Università degli Studi di Padova

Dipartimento di Scienze del Farmaco

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE FARMACOLOGICHE
INDIRIZZO: FARMACOLOGIA, TOSSICOLOGIA E TERAPIA
CICLO XXVII

DEVELOPMENT OF NEW MOLECULAR DIAGNOSTIC TESTS FOR PERSONALIZED MEDICINE: ANALYSIS OF *IL28B* POLYMORPHISMS AND *MPL/CALR* MUTATIONS

Direttore della Scuola: Ch.mo Prof. Pietro Giusti
Coordinatore d'indirizzo: Ch.mo Prof. Pietro Palatini
Supervisore interno: Ch.mo Prof. Andrea Cignarella
Supervisore esterno: Dott. Dino Paladin

Dottoranda: Ilaria Schiavo

CONTENTS

1 RIASSUNTO

L'attività di ricerca di questo Dottorato ha avuto come obiettivo quello di sviluppare test genetici di interesse commerciale nell'ambito della medicina personalizzata.

La prima parte di questo progetto ha riguardato lo sviluppo di due saggi per la discriminazione allelica di due polimorfismi a singolo nucleotide (SNP) vicini al locus di *IL28B* che svolgono un ruolo fondamentale nella gestione dei pazienti con epatite C cronica. L'epatite C è uno dei maggiori problemi di salute pubblica a livello globale; si stima infatti che vi siano circa 170 milioni di individui affetti da epatite C al mondo. Per molti anni la terapia definita come "standard of care" (SOC) per la cura dell'epatite C si è basata sulla somministrazione combinata di interferone alfa peghilato e ribavirina (PEG-IFN/RBV). Con questa terapia, però, circa l'80% dei pazienti infettati dai genotipi 2 e 3 di HCV e solo il 50% di quelli infettati dai genotipi 1 e 4 erano in grado di raggiungere una risposta virologica sostenuta (SVR) con l'eradicazione completa del virus. Negli ultimi anni, diversi studi hanno dimostrato in modo indipendente che fattori genetici legati all'ospite, quali il genotipo dei polimorfismi rs12979860 e rs8099917 a monte del gene *IL28B*, sono correlati alla probabilità di raggiungere SVR. Si è infatti osservato che i pazienti omozigoti per l'allele favorevole in entrambi i polimorfismi (CC per rs12979860 e TT per rs8099917) raggiungevano una *clearance* virale spontanea nelle infezioni acute oppure una SVR dopo il trattamento in quelle croniche, con una probabilità 2-3 volte maggiore rispetto ai pazienti eterozigoti o omozigoti per l'allele sfavorevole (CT o TT per rs12979860 e TG o GG per rs8099917). Recentemente, l'uso di nuovi agenti antivirali ad azione diretta (DAA) che inibiscono le proteasi coinvolte nel ciclo replicativo di HCV, in combinazione con la terapia standard ha permesso di ottenere un aumento significativo dei tassi di SVR nei pazienti indipendentemente dal genotipo virale dell'infezione. Due di questi, boceprevir e telaprevir, sono stati approvati dall'FDA nel 2011 e molti altri sono attualmente in fase di sviluppo per arrivare a definire una nuova terapia SOC non più basata sull'utilizzo di interferone. Nonostante l'effetto dei genotipi dei polimorfismi di *IL28B* sulla cinetica virale sembri risultare indebolito dall'uso di questi potenti agenti antivirali nella terapia, dati sperimentali indicano che le varianti genetiche di questi SNP rimarranno fortemente informative anche nel contesto dei futuri regimi *IFN-free*. Pertanto, la genotipizzazione degli SNP di *IL28B* prima della terapia è fondamentale allo scopo di evitare un trattamento inefficace caratterizzato non solo da tempi lunghi e costi elevati ma soprattutto da pesanti effetti collaterali (come ad esempio sindrome di tipo influenzale, anormalità ematologiche ed eventi avversi neuropsichiatrici).

Lo scopo di questo studio è stato quindi sviluppare due test per la genotipizzazione di rs12979860 e rs8099917 basandosi sul metodo della *real-time* PCR con sonde TaqMan®. Sono stati disegnati i *primer* e le sonde per l'ibridazione specifica con la sequenza target di DNA. Le sonde sono state marcate con fluorofori specifici per ciascun allele. La specificità dell'appaiamento dei *primer* alla sequenza genomica di interesse è stata verificata mediante il

tool bioinformatico di allineamento *Blast* mentre l'assenza di strutture secondarie o *folding* indesiderati è stata confermata con il programma *mfold*. Utilizzando campioni clinici con genotipo noto degli SNP di *IL28B,* i controlli positivi di entrambi i saggi sono stati preparati mediante clonaggio della regione target all'interno di un vettore plasmidico. La corretta inserzione della sequenza target nel plasmide è stata accertata mediante sequenziamento di Sanger. Il protocollo di reazione è stato ottimizzato confrontando tra loro diverse condizioni sperimentali con cui sono state testate diverse *master mix* di amplificazione e concentrazioni di *primer* e sonde. I protocolli di reazione sono stati ottimizzati per funzionare con i seguenti strumenti: Applied Biosystems StepOne™ e StepOnePlus™ Real-Time PCR System, Applied Biosystems 7500 Fast e 7500 Fast Dx Real-Time PCR Systems, Applied Biosystems 7300 Real-Time PCR System, Bio-Rad Dx Real-Time System e Bio-Rad CFX96™ Real-Time PCR Detection System. Sensibilità e specificità diagnostiche sono state calcolate analizzando per ciascun saggio circa 200 campioni precedentemente genotipizzati con un IVD di riferimento, e sono risultate rispettivamente pari al 99,50% per rs12979860 e al 99,48% per rs8099917. Il *range* di concentrazioni del DNA entro il quale le *performance* del test rimangono invariate è stato determinato usando campioni di DNA a diverse concentrazioni. Si è visto che entrambi i saggi assegnano correttamente il genotipo a campioni con concentrazioni di DNA comprese tra 2 e 250 ng/µL. In seguito ad uno studio di stabilità durante il quale sono state misurate le *performance* del test a vari intervalli di tempo, è stato possibile attribuire ai reagenti di entrambi i saggi una *shelf-life* di 12 mesi.

I test messi a punto in questo studio, REALQUALITY RS-IL28B rs12979860 e REALQUALITY RQ-IL28B rs8099917, sono stati infine notificati al Ministero della Salute e immessi nel mercato come kit commerciali CE-IVD.

Questi dispositivi, permettendo la genotipizzazione dei due SNP di *IL28B* nella stessa seduta, rappresentano uno dei sistemi commerciali più completi per la gestione dei pazienti con HCV. Inoltre, entrambi i kit hanno dimostrato di essere molto sensibili ed affidabili anche in condizioni non ottimali (ad esempio dopo ripetuti cicli di gelo-scongelo della mix di amplificazione oppure utilizzando campioni degradati).

