Membrane Matrix provided by Corning was

thawed by submerging the vial in ice. After that, 50 µL of Matrigel Matrix was added to each well of a 96-well plate keeping it on ice all times. Then, the multiwell plate was incubated 30 minutes at 37 °C and 5% CO_2 to let the gel to form. GO was diluted in Medium 200 with 10% FBS at different final concentrations. Cells were harvested, counted and directly suspended in the previously prepared GO suspensions at the density of 1 x 10^4 cells per well. Negative controls samples were also prepared suspending cells in serum free medium. The cells were then seeded on the top of the preformed gel and incubated 6 h at 37 °C. After the incubation time, cells were washed two times with PBS with Ca^{2+} and Mg^{2+} and fixed with 4% paraformaldehyde for 20 min at room temperature. Images of the samples were acquired by a Microscope Leica 6000 at 5x magnification and analyzed with ImageJ Plugin Angiogenesis Analyzer. Data were reported as the average ± standard error of the number of meshes formed for each sample. The graph reports the average of six independent experiments. One-way Anova was performed to determine statistical significance respect to the positive control.

Wound healing assay

Cells were seeded at 90% of confluence in a 24-well plate. When a cell monolayer for each well was formed, a "wound" was created using a 10 µL sterile tip. Then, cells were incubated with increasing concentrations of GO, previously dispersed in complete Medium 200, and non-treated samples were considered as a control. Four images for each sample were acquired by Microscope Leica 6000 after 0, 6 and 24 h from the wound to determine the migration rate. The width of the wound was measured using ImageJ (Fiji) and reported as the average ± standard deviation. The graph reports the average of three independent experiments. One-way Anova was performed to determine statistical significance of the treatments respect to the non-treated cells at comparable time-points.

Detection of reactive oxygen species (ROS) by H_2DCFDA

The production of reactive oxygen species (ROS) was determined by 2',7'-dichlorodihydrofluorescein diacetate (H_2DCFDA) (Invitrogen). Briefly, Huvec cells were seeded with 1 × 10^4 cells in 100 µL of complete cell culture medium per well in a 96-well plate and grown overnight under standard cell culture conditions. Then, medium was removed and cells were incubated with increasing concentration of GO, previously dispersed in cell culture

medium, and incubated for 24 h at 37 °C. Cells were washed two times with Hanks' Balanced Salt solution (HBSS) and incubated with H_2DCFDA diluted in HBSS at the final concentration of 5 μM for 30 minutes at 37 °C. Positive control samples were treated with 1 mM H_2O_2 for 15 minutes at 37 °C before the staining with H_2DCFDA. A Biotek Synergy HT plate reader was used for the DCF fluorescence intensity, setting the excitation filter at 480 nm and the emission filter at 520 nm. After the detection of the fluorescence cells were washed twice with HBSS and incubated with MTS assay solution for 2 h at 37 °C. The absorbance at 490 nm wavelength of each sample was measured using the plate reader as explained before. Data from H_2DCFDA assay was normalized to the number of cells and the results were reported as a percentage. More precisely, the results were reported as the ratio between the fluorescence intensity of the DCF and the absorbance of the MTS assay. The results were reported as a percentage where the untreated cells represented 100%. The graph reports the average of three independent experiments. One-way Anova was performed to determine statistical significance.

Metabolomics analysis by Nuclear Magnetic Resonance (NMR)

Sample preparation:

Briefly, 2×10^6 Huvec cells were seeded in a 10 cm plate and were grown overnight at standard cell culture conditions. The medium was removed, and cells were incubated with the same medium containing 25 □g/mL GO (treatment T) or with medium without GO (controls C) for 24 h at 37 °C: six replicates for each condition T and C. After 24 h, 2 mL of medium for each sample were collected, immediately frozen in liquid nitrogen and stored at -80°C until NMR analysis. Just before the analysis, the media were thawed on ice and 400 □L of each sample were diluted with 100 □L of a prepared NMR buffer for a final concentration of 150 mM buffer Phosphate pH 7.4, 1 mM 2,2',3,3'-deuterotrimethylsilylproprionic acid (TSP), 0.04% Sodium Azide, and 10% D_2O (for the lock signal), into 5 mm NMR tube.