La seconda parte di questo progetto ha avuto come scopo lo sviluppo di due saggi, uno per la rilevazione delle mutazioni note nell'esone 9 di *CALR*, l'altro per la rilevazione e la semi-quantificazione delle due mutazioni più frequenti di *MPL* (W515L e W515K). Le mutazioni somatiche nel gene codificante per la calreticulina (*CALR*) scoperte recentemente, assieme a quelle nei geni *JAK2* e *MPL*, sono ritenute essere le mutazioni "driver" alla base della sottoclasse di neoplasie mieloproliferative (MPN) BCR-ABL1-negative. La policitemia vera (PV), la trombocitemia essenziale (ET) e la mielofibrosi primaria (PMF) fanno parte di questo gruppo di MPN. Le mutazioni in *JAK2*, *CALR* e *MPL*, pur non essendo specifiche per la malattia, sono

presenti nella maggior parte dei pazienti con tali neoplasie in modo mutuamente esclusivo. La mutazione V617F nell'esone 14 di *JAK2* è la più frequente in tutte le patologie considerate, essendo presente nella quasi totalità dei pazienti con PV (nei rimanenti casi si riscontrano mutazioni nell'esone 12 di *JAK2*) e nella maggior parte dei pazienti con ET o PMF. Le mutazioni di *CALR* sono presenti con una frequenza inferiore rispetto a *JAK2* V617F nelle ET e PMF. Infine, le mutazioni nel codone 515 dell'esone 10 di *MPL* sono quelle presenti con la minor frequenza nelle ET e PMF. Le mutazioni in questi tre oncogeni, oltre che per la diagnosi e la prognosi, vengono anche utilizzate per monitorare la progressione della malattia verso forme più aggressive, quali leucemia e mielofibrosi. I criteri della WHO (World Health Organization) per la diagnosi delle MPN BCR-ABL1-negative, in particolare ET e PMF, sono attualmente in stato di revisione a seguito della proposta avanzata nel 2014 per l'inserimento delle mutazioni di *CALR* tra i criteri diagnostici maggiori accanto alle mutazioni di *JAK2* (V617F) e *MPL* già presenti. È stato inoltre osservato che le mutazioni di *CALR*, rispetto a *JAK2* V617F e alle mutazioni di *MPL*, sono associate ad un fenotipo caratterizzato da livelli inferiori di emoglobina, minor numero di leucociti, maggior numero di piastrine e maggiore sopravvivenza priva di trombosi.

Siccome AB ANALITICA ha già sviluppato dei test per l'identificazione, la semi-quantificazione e la quantificazione assoluta dell'allele *JAK2* V617F, questo studio aveva l'obiettivo di sviluppare due test per la rilevazione delle mutazioni in *CALR* ed *MPL* al fine di completare il pannello delle principali mutazioni associate alle MPN BCR-ABL1-negative. Dovendo rilevare mutazioni di tipo diverso (inserzioni e/o delezioni in *CALR* e due mutazioni puntiformi in *MPL*), si è scelto di usare per i due saggi due differenti tecnologie, vale a dire *end-point* PCR con elettroforesi in gel d'agarosio per le mutazioni di *CALR* e *real-time* PCR per le mutazioni W515L e W515K di *MPL*. Seguendo il modello descritto in precedenza, sono stati disegnati i *primer* di entrambi i test e le sonde TaqMan® MGB per il saggio di MPL W515L/K ed è stato verificato il loro appaiamento specifico alla regione target. I controlli positivi plasmidici di entrambi i sistemi sono stati preparati utilizzando kit commerciali per il clonaggio. Il protocollo del saggio per la rilevazione delle mutazioni di *CALR* è stato ottimizzato dopo aver testato diverse condizioni sperimentali sia per la reazione di amplificazione (variando *master mix*, concentrazione dei *primer*, numero di cicli di amplificazione) che per la visualizzazione (variando tipo di agarosio, densità del gel, agente intercalante del DNA, volumi da caricare nel gel). Una fase preliminare di messa a punto del saggio per la rilevazione delle mutazioni di *MPL* è stata effettuata sullo strumento StepOnePlus™ (Applied Biosystems) testando diverse *master mix* e concentrazioni di *primer* e sonde.

Sensibilità e specificità diagnostiche del saggio delle mutazioni di *CALR* sono state calcolate utilizzando un totale di 65 campioni clinici di cui 36 *wild-type* e 29 mutati. Tutte le mutazioni analizzate in questo studio (Tipo-1, Tipo-2 e mutazioni rare) sono state precedentemente determinate mediante sequenziamento bidirezionale con cui si è visto che alterano la lunghezza

della sequenza di almeno 4 paia di basi rispetto alla sequenza *wild-type*. Il saggio ha dimostrato di saper discriminare correttamente lo status di tutti i campioni testati (mutato vs *wild-type*), perciò gli è stata attribuita un'accuratezza (osservato/atteso) del 100%. Testando diluizioni seriali di allele mutato in un *background* di allele *wild-type* è stato identificato il limite di rilevabilità del saggio pari al 10% di allele mutato. Il test ha inoltre dimostrato di discriminare correttamente campioni con concentrazioni di DNA tra 25 e 100 ng/reazione. Nonostante lo studio di stabilità sia attualmente in corso, è stata attribuita ai reagenti una *shelf-life* di 6 mesi secondo la norma europea EN 13640:2002[1]. I *primer* per il sequenziamento di Sanger (facoltativo) che consente l'identificazione della mutazione sono risultati altamente specifici per la regione *target*.

Nelle prove preliminari del saggio per la rilevazione delle mutazioni W515L/K di *MPL*, è stato stabilito un *cut-off* dell' 1,5% di allele mutato con cui sono stati correttamente discriminati 11/11 campioni *wild-type*. Il saggio è stato in grado di rilevare e semi-quantificare correttamente 1/1 campione con la mutazione W515K e 4/4 campioni con la mutazione W515L. Utilizzando però campioni con una mutazione diversa nel codone 515 o nelle sue vicinanze non è stato riscontrato alcun segnale di amplificazione relativo alla sequenza mutata, indice di una elevata specificità delle sonde per le due mutazioni W515L e W515K.

In conclusione, questo studio ha portato allo sviluppo del dispositivo GENEQUALITY CALR MUTATION che è stato notificato al Ministero della Salute ed è il primo kit CE-IVD che permette di rilevare ed identificare le mutazioni nell'esone 9 del gene *CALR*.

Per quanto riguarda il saggio di MPL W515L/K, è stato deciso ed è attualmente in corso il disegno di nuove sonde per la rilevazione di tutte le mutazioni nel codone 515 di *MPL.*

2 ABSTRACT

The research during this doctorate focused on the development of commercial genetic tests in the context of personalized medicine.

In the first part of this project two assays were developed that allow allelic discrimination of two informative single nucleotide polymorphisms (SNPs) near the *IL28B* locus, which are useful for the management of patients with chronic hepatitis C infection. Hepatitis C is a significant health problem worldwide, with approximately 170 million infected people. For many years the standard of care (SOC) for HCV therapy has consisted in the administration of pegylated interferon α and ribavirin (PEG-IFN/RBV). With this therapy, however, around 80% of patients infected by genotype 2 or 3 of HCV and only 50% of those infected by genotype 1 or 4 reached a sustained virological response (SVR), which indicates complete eradication of the virus. In the last years, several studies have demonstrated that, among others, host genetic factors, such as different genotypes of rs12979860 and rs8099917 SNPs located upstream of the *IL28B* gene, determine the probability of attaining SVR. Patients homozygous for the favorable allele in both SNPs (CC for rs12979860 and TT for rs8099917) had a two- to three-fold higher likelihood to achieve spontaneous viral clearance in acute infections or SVR after treatment compared to patients heterozygous or homozygous for the unfavorable alleles (CT or TT for rs12979860 and TG or GG for rs8099917). Recently developed direct-acting antiviral agents (DAAs) inhibiting proteases involved in viral replication have been shown to significantly increase SVR rates in patients independently of the viral genotype, when combined with standard therapy. Two of those DAAs, boceprevir and telaprevir, were approved by the FDA in 2011 and many others are currently under investigation in order to develop an IFN-free SOC. Although the *IL28B* SNP genotype may have less effect on viral kinetics in the new combination therapies, there is evidence that it will continue to be relevant also in the context of IFN-free regimens. Therefore, genotyping of *IL28B* SNPs prior to therapy will remain fundamental to avoid ineffective treatments that are not only long and costly but, most importantly, are associated with significant side effects (e.g., flu-like syndrome, hematologic abnormalities and adverse neuropsychiatric events).

The objective of this study was the development of two real-time PCR based tests for rs12979860 and rs8099917 genotyping using TaqMan® probes. Primers and probes that specifically bind the target region were designed. Probes were labeled with different fluorophores for each allele. Primer specificity to the target genomic sequence was determined using *Blast* and the probability of unwanted folding and secondary structures was checked with *mfold*. Starting from samples with a known genotype, positive controls for both assays were prepared by cloning the targeted gene region into a plasmid vector. Correct insertion of the target sequence into the plasmid was checked with Sanger sequencing. The reaction protocol was optimized by testing and comparing several amplification conditions with different master

mixes and primer/probe concentrations. The reaction protocols were optimized to run on the following real-time PCR systems: Applied Biosystems StepOne™ and StepOnePlus™ Real-Time PCR System, Applied Biosystems 7500 Fast and 7500 Fast Dx Real-Time PCR Systems, Applied Biosystems 7300 Real-Time PCR System, Bio-Rad Dx Real-Time System and Bio-Rad CFX96™ Real-Time PCR Detection System. Both *IL28B* diagnostic devices were demonstrated to have high diagnostic specificity and sensitivity, 99.50% and 99.48% for rs12979860 and rs8099917, respectively, by analyzing around 200 samples that had been previously genotyped with a reference IVD. Using different concentrations of clinical samples, the assays were shown to work with DNA concentrations of 2 to 250 ng/µL. A reagent shelf-life of 12 months was determined following a stability study that assessed assay performances at different time points.
At the end of this study, the final assays REALQUALITY RS-IL28B rs12979860 and REALQUALITY RQ-IL28B rs8099917 were notified with the Ministry of Health and placed on the market as CE-IVD commercial kits. By allowing simultaneous genotyping of both *IL28B* SNPs, these devices represent one of the most complete systems available for HCV patients management. Moreover, both kits have been shown to be highly sensitive and robust even under suboptimal conditions (after several freeze-thaw cycles of the amplification mix and with degraded DNA samples).

The second part of this project focused on the development of two assays for the detection of the complete panel of mutations found in exon 9 of the *CALR* gene and for the detection and semi-quantification of the two most frequent mutations (W515L and W515K) in the *MPL* gene. Recently discovered somatic mutations in the gene encoding for calreticulin (*CALR*) have been demonstrated – together with *JAK2* and *MPL* mutations – to be the driver mutations responsible for a subclass of myeloproliferative neoplasms (BCR-ABL1-negative MPN). To this group of MPNs belong polycythemia vera (PV), essential thrombocytemia (ET) and primary myelofibrosis (PMF). Although mutations in these three genes are not disease-specific, they have been shown to be mutually exclusive and present in the vast majority of patients with these types of neoplasms. According to relative mutation frequencies the V617F mutation in exon 14 of *JAK2* is at the first place, being present in almost all PV cases (the remaining cases carry mutations in exon 12 of this gene) and in most cases of ET and PMF. *CALR* mutations are the second most common genetic alterations in ET and PMF, whereas the mutations in codon 515 of exon 10 of the *MPL* gene are the third most frequent mutation encountered in ET and PMF. Mutations in *JAK2*, *CALR* and *MPL* are considered to have high diagnostic and prognostic value, and are used to monitor disease progression towards myelofibrosis and leukemia. In 2014, a new revision of the World Health Organization (WHO) criteria for diagnosis of BCR-ABL1-negative MPNs has been proposed including *CALR* mutations as major criterion for ET and PMF together with the above-mentioned *JAK2* (V617F) and *MPL* mutations. Compared to patients with *JAK2* V617F and

MPL mutations, carriers of *CALR* mutations have lower hemoglobin levels, lower leukocyte counts, higher platelet counts and improved thrombosis-free survival.

Since tests for the identification, semi-quantification and absolute quantification of the *JAK2* V617F allele are already available at AB ANALITICA, this study focused on the development of tests for detection of mutations in the *CALR* and *MPL* genes. Different molecular biology technologies were used for these two assays. The test for CALR mutations detection was based on end-point PCR and agarose gel electrophoresis, whereas the assay for analysis of the two most common MPL mutations (W515L and W515K) used real-time PCR. Primers for both assays and TaqMan® MGB probes of the MPL W515L/K assay were designed and checked as described above. Plasmid positive controls were prepared using commercial cloning kits. The protocol of the CALR MUTATION assay was optimized by testing several amplification settings and conditions, different master mix, primer concentration and number of amplification cycles, as well as by refining the visualization step (variation of agarose type, agarose density, DNA-intercalating agent, loading volumes). The MPL W515L/K assay was tested on the Applied Biosystems StepOnePlus™ Real-Time PCR System with different master mixes and concentrations of primers and probes.

The diagnostic specificity and sensitivity of CALR MUTATION assay were determined on a total of 65 clinical samples: 36 wild-type samples and 29 mutated samples containing the two most common *CALR* mutations (Type-1 and Type-2) as well as the so-called rare *CALR* mutations with a sequence length that differs by at least 4 base pairs from the wild-type sequence. The assay was able to correctly discriminate all samples tested, hence the test was assigned a diagnostic specificity and sensitivity of 100%. In addition, the assay correctly typed samples in a DNA concentration range of 25 to 100 ng/reaction and the measured detection limit was 10% of mutated allele on wild-type background. Although the stability study is still on-going, a shelf-life of 6 months was assigned to the device according to the European Standard EN 13640:2002[1]. The primers for optional Sanger sequencing included with the device were shown to be highly specific for the target region.

During the development of the assay for detection of W515L/K mutations in *MPL*, a cut-off of 1.5% of mutated allele was established. This cut-off allowed the correct identification of all tested wild-type samples (11 in total). The assay was able to detect and correctly semi-quantify 1/1 sample with the W515K mutation and 4/4 samples with the W515L mutation. When using samples with other mutations in or close to codon 515, no amplification signal of the mutated sequence was generated, due to the high specificity of the probes for the W515L and W515K mutations.

As a conclusion of this study, the completed device GENEQUALITY CALR MUTATION was notified with the Ministry of Health and placed on the market as the first commercial CE-marked IVD device for detection and identification of mutations in exon 9 of the *CALR* gene.

Concerning the MPL W515L/K assay, the design of new probes for the detection of all mutations in codon 515 of *MPL* was decided and is currently on going.

3 INTRODUCTION

3.1 PERSONALIZED MEDICINE

Over the last decades, the traditional "one-drug-fits-all" approach of medicine based on the clinical practice of "trials-and-errors" has rapidly given way to the "personalized medicine" approach based on the evidence that patients with similar symptoms respond differently to the same treatment according to information encoded by the individual's genetic makeup (Fig. 1).
The concept of personalized medicine is not new, but dates back many hundreds of years:

> "*It's far more important to know what person the disease has than what disease the person has"* (Hippocrates).

However, only in the 19th century, thanks to developments in chemistry, histochemistry and microscopy, scientists began to understand the causes underlying disease. From here, advances in science and technology increasingly influenced health care decisions. In the 20th century, the growth of the pharmaceutical and medical devices industry ushered in the age of genetics, imaging, and data mining. In the middle of the century, the focus on identifying key enzymes that play a role in inter-individual variation in drug metabolism and response laid the foundation for pharmacogenetics. Sequencing of the human genome at the turn of the 21st century triggered the transformation of personalized medicine from an idea to practice. Rapid developments in genomics, together with advances in a number of other areas, such as computational biology, medical imaging, and regenerative medicine, are creating the possibility for scientists to develop tools to truly personalize diagnosis and treatment. Despite the huge progress that has been made in medical fields, we are still far from understanding the biological mechanisms leading to inter-individual differences in response to treatment[2].
Encompassing many scientific disciplines, personalized medicine is a concept that can be described with definitions that range from the extremely broad to the very narrow. Over time, several definitions of "personalized medicine" have been proposed:

- *"The use of new methods of molecular analysis to better manage a patient's disease or predisposition to disease."*– Personalized Medicine Coalition
- *"Providing the right treatment to the right patient, at the right dose at the right time."* – European Union
- *"The tailoring of medical treatment to the individual characteristics of each patient."* – President's Council of Advisors on Science and Technology
- *"Health care that is informed by each person's unique clinical, genetic, and environmental information."* – American Medical Association
- *"A form of medicine that uses information about a person's genes, proteins, and environment to prevent, diagnose, and treat disease."* – National Cancer Institute, NIH

The goal of personalized medicine is to streamline clinical decision-making by selecting those patients that most likely will benefit from a certain treatment according to their genomic

information. The effects of personalized medicine consist of the reduction of healthcare costs for ineffective treatments and the minimization of drug-related adverse events with the increase of patient adherence to therapy.