NMR analysis:

All the NMR experiments were recorded with a Bruker Ultrashield Plus FT-NMR 600 MHz ADVANCE III equipped with a Cryoprobe™ QCI $^{1}H/^{19}F–^{13}C/^{15}N–D$ with a SampleJet™ autosampler with temperature control. For each sample, the probe was automatically locked, tuned, matched and shimmed. Before measurement, samples were kept for 5 min inside the NMR probehead for temperature equilibration at 298 °K.

Two NMR spectra were recorded for each sample: a mono-dimensional (1D) ^{1}H NMR spectrum with standard pulse sequence water suppression (noesygppr1d, Bruker), with 128 scans, 64 k data points, a spectral width of 20 ppm, an acquisition time of 2.7 s, a relaxation delay of 4 s and a mixing time of 100 ms and a 1D ^{1}H spin−echo Carr−Purcell−Meiboom−Gill sequence [17, 18] (cpmgpr1d, Bruker), to suppress large NMR signals arising from high molecular weight molecules (i.e. serum proteins). In the cpmg experiments the total echo time was to 51.2 ms consisting of 128 repetitions with a τ time of 200 μs and a 180° pulse of approximately 36 μs. Each spectrum was recorded with total of 128 scans 64k data points, a repetition time of 4 s and an acquisition time of 2.7 s. The free induction decay was multiplied with an exponential window function with 0.3 Hz line broadening prior to Fourier transformation.
All the ^{1}H NMR chemical shifts are referenced to the TSP signal.

Metabolomics analysis:

The spectra were analyzed for the metabolomics studies by Workflow4Metabolomics W4M 3.0 (http://workflow4metabolomics.org) an online free infrastructure [19] built on the Galaxy environment which include from spectra processing, bucketing, and normalization to statistical multivariate (i.e. Principal Component Analysis PCA, Partial least squares discriminant analysis PLS-DA) and univariate analysis (i.e. One Way analysis with Multiple Responses). For the bucketing we used a width of 0.04 ppm, and the samples were normalized based on the total intensity: each bucket integration is divided by the integration of the total spectrum. The significant NMR buckets, resulted from W4M analysis, were assigned by the free software version NMR Suite 8.2 (Chenomx, Edmonton, Canada) in combination with the Human Metabolome Database (http://www.hmdb.ca). The identified metabolites were quantified using the PULCON method [20] and TOPSPIN 3.2 Bruker software. One-way analysis of variance (t-test) was then performed on relative consumption/production of the identified metabolites of the controls and treatments by GraphPad Prism 5 software, and values within a range of $p < 0.05$ were regarded as significant.

1.4 Conclusions

Overall, our findings elucidated the biological effect of GO on primary human endothelial cells, suggesting a mechanism of GO-mediated cytotoxicity by taking advantages from an integrated approach based on cytotoxicity, angiogenesis, and metabolomics combined analysis. Our data demonstrated that GO has a cytotoxic effect on Huvec cells compared to HeLa cells, despite a comparable efficiency in the material uptake. Interestingly, our analysis showed that the GO acted at different cellular levels affecting various cellular processes, such as cell viability, oxidative stress, migration and angiogenesis, directly and indirectly. Actually, the steric hindrance of GO intracellular aggregates perturbed the correct assembly of cytoskeleton and the distribution of mitochondria, leading to defects in cell migration and an increase in oxidative stress. Moreover, preliminary metabolomics analysis allowed a comprehensive characterization of the metabolites altered by GO exposure, indicating the possible biochemical pathways underlying the angiogenesis processes. These results represent a step forward to understand GO-mediated toxicological mechanisms and, at the same time, a starting point for future investigations, which are actually ongoing in our group, to further elucidate the cellular processes affected by GO.