Terms like “precision medicine”, “stratified medicine”, “targeted medicine” and “pharmacogenomics” are synonyms of “personalized medicine”. “Precision medicine”, defined by the National Academy of Sciences (NAS) as “the use of genomic, epigenomic, exposure and other data to define individual patterns of disease, potentially leading to better individual treatment”[3], is the concept closest to “personalized medicine”. “Stratification” of patients, with a particular disease and that will benefit from a certain treatment or have low risk to develop side effects into subgroups for determined characteristics, is a fundamental aspect of personalized medicine. “Pharmacogenomics” (PGx) – the study of variations of DNA and RNA characteristics as related to drug response – is a relatively recent and highly dynamic area of personalized medicine. Pharmacogenomics spawned from the convergence of advances in pharmacology and genomics and its worldwide diffusion takes advantage of the new high-throughput sequencing technologies that dramatically reduced the cost of sequencing. Pharmacogenomics, in a wider sense than pharmacogenetics that correlates the variations in a single candidate gene to a treatment response, seeks to explain interindividual differences in drug metabolism (pharmacokinetics) and physiological drug response (pharmacodynamics), identifying responders and non-responders to a therapy and predicting the efficacy and/or toxicity of a drug.

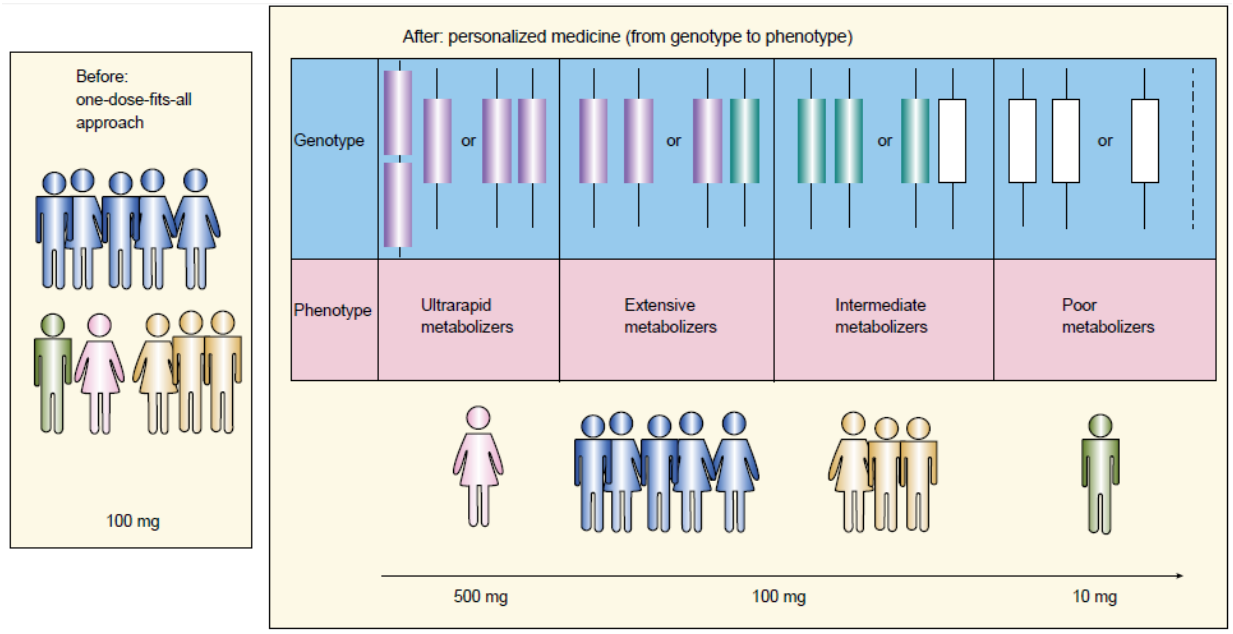


Fig. 1: Representation of the "one-dose-fits-all" approach versus "personalized medicine". The left panel shows a situation in which everyone gets the same dose of a drug, regardless of genotype. The right panel shows a personalized medicine approach in which the dose of the drug is selected based upon genotypical, and therefore phenotypical, variability of the metabolizing enzyme (Source:[2]).

The success of personalized medicine relies on the identification of predictive biomarkers and the development of diagnostic devices that allow their accurate detection and measuring. Many of the diagnostic tests used in personalized medicine are *in vitro* diagnostic devices (IVDs), which test human matrices for genetic alterations (e.g. mutations in oncogenes, allelic variations of SNPs related to drug response). IVD kits are developed by a device manufacturer and sold to labs, hospitals and physicians offices where the test kit is used to run the tests. In clinical practice, for example, many diagnostic tests are used to identify appropriate patients for a given therapy or patients who should not receive a particular therapy because of an increased risk of a serious side effect. Other commonly used tests help to characterize a disease or condition, such as cancer, to determine what type of treatment is potentially most appropriate.

To date, the FDA (Food and Drug Administration) has contributed to the purpose of personalized medicine with the introduction of information on genomic biomarkers (including gene variants, functional deficiencies, expression changes, chromosomal abnormalities, and others) in the label of more than 100 approved drugs[2].

The increasing number of published gene-disease association studies over the last years and the thousands of new drug targets unveiled by DNA sequencing and characterization of the human genome are effective indicators of the success of personalized therapeutics that has become part of routine clinical practice.

However, the beginning of the "personalized medicine era" has opened new issues that must be considered:

1) *Limited understanding of the intrinsic biology of disease* that is the most important limiting factor of this field;
2) *Common conditions involving multiple genes/biomarkers* that are complex to correlate and require huge investments in clinical research;
3) *An outdated disease classification system* and the need of a "new taxonomy" that describes diseases on the basis of the new biological insights in addition to traditional signs and symptoms;
4) *Lack of infrastructure* that serves to collect, analyze, integrate, share and mine the huge amount of information deriving from the new high-throughput sequencing methods;
5) *Clinical implementation of new diagnostics* in clinical practice with the need of tools that would help clinicians to adopt the appropriate test and to interpret the results afterwards;
6) *Investment uncertainties* due to smaller size of the targeted population;
7) *Access to personalized therapeutics* for patients that do not have the targeted biomarker and do not benefit from this therapy;
8) *Intellectual property rights* concerning patent protection for diagnostic tools, genes and biomarkers;
9) *Reimbursement policies* that have to be redefined to fit the changes that personalized medicine will bring to the healthcare system;
10) *Patient privacy and confidentiality* concerning the protection of patient genetic data.