1.5 Bibliography

1. Stevens, T., Garcia, J. G. N., Shasby, D. M., Bhattacharya, J. & Malik, B. Mechanisms regulating endothelial cell barrier function. *Am J Physiol Lung Cell Mol Physiol* **279**, 419–422 (2011).
2. Potente, M., Gerhardt, H. & Carmeliet, P. Basic and therapeutic aspects of angiogenesis. *Cell* **146**, 873–887 (2011).
3. Novoselov, K. S., Fal'ko, V I, Colombo, L, Gellert, P. R., Schwab, M. G. Kim, K. A roadmap for graphene. *Nature* **490**, 192–200 (2012).
4. Chung, C., Kim Y.K., Shin D. Ryoo, S.R., Hong, B.H., Min, D. H. Biomedical applications of Graphene and Graphene Oxide. *Acc. Chem. Res.* **46**, 2211–2224 (2013).
5. Pratap, D., Eugenio, C., Singh, B. & Singh, S. Graphene oxide: An efficient material and recent approach for biotechnological and biomedical applications. *Mater. Sci. Eng. C* **86**, 173–197 (2018).
6. Lakshmanan, R. & Maulik, N. Graphene-based drug delivery systems in tissue engineering and nanomedicine. *Can. J. Physiol. Pharmacol.* **96(9)**, 869–878 (2018).
7. Fadeel, Bengt Flahaut, Emmanuel Bussy, Cyrill Merino, Sonia Va, Ester Mouchet, Florence Evariste, Lauris Gauthier, Laury Koivisto, Antti J Vogel, U. *et al.* Safety Assessment of Graphene-Based Materials: Focus on Human Health and the Environment. *ACS Nano* **12**, 10582–10620 (2018).
8. Mukherjee, S., Sriram, P., Barui, A. K. & Nethi, S. K. Graphene Oxides Show Angiogenic Properties. *Adv. Healthc. Mater.* **4**, 1722–1732 (2015).
9. Lai, P. X. *et al.* Ultrastrong trapping of VEGF by graphene oxide: Anti-angiogenesis application. *Biomaterials* **109**, 12–22 (2016).
10. Jeong, J. *et al.* In vivo toxicity assessment of angiogenesis and the live distribution of nano-graphene oxide and its PEGylated derivatives using the developing zebrafish embryo. *Carbon N. Y.* **93**, 431–440 (2015).
11. Xue, D., Wang, S., Joomun, M. U., Tan, Y. & Lin, S. Immunomodulatory properties of graphene oxide for osteogenesis and angiogenesis. *Int. J. Nanomedicine* **13**, 5799–5810 (2018).
12. Grodzik, M. & Winnicka, A. In vitro and in vivo effects of graphene oxide and reduced

graphene oxide on glioblastoma. *Int. J. Nanomedicine* **10**, 1585–1596 (2015).