3.2 HEPATITIS C VIRUS INFECTION

3.2.1 HCV EPIDEMIOLOGY AND TRANSMISSION

Hepatitis C is a disease with a significant global impact; according to the World Health Organization 130-170 million people are infected with the hepatitis C virus (HCV), corresponding to 2-2.5% of the world's total population. In Europe, it is estimated that there are 2-5 million HCV-positive persons with chronic hepatitis C being the most common chronic liver disease and the major cause of liver transplants. Since the vast majority of acute cases of hepatitis C are asymptomatic[4], it is difficult to determine the number of new HCV infections as well as the age of infection upon diagnosis. Nevertheless, it has to be assumed that the number of new infections has considerably decreased over the past decades thanks to the introduction of routine testing for HCV before blood transfusions in the early 1990s, soon after HCV discovery. In addition to blood transfusion other possible routes of infection exist:

- injection drug use with up to 70% of drug users showing seroprevalence of anti-HCV antibodies[5];
- organ transplantation;
- sexual habits with particular attention to risk factors like male homosexual intercourses, a high number of sex partners and unprotected intercourses;
- perinatal transmission in HCV RNA-positive mothers with an estimated risk lower than 5%[6];
- hemodialysis with particular attention to risk factors like blood transfusions, duration of hemodialysis, prevalence of HCV infection in the dialysis unit and type of dialysis;
- other rare transmission routes like needlestick injury for healthcare workers, tattooing and body piercing[7].

3.2.2 CLINICAL MANIFESTATIONS

HCV infection may be present in acute or chronic form. About 80% of acute HCV infections evolve into chronic disease that slowly progresses from an asymptomatic state to cirrhosis and hepatocellular carcinoma over many years.

ACUTE HEPATITIS C

During the first weeks after infection (incubation period), HCV RNA in blood cannot be detected by PCR. In this time period the level of aminotransferases begin to increase to 10-30 times the upper limit of normal (800U/L). Over the first months of infection, the majority of patients show no clinical sign of the disease[4] and only 25% of cases present jaundice at this stage. Other possible symptoms are malaise, nausea and right upper quadrant pain. After some weeks,

spontaneous viral clearance occurs in about 20% of cases. Fulminant hepatic failure in acute hepatitis C is a rare event, usually associated with concurrent chronic HBV infection[8].

CHRONIC HEPATITIS C

Hepatitis C becomes chronic after persistence of more than 6 months, which occurs in 80-100% of HCV infections[4,9]. Long-lasting establishment of the virus in the human body may be due to the ability of the HCV genome to rapidly mutate, which protects the virus from immune recognition. Host factors that have been associated with successful HCV clearance are HCV-specific CD4 T cell responses, high titers of neutralizing antibodies against HCV structural proteins, *IL28B* gene polymorphisms and specific *HLA*-DRB1 and -DQB1 alleles[10–12]. However, once chronic infection is established, there is a very low rate of spontaneous clearance. A common symptom in this phase of the disease is fatigue but also nausea, weakness, myalgia, arthralgia and weight loss may occur. Aminotransferase levels are usually slightly above the upper limit of normal (between 2 and 5 times)[13] but no correlation between concentrations of aminotransferases and liver histology has been observed[14].

EXTRAHEPATIC MANIFESTATIONS

Around 30 to 40% of patients with chronic hepatitis C have an extrahepatic manifestation of HCV[15]. There is a wide variety of extrahepatic manifestations described as being associated with HCV:

- hematologic manifestations (essential mixed cryoglobulinemia, lymphoma)
- autoimmune disorders (thyroiditis, presence of various autoantibodies)
- renal disease (membranoproliferative glomerulonephritis)
- dermatologic disease (porphyria cutanea tarda, lichen planus)
- diabetes mellitus

3.2.3 NATURAL HISTORY

Survival of chronic hepatitis C patients is generally not impaired until cirrhosis has developed. Chronic infection does not always lead to liver damage but the risk to develop chronic active hepatitis, cirrhosis and hepatocellular carcinoma is certainly enhanced (respectively by 23%, 51% and 5% more than 20 years after transfusion[16]).

CIRRHOSIS AND HEPATIC DECOMPENSATION

Complications of hepatitis C like decompensation and hepatocellular carcinoma occur almost exclusively in patients who have developed cirrhosis. Cirrhotic patients generally present asymptomatic hepatomegaly and/or splenomegaly, elevated serum bilirubin concentration, hyperalbuminemia or low platelet count. Less than half of cirrhosis cases are associated with

spider angioma, caput medusae, palmar erythema, testicular atrophy, or gynecomastia. The risk of progression to hepatic decompensation is estimated to be close to 5% per year in cirrhotics[17] and leads to ascites, followed by variceal bleeding, encephalopathy and jaundice. Roughly half of the patients survive over a 5-year period at this stage of the disease[18] and liver transplantation is the only effective therapy. The risk to develop hepatocellular carcinoma in cirrhotics is lower than 3% per year[19,20] and usually has a very poor outcome.

DISEASE PROGRESSION

Individual differences in disease progression may be associated with the following factors:

- age and gender: acquisition of HCV infection after the age of 40 to 55 and male gender may be associated to more rapid progression of liver injury[21]. In contrast, children have a relatively low risk of disease progression[22];
- ethnicity: disease progression appears to be slower and changes in liver histology less severe in African-Americans[23];
- HCV-specific cellular immune response;
- alcohol intake increases HCV replication, enhances the progression of chronic HCV, and accelerates liver injury[24];
- other host factors: genetic polymorphisms of certain genes might influence the fibrosis progression rate[25];
- viral coinfection: disease progression is faster in HCV patients coinfected with either HIV or HBV infections[13];
- geography and environmental factors;
- use of steroids that increases the HCV viral load;
- viral factors: infections with more genotypes have worse outcome as compared to monoinfection.

3.2.4 HCV VIRUS

HCV is a small-enveloped virus of the *Flaviviridae* family with a RNA-based genome (one single-stranded positive-sense molecule) of approximately 9.6 kb. Despite the huge genetic variability of this virus, which is able to rapidly mutate in response to environmental changes, six major genotypes have been described, each with a large number of subtypes[26,27]. Determination of the HCV genotype plays an important role for the initiation of anti-HCV treatment since the patient response to an antiviral drug regimen varies significantly with regard to the specific viral genotype of infection. For example, a patient infected with genotype 1 (GT1) of HCV is less responsive (i.e. HCV GT1 is more resistant) to standard therapy based on the combination of pegylated interferone α and ribavirin (PEG-IFN/RBV)[28,29].

In the bloodstream HCV virions may be present in various forms such as bound to very low density lipoproteins (VLDL), bound to low density lipoproteins (LDL), complexed with immunoglobulins, and free circulating[30–34]. It is possible that binding LDL serves HCV to enter hepatocytes[33,35] or as protection against neutralization by HCV-specific antibodies[36].

VIRAL STRUCTURE

HCV is a spherical virus of approximately 50 to 55 nm[37–39] with an outer layer comprising E1 and E2 proteins[39]. This outer layer surrounds the lipid bilayer that contains the viral nucleocapsid consisting of the HCV core (C) protein. An inner spherical structure with a diameter of approximately 30-35 nm has been observed[38] representing the nucleocapsid that harbours the genomic viral RNA[40].

The HCV genomic RNA molecule serves as mRNA for translation of viral proteins and contains a single open reading frame (ORF) coding for a precursor polypeptide of about 3000 amino acid residues (Fig. 2 A). During viral replication the polypeptide is cleaved by viral and host enzymes into 3 structural (C, E1, E2) and 7 non-structural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A, NS5B). The structural genes are located at the 5' terminus of the ORF and the coding regions for the non-structural proteins follow downstream. Structural proteins are essential components of HCV virions, whereas non-structural proteins are involved in RNA replication and virion morphogenesis. The ORF is flanked by 5' and 3' non-translated regions (NTR) containing highly conserved regions involved in the regulation of viral replication that make them perfect targets for antiviral therapeutics[13]. Both NTRs appear to work together in a long-range RNA-RNA interaction possibly resulting in temporary genome circularization[41].

The translation process leads to a precursor polyprotein (Fig. 2 B) with the C, E1, E2 and p7 proteins at the N-terminus that are subsequently cleaved by signal peptidases[42]. Non-structural proteins are processed by two virus-encoded proteases (NS2/NS3 and NS3/NS4A).

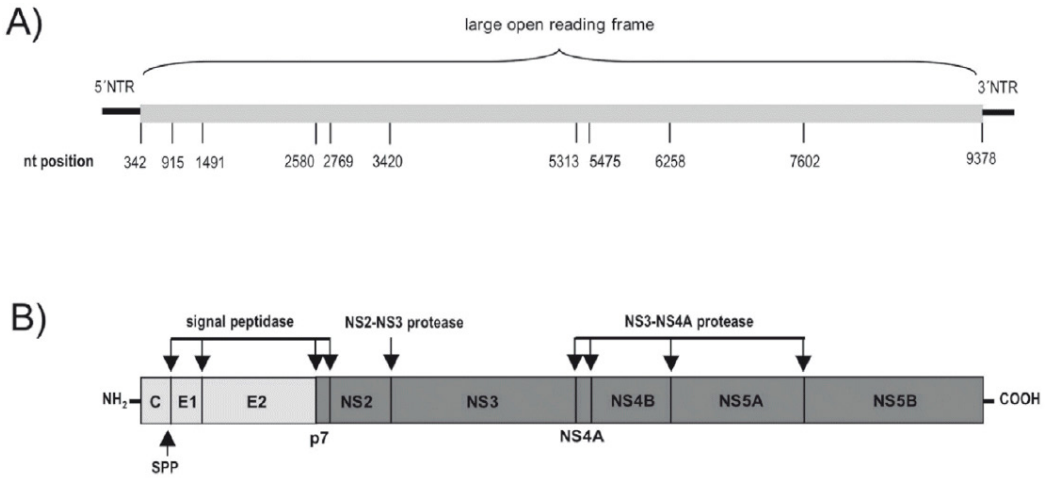


Fig. 2: HCV structure. A) ORF with nucleotide positions in the viral genome. NTR, non-translated region. B) HCV precursor polyprotein with cleavage sites for the signal peptide peptidase (SPP) and the viral peptidases (NS2/NS3 and NS3/NS4A). (Source:[13])

Core

The core protein has a molecular weight of 21 kilodalton (kDa) and consists of homo-multimers located mainly at the endoplasmic reticulum (ER) membrane[43]. It has both structural and regulatory functions; beside formation of the viral capside, it serves for particle assembly, viral RNA binding and regulation of RNA translation[44,45].

E1 and E2

After cleavage from the precursor polyprotein these proteins are N-glycosilated inside the ER lumen[46]. E1 and E2 are type I transmembrane proteins with a large hydrophilic ectodomain and a short transmembrane domain (TMD) of 30 amino acids (aa). The TMD is responsible for anchoring of the envelope proteins in the ER membrane[47–50] and contributes to the formation of E1-E2 heterodimers[51]. The E1-E2 complex is involved in LDL receptor binding. Two hypervariable regions have been identified within the coding region of E2 (HVR1 and HVR2) with a sequence variability of up to 80% in their aa sequences[52,53]. Such high variability of HVRs reduces the activity of neutralizing antibodies targeting them[54].

p7 protein

The p7 protein is a membrane protein localized in the ER where it forms an ion channel[55–57] and plays an essential role in the formation of infectious virions[55,58].

NS2

The non-structural protein 2 interacts with viral factors (E1-E2 glycoprotein and NS3/NS4A complexes) during early stages of virion assembly and morphogenesis[59], and with host factors (liver-specific pro-apoptotic CIDE-B protein) inhibiting apoptosis[60]. Other functions of HCV NS2

include inhibition of cell growth, induction of cell cycle arrest in the S phase[61] and inhibition of cellular IFN-β production involved in the unspecific antiviral cellular response[62].

NS3

The non-structural protein 3 with its ATPase/helicase activity plays an important role during viral replication[63,64]. However, it is the combination with the NS4A cofactor forming the NS3/NS4A protease that is essential for viral infectivity. This protein complex was demonstrated to be a promising target for antiviral therapies after the approval by FDA and EMA (European Medicines Agency) in 2011 of two potent NS3/NS4A inhibitors, boceprevir[65] and telaprevir[66] to be used in combination with IFN-α and ribavirin.

NS4A

The non-structural protein 4A, beside its role as cofactor of the NS3 serine protease, increases the binding stability of NS3 to the ER[67].

NS4B

The non-structural protein 4B is an ER integral protein with a central function in the formation of the HCV RNA replication complex[68] through its nucleotide- and RNA-binding capability[69,70] and the induction of the ER-derived replication platform[71].

NS5A

The non-structural protein 5A is a membrane-associated phosphoprotein with an important role during viral replication[72]. Moreover, it contains the IFN-α sensitivity-determining region (ISDR) that plays a significant role in the response to IFN-α-based therapy[73,74].

NS5B

The non-structural protein 5B is a HCV RNA-dependent RNA polymerase[75] that is essential for viral replication. Using the genomic HCV RNA as a template, the NS5B promotes the synthesis of minus-strand RNA that then serves as a template for the synthesis of genomic positive-strand RNA by the polymerase. It is a non-proofreading error-prone enzyme that incorporates wrong ribonucleotides at a rate of approximately 10^{-3} per nucleotide per generation. Considering the high rate of viral replication, this protein is responsible of the pronounced intra-patient as well as inter-patient HCV differentiation.

F protein (ARFP)

The F (frameshift) or ARF (alternate reading frame) protein is the result of a ribosomal frameshift within the core protein-encoding region. It is localized in the cytoplasm and short-lived[76] but its functions in the viral life-cycle are still unknown. However, detection of anti-F

protein antibodies in the serum of HCV-positive subjects indicates that the protein is expressed during infection *in vivo*[77,78].

VIRAL LIFECYCLE

Adsorption and viral entry

Virus entry into the target cell is a complex mechanism involving several host factors that is not yet fully understood. Viral binding to the cell surface may occur through the binding of HCV E2 envelope glycoprotein and the glycosaminoglycan heparan sulfate[79–81] or via LDL receptor binding, as previously described[33,35,82–84]. After attachment to the cell, the HCV E2 glycoprotein binds to several host factors in a cascade of interactions: SR-BI[85], CD81[86,87], claudin-1[88], claudin-6 and claudin-9[89,90], occludin[91,92], two receptor tyrosine kinase and the Niemann–Pick C1-like 1 cholesterol uptake receptor[93,94]. After binding to different host membrane factors, HCV enters target cells by clathrin-mediated endocytosis[95–98] followed by a fusion step of the viral envelope with the endosome membrane triggered by the acidic endosomal compartment[97,99,100] (Fig. 3). HCV can also spread via cell-to-cell transmission[101] with different mechanisms that however seem to involve CD81 as shown by *in vitro* studies[102].

Translation and post-translational processes

After viral envelope fusion with the endosomic membrane, the genomic HCV RNA is released into the cell cytoplasm. Its 5' NTR allows ribosome-binding and the following cap-independent initiation of translation[103–106]. After assembly of the translational active complex, the HCV polypeptide precursor is produced and post-translationally cleaved to form the 10 functional viral proteins. Cleavage is carried out by two cellular and two viral peptidases. The cellular peptidases are the signal peptidase (SP) that is involved in the formation of the immature protein including the core protein, E1, E2 and p7[107], and the signal peptide peptidase (SPP) that is responsible for the cleavage of the E1 signal sequence leading to the core mature form[42]. The two viral enzymes responsible for cleavage of the non-structural proteins NS2 to NS5B are the NS2/NS3 protease that autocatalytically cleaves the junction between NS2 and NS3[108], and the NS3/NS4A protease that cleaves the remaining functional proteins[109–113]. The E1 and E2 proteins remain within the ER lumen where they are subsequently N-glycosylated[46].