13. Zhang, B., Wei, P., Zhou, Z. & Wei, T. Interactions of graphene with mammalian cells : Molecular mechanisms and biomedical insights. *Adv. Drug Deliv. Rev.* **105**, 145–162 (2016).
14. Matsui, J., Wakabayashi, T., Asada, M., Yoshimatsu, K. & Okada, M. Stem Cell Factor/c-kit Signaling Promotes the Survival, Migration, and Capillary Tube Formation of Human Umbilical Vein Endothelial Cells. *J. Biol. Chem.* **279**, 18600–18607 (2004).
15. Lam, K. H. *et al.* Nobiletin, a polymethoxylated flavonoid from citrus, shows anti-angiogenic activity in a zebrafish in vivo model and HUVEC in vitro model. *J. Cell. Biochem.* **112**, 3313–3321 (2011).
16. Cibecchini, G. *et al.* Antiangiogenic Effect of Graphene Oxide in Primary Human Endothelial Cells. *ACS Appl. Mater. Interfaces* **12**, 22507–22518 (2020).
17. Guarnieri, D. *et al.* Biotransformation and Biological Interaction of Graphene and Graphene Oxide during Simulated Oral Ingestion. *Small* **14**, 1–11 (2018).
18. Bramini, M. *et al.* Graphene Oxide Nanosheets Disrupt Lipid Composition, Ca^2 Homeostasis, and Synaptic Transmission in Primary Cortical Neurons. *ACS Nano* **10**, 7154–7171 (2016).
19. Chiacchiaretta, M. *et al.* Graphene Oxide Upregulates the Homeostatic Functions of Primary Astrocytes and Modulates Astrocyte-to-Neuron Communication. *Nano Lett.* **18**, 5827–5838 (2018).
20. Bramini, M. *et al.* An Increase in Membrane Cholesterol by Graphene Oxide Disrupts Calcium Homeostasis in Primary Astrocytes. *Small* **15**, 1900147 (2019).
21. Lv, M. *et al.* Effect of graphene oxide on undifferentiated and retinoic acid-differentiated SH-SY5Y cells line. *Nanoscale* **4**, 3861–3866 (2012).
22. Chang, Y. *et al.* In vitro toxicity evaluation of graphene oxide on A549 cells. *Toxicol. Lett.* **200**, 201–210 (2011).
23. Horváth, L. *et al.* Evaluation of the toxicity of graphene derivatives on cells of the lung luminal surface. *Carbon N. Y.* **64**, 45–60 (2013).
24. Kucki, M. *et al.* Interaction of graphene-related materials with human intestinal cells: an in vitro approach. *Nanoscale* **8**, 8749 (2016).
25. Zheng, P. & Wu, N. Fluorescence and Sensing Applications of Graphene Oxide and Graphene Quantum Dots: A Review. *Chem. Asian J* **12**, 2343–2353 (2017).

26. Guarnieri, D. *et al.* Transport across the cell-membrane dictates nanoparticle fate and toxicity: a new paradigm in. *Nanoscale* **6**, 10264–10273 (2014).
27. Durymanov, M., Kroll, C., Permyakova, A. & Reineke, J. Role of Endocytosis in Nanoparticle Penetration of 3D Pancreatic Cancer Spheroids. *Mol. Pharm.* **16**, 1074–1082 (2019).
28. Sahay, G., Alakhova, D. Y. & Kabanov, A. V. Endocytosis of nanomedicines. *J. Control. Release* **145**, 182–195 (2010).
29. Turco, A. *et al.* Sputtering-Enabled Intracellular X‑ray Photoelectron Spectroscopy: A Versatile Method To Analyze the Biological Fate of Metal Nanoparticles. *ACS Nano* **12**, 7731–7740 (2018).
30. Guarnieri, D. *et al.* Particle size affects the cytosolic delivery of membranotropic peptide-functionalized platinum nanozymes. *Nanoscale* **9**, 11288–11296 (2017).
31. Tokmakov, A. A. & Sato, K. Activity and intracellular localization of senescence-associated β-galactosidase in aging Xenopus oocytes and eggs. *Exp. Gerontol.* **119**, 157–167 (2019).
32. Chazotte, B. & Chazotte, B. Labeling Lysosomes in Live Cells with LysoTracker Labeling Lysosomes in Live Cells with LysoTracker. *Cold Spring Harb Protoc* (2011).
33. Liu, Y., Liu, C. Y. & Liu, Y. Investigation on fluorescence quenching of dyes by graphite oxide and graphene. *Appl. Surf. Sci.* **257**, 5513–5518 (2011).
34. Treossi, E., Melucci, M., Liscio, A., Gazzano, M. & Samorı, P. High-Contrast Visualization of Graphene Oxide on Dye-Sensitized Glass, Quartz, and Silicon by Fluorescence Quenching. *J. AM. CHEM. SOC.* **131**, 15576–15577 (2009).
35. Gao, L. *et al.* Graphene oxide–DNA based sensors. *Biosens. Bioelectron.* **60**, 22–29 (2014).
36. Feng, B. *et al.* A Graphene Oxide-Based Fluorescent Biosensor for the Analysis of Peptide–Receptor Interactions and Imaging in Somatostatin Receptor Subtype 2 Overexpressed Tumor Cells. *Anal. Chem.* **85**, 7732–7737 (2013).
37. Rezzola, S. *et al.* In vitro and ex vivo retina angiogenesis assays. *Angiogenesis* **17**, 429–442 (2014).
38. Guarnieri, D., Malvindi, M. A., Belli, V., Pompa, P. P. & Netti, P. Effect of silica nanoparticles with variable size and surface functionalization on human endothelial cell viability and angiogenic activity. *J. Nanoparticle Res.* **16**, 2229 (2014).

39. Griendling, K. K. Endothelial Cell Migration During Angiogenesis. *Circ. Res.* **100**, 782–794 (2007).
40. Clainche, C. L. E. & Carlier, M. Regulation of Actin Assembly Associated With Protrusion and Adhesion in Cell Migration. *Physiol Rev* **88**, 489–513 (2019).
41. Zhou, H. *et al.* The inhibition of migration and invasion of cancer cells by graphene via the impairment of mitochondrial respiration. *Biomaterials* **35**, 1597–1607 (2014).
42. Hussein, K. H., Abdelhamid, H. N. & Zou, X. Ultrasonicated graphene oxide enhances bone and skin wound regeneration. *Mater. Sci. Eng. C* **94**, 484–492 (2019).
43. Lu, B. *et al.* Graphene-based composite materials beneficial to wound healing. *Nanoscale* **4**, 2978–2982 (2012).
44. Fan, Z. *et al.* A Novel Wound Dressing Based on Ag/Graphene Polymer Hydrogel: Effectively Kill Bacteria and Accelerate Wound Healing. *Adv. Funct. Mater.* **24**, 3933–3943 (2014).
45. Labrousse, A M Zappaterra, M D Rube, D A Bliek, V. Der *et al.* Cell Migration: Integrating Signals from Front to Back. *Science (80-.).* **302**, 1704–1710 (2003).
46. Frontiñán-rubio, J. *et al.* Differential effects of graphene materials on the metabolism and function of human skin cells†. *Nanomedicine* **10**, 11604–11615 (2018).
47. Luca, E. De *et al.* Multifunctional Platinum@BSA−Rapamycin Nanocarriers for the Combinatorial Therapy of Cerebral Cavernous Malformation. *ACS omega* **3**, 15389–15398 (2018).
48. Green, D. R. & Reed, J. C. Mitochondria and Apoptosis. *Science (80-.).* **281**, 1309–1313 (1998).
49. Grif, E. J. & Rutter, G. A. Mitochondrial calcium as a key regulator of mitochondrial ATP production in mammalian cells. *Biochim. Biophys. Acta* **1787**, 1324–1333 (2009).
50. Gunter, T. E., Buntinas, L., Sparagna, G., Eliseev, R. & Gunter, K. Mitochondrial calcium transport: mechanisms and functions. *Cell Calcium* **28**, 285–296 (2000).
51. Dan, J., Aj, L., Zhang, X., Soldati, T. & V, P. E. B. Redox Biology Reactive oxygen species and mitochondria: A nexus of cellular homeostasis. *Redox Biol.* **6**, 472–485 (2015).
52. Minin, A. A. *et al.* Regulation of mitochondria distribution by RhoA and formins. *J. Cell Sci.* **119**, 659–670 (2005).
53. Johnson, L. V, Walsh, M. L. & Chen, L. A. N. B. Localization of mitochondria in living cells with rhodamine 123. *Cell Biol.* **77**, 990–994 (1980).