HCV RNA replication

The complex process of HCV RNA replication is poorly understood. As previously mentioned, the principal actors of viral RNA replication are NS5B, an RNA-dependent RNA polymerase (RdRp)

of HCV[75] and NS4B that induces an ER-derived membranous web containing most of the nonstructural HCV proteins[71] necessary for replication. The RdRp uses the previously released genomic positive-strand HCV RNA as a template for the synthesis of an intermediate minus-strand RNA assisted by the NS3 helicase that unwinds putative secondary structures of the template RNA[63,64]. The newly synthesized antisense RNA molecule serves as a template for the synthesis of several plus-strand RNAs that are used as genomic RNA for HCV progeny.

Assembly and release

The precise mechanisms for the formation and release of infectious HCV particles are still unknown, but it is possible that viral assembly takes place within the ER[114] and lipid droplets (LD) are involved in infectious viral particle formation[115–119].

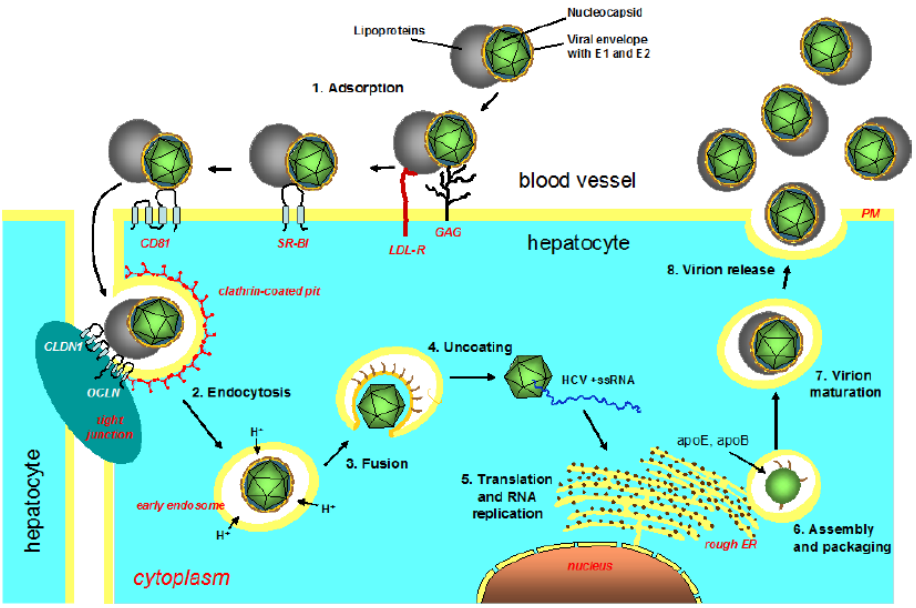


Fig. 3: HCV lifecycle. Cellular components are in red. HCV +ssRNA, single stranded genomic HCV RNA with positive polarity; rough ER, rough endoplasmic reticulum; PM, plasma membrane. (Source:[13])

3.2.5 HCV THERAPY

Before HCV discovery, the first treatment for non-A, non-B hepatitis was based on interferon α (IFN) monotherapy as it had been observed to normalize transaminase levels and improve liver histology in some patients[120]. With the identification of HCV as the infectious agent, clinicians started to measure the success of a therapy in terms of long-lasting disappearance of HCV RNA from serum, the so-called "sustained virological response" (SVR) (Tab. 1). Combination therapy of IFN and ribavirin (RBV), a nucleoside inhibitor, increased SVR rates from 5-20% (with IFN monotherapy) to 40-50%[121,122]. Distinct HCV genotypes (HCV GT) were found to correlate with different SVR rates. Patients with the most frequent HCV genotype, GT1, require longer treatment and show lower SVR rates compared to patients with HCV GT2 and HCV GT3. The development of pegylated interferon α (PEG-IFN) improved IFN pharmacokinetics, allowing more convenient dosing intervals and resulting in higher SVR, especially for HCV GT1. Two PEG-IFN α are available and equally effective in regard to SVR[123]: PEG-IFN α-2b (peg-Intron®, Merck) and PEG-IFN α-2a (PEGASYS®, Roche). In 2011, two inhibitors of the HCV protease (PI), boceprevir (Victrelis®, Merck) and telaprevir (Incivek®, Vertex; Incivo®, Johnson & Johnson) were the first direct antiviral agents (DAA) to be approved by FDA and EMA for the treatment of patients with HCV GT1. It was observed that triple therapy (DAA/PEG-IFN/RBV) improved SVR rates up to 75% in naïve HCV GT1 patients[124,125] and up to 88% in treatment-experienced HCV GT1 patients[126,127]. The combination of these PIs with standard therapy is required since DAA monotherapy results in rapid emergence of drug resistance. Both boceprevir (BOC) and telaprevir (TLV) can be combined with PEG-IFN α-2a or PEG-IFN α-2b[128].

Abbreviation	Term	Description
SVR	Sustained Virological Response	HCV RNA negative 6 months after the end of therapy
RVR	Rapid Virological Response	HCV RNA negative after 4 weeks of therapy
EVR	Early Virological Response	HCV RNA decline ≥2 $\log_{10}$ at week 12
cEVR	Complete Early Virological Response	HCV RNA negative at week 12
NR (BOC)	Non-response (boceprevir)	HCV RNA ≥100 IU/mL at week 12; or HCV RNA positive at week 24, futility rule for BOC
NR (TLV)	Non-response (telaprevir)	HCV RNA ≥1000 IU/mL at week 4 or week 12; or HCV RNA positive at week 12, futility rule for TLV
BT	Breakthrough	HCV RNA was LLD but increased to ≥100 IU/mL or increase of HCV RNA ≥1 $\log_{10}$ during therapy
RL	Relapse	HCV RNA negative at EOT and recurrence of HCV RNA during the follow-up of 6 months
PR	Partial Response	HCV RNA decline ≥2 $\log_{10}$ at week 12 but positive at week 24 during PEG-IFN/RBV
NULR	Null response	HCV RNA decline <2 $\log_{10}$ at week 12 during PEG-IFN/RBV
LI	Lead-In	4 weeks PEG-IFN/RBV before adding a PI

Tab. 1: Relevant definitions for HCV treatment. LLD, lower limit of detection (<10-15 IU/mL; here indicated as negative); EOT, end of treatment; RGT, response-guided therapy. (Modified from [13])

DOUBLE THERAPY (PEG-IFN/RBV)

Until recently, the double therapy with PEG-IFN/RBV was considered the standard of care (SOC) for patients with chronic HCV infection. The innate immune response to viral infections is primarily mediated by IFN through regulation of hundreds of IFN-stimulated genes (ISGs) with anti-viral effect. IFNs induce ISG transcription by activating the JAK-STAT pathway[129]. Type I and Type III IFNs bind to specific cell surface receptors (Type I IFNs to IFNAR, Type III IFNs to the receptor complex formed by IL10R2 and IFNLR1) and activate the receptor-associated tyrosine kinases JAK1 and TYK2. The kinases then phosphorylate and activate STAT1 and STAT2. Activated STATs translocate to the nucleus, where they bind specific DNA elements in promoters of IGSs[130]. The capacity of HCV to interfere with the IFN pathway at many different levels seems to be the mechanism underlying HCV success to establish a chronic infection[131]. Patients not responding to the therapy with PEG-IFN/RBV showed pre-activation of their IFN system in hepatocytes, indicating defects downstream of ISG expression that render these patients refractory to IFN therapy[130]. By contrast, in patients without a pre-activated IFN system, PEG-IFN induces robust up-regulation of many ISGs in the liver within 4 hours[130]. A possible explanation is that injection of exogenous IFN-α during treatment in patients without pre-activated IFN system would very rapidly induce an antiviral state in contrast to a slow buildup of the antiviral state due to pre-activation of the IFN system. A rapid activation of antiviral defenses would not allow the virus to escape and adapt, making it resistant to therapy[130]. Moreover, it was observed that pre-activation of the endogenous IFN system occurred more frequently in patients infected with HCV GT1 and GT4 than in those with HCV GT2 or GT3, thus accounting for the different SVR rates in patients infected by different HCV genotypes (SVR <50% in GT1 and SVR >80% in GT2 and GT3)[130]. It is possible that HCV GTs 2 and 3 are more prone to actively prevent the activation of the innate immunity in the liver resulting in higher susceptibility to IFN-α-based therapies[130].