54. Boldogh, I. R. & Pon, L. A. Interactions of mitochondria with the actin cytoskeleton. *Biochim. Biophys. Acta* **1763**, 450–462 (2006).
55. Dluya, T., Bernstein, P. S. & Jahng, W. J. Altered Cytoskeleton as a Mitochondrial Decay Signature in the Retinal Pigment Epithelium. *Protein J.* **35**, 179–192 (2016).
56. Rohlenova, K., Veys, K., Miranda-santos, I., Bock, K. De & Carmeliet, P. Endothelial Cell Metabolism in Health and Disease. *Trends Cell Biol.* **28**, 224–236 (2018).
57. Treps, L., Conradi, L., Harjes, U. & Carmeliet, P. Manipulating Angiogenesis by Targeting Endothelial Metabolism: Hitting the Engine Rather than the Drivers-A New Perspective? *Pharmacol Rev* 872–887 (2016).
58. Beckonert, O. *et al.* Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2**, 2692–2703 (2007).
59. Kostidis, S., Addie, R. D., Morreau, H., Mayboroda, O. A. & Giera, M. Quantitative NMR analysis of intra-and extracellular metabolism of mammalian cells: A tutorial. *Anal. Chim. Acta* **980**, 1–24 (2017).
60. Emwas, A. *et al.* NMR Spectroscopy for Metabolomics Research. *Metabolites* **9**, 2–39 (2019).
61. Xiao, W., Wang, R., Handy, D. E. & Loscalzo, J. NAD(H) and NADP(H) Redox Couples and Cellular Energy Metabolism. *Antioxid. redox Signal.* **28**, 1–56 (2018).
62. Circu, M. L. & Aw, T. Y. Free Radical Biology & Medicine Reactive oxygen species , cellular redox systems , and apoptosis. *Free Radic. Biol. Med.* **48**, 749–762 (2010).
63. Oberkersch, R. E. & Santoro, M. M. Role of amino acid metabolism in angiogenesis. *Vascul. Pharmacol.* **112**, 17–23 (2019).
64. Min, H., Lee, D., Phuoc, N., Won, S. & Hill, J. Uptake of nanopolystyrene particles induces distinct metabolic pro fi les and toxic effects in Caenorhabditis elegans. *Environ. Pollut.* **246**, 578–586 (2019).
65. Kim, B., Li, J., Jang, C. & Arany, Z. Glutamine fuels proliferation but not migration of endothelial cells. *EMBO J.* **36**, 2321–2333 (2017).
66. Sidney M Morris Jr. Recent advances in arginine metabolism: roles and regulation of the arginases. *Br. J. Pharmacol.* **157**, 922–930 (2009).
67. Cooke, J. P. NO and angiogenesis. *Atheroscler. Suppl.* **4**, 53–60 (2003).
68. Li, H. U. I. *et al.* Regulatory role of arginase I and II in nitric oxide, polyamine, and

proline syntheses in endothelial cells. *Am J Physiol Endocrinol Metab* **280**, 75–82 (2001).

List of publications

Cibecchini, G., Cecere, P., Tumino, G., Morcia, C., Ghizzoni, R., Carnevali, P., Terzi, V., Pompa, P.P. A naked-eye, fast assay for varietal traceability in durum wheat production chain. Foods, 2020.

Cibecchini, G., Veronesi, M., Catelani, T., Bandiera, T., Guarnieri, D., Pompa, P.P. Antiangiogenic Effect of Graphene Oxide in Primary Human Endothelial Cells. ACS Applied Materials & Interfaces, 2020.

Pisani, A., Donno, R., Gennari, A., Cibecchini, G., Catalano, F., Marotta, R., Pompa, P.P., Tirelli, N., Bardi, G. CXCL12-PLGA/Pluronic Nanoparticle Internalization Abrogates CXCR4-Mediated Cell Migration. Nanomaterials 2020.

Mastronardi, V., Udayan, G., Cibecchini, G., Brescia, R., Fichthorn, KA., Pompa, P.P. Synthesis of Citrate-Coated Penta-twinned Palladium Nanorods and Ultrathin Nanowires with a Tunable Aspect Ratio. ACS Applied Materials & Interfaces, 2020.

Cagliani, R., Gatto, F., Cibecchini, G., Marotta, R., Catalano, F., Sanchez-Moreno, P., Pompa, P.P., Bardi, G. CXCL5Modified Nanoparticle Surface Improves CXCR2(+) Cell Selective Internalization. Cells, 2019.