TELAPREVIR

Telaprevir (TLV) is a selective peptidomimetic NS3/NS4A PI that forms a covalent but reversible enzyme-inhibitor complex. Its efficacy in combination with PEG-IFN/RBV in both treatment-naïve and treatment-experienced patients with HCV GT1 was established by multicenter studies that introduced the new concept of "response-guided therapy" (RGT), in which treatment duration is determined by viral response early in the course of therapy. According to the prescribing information, the three-drug regimen is given for 12 weeks, followed by RGT of either 12 or 36 additional weeks PEG-IFN/RBV depending on both viral response and previous treatment status[132]. Stopping rules for discontinuation of the entire regimen include HCV RNA greater than 1000 IU/mL at weeks 4 or 12 or detectable HCV RNA at week 24[133]. The two major

adverse effects of TLV are rash and anemia (PROVE 1, PROVE 2, PROVE 3 trials). As TLV is a potent CYP3A inhibitor, its use is contraindicated in those patients concurrently taking medications that are also highly dependent on CYP3A clearance or that strongly induce CYP3A[133].

BOCEPREVIR

Like telaprevir, boceprevir (BOC) is a selective peptidomimetic NS3/NS4A PI that forms a covalent but reversible enzyme-inhibitor complex. According to the prescribing information, all patients with HCV GT1 should receive a 4-week lead-in of PEG-IFN/RBV, and BOC three times a day in combination with PEG-IFN/RBV thereafter. Duration of therapy is determined by RGT based on HCV RNA level at treatment weeks 8, 12 and 24[134]. A uniform therapy discontinuation rule is based on HCV RNA of 100 IU/mL or greater at week 12 of treatment. A modest increment in neutropenia of unclear clinical significance and dysgeusia were noted with BOC[133]. As is the case with TLV, BOC is contraindicated for concurrent use of medications that either are highly dependent on CYP3A for clearance or induce CYP3A[133].

TRIPLE THERAPY

The high replication rate and the error-prone nature of viral RNA polymerases generate a viral quasispecies from which variants resistant to viral inhibitors can be selected[135]. It was demonstrated that the use of DAA, especially in monotherapy, constitutes a selective pressure that can favor the emergence of resistance-associated amino acid variants (RAV)[136,137]. It seems that the development of RAVs is associated with the emergence of HCV quasispecies, since they occur more frequently in genotype 1a as opposed to genotype 1b patients, indicating that the latter have a higher genetic barrier to resistance[138,139]. RAVs are usually less fit in terms of replication and/or infectious virus production and, therefore, present in much smaller quantities than wild-type virus[140]. Over weeks to months after cessation of DAA therapy, RAVs tend to regress allowing wild-type virus to re-emerge[141,142]. Concomitant ribavirin administration, fundamental to reduce the risk of virological breakthrough (PROVE 2 study), as well as strict adherence to stopping rules are essential to minimize RAV emergence in patients undergoing treatment[133].

NOVEL THERAPEUTIC AGENTS

Beyond the two approved agents, there are many other antiviral agents targeting both viral and host factors in clinical trials:

- second-generation protease inhibitors binding the active site of the protease (asunaprevir [BMS-650032], danoprevir [ITMN191/R7227], vaniprevir [MK-7009], simeprevir [TMC435]);
- Clemizole hydrochloride that blocks RNA binding to the NS4B in cell culture[70];
- Daclatasvir (BMS-790052), the major inhibitor of NS5A;
- two classes of NS5B inhibitors including nucleoside analogues (Nuc), which are active site inhibitors mimicking the natural polymerase substrate and causing chain termination (mericitabine [R7128]), and non-nucleoside inhibitors (NNI) that bind at four different sites outside of the polymerase active center;
- Alisporivir that inhibits HCV replication by preventing cellular cyclophilin A to induce NS5A isomerization during viral replication[143];
- Miravirsen (SPC3649), a locked nucleic acid-modified oligonucleotide that targets miR-122, making it unavailable for HCV RNA replication[144];
- ITX-5061, the first-in-class HCV entry inhibitor.

The availability of numerous drugs belonging to different classes is currently stimulating multiple DAA and host targeting agent (HTA) combination trials. The goal is to have a novel all-oral HCV therapy with lower chance of resistance development compared with PEG-IFN/RBV[145].

3.2.6 PREDICTORS OF TREATMENT RESPONSE

Predictors of IFN-based HCV treatment outcome have to be sought within viral and host factors. Viral factors include viral genotype, baseline viral load and amino acid substitutions in the core[146] as well as in the NS5A protein (containing the IFN-α sensitivity-determining region)[74] in patients infected with HCV GT1. On the other hand, host factors associated to treatment failure include older age, ethnicity (lower SVR rates in African-American population)[147–149], insulin resistance, advanced fibrosis and hepatic steatosis[123,150].

In 2009, four groups[12,151–153] independently demonstrated a positive association between SVR on treatment with PEG-IFN/RBV in patients infected with HCV GT1 and the genetic variants of two single nucleotide polymorphisms (SNPs) near the *IL28B* locus on chromosome 19. In their genome-wide association studies (GWAS) the rs12979860[151] and the rs8099917[12,152,153] SNPs, respectively located 3 and 8 kb upstream of the *IL28B* gene, were found to be the strongest host-associated genetic predictors of SVR in HCV GT1 patients. These SNPs are in linkage disequilibrium except in patients of African ancestry, where a lower frequency of the favorable *IL28B* genotype was found, which partially explains their lower response rates to the PEG-IFN/RBV therapy compared to other ethnicities[147–149,154,155]. Further studies indicated that the favorable genotype (CC) of *IL28B* rs12979860 in HCV GT1-infected patients correlates with improved early viral suppression (EVR) and SVR in those patients who did not previously experience RVR (rapid virological response)[155]. The same predictive value of these *IL28B* polymorphisms has also been observed in HIV/HCV coinfected patients[156] and in recipients and donors before liver transplant[157,158]. Regarding patients with HCV strains other than genotype 1, studies indicate that the correlation of the *IL28B* SNP genotype with the SVR might be weaker, except for HCV GT4[12,159–164]. However, since patients with genotypes 2 and 3 of HCV normally show much higher SVR rates compared to patients infected by genotypes 1 and 4 (70-85% and 40-50%, respectively[165]) and considering the very low number of studies on HCV genotypes 5 and 6 infections, this lower correlation does not pose a serious problem.

In addition to treatment-induced viral clearance in CHC patients, independent studies showed a strong association between the presence of favorable genotypes of the two *IL28B* polymorphisms (CC and TT) and an elevated immune response during the acute phase of HCV infection resulting in higher frequency of spontaneous clearance[11,166]. Therefore genotyping patients for the *IL28B* polymorphisms also during acute HCV infection is recommended for an early therapeutic intervention[167,168].

The clinical significance of the rs12979860 and rs8099917 genetic variants on the progression of the disease (liver fibrosis, hepatocarcinogenesis) is still controversial and matter of debate[169].

REGIMENS INCLUDING DAA

Several trials for the first-generation protease inhibitors boceprevir and telaprevir showed that the favorable genotype of *IL28B* SNP rs12979860 was associated with higher SVR rates regardless of the patient's therapeutic status (treatment-naïve or treatment-experienced)[170]. In addition to the *IL28B* SNP rs12979860 genotype, liver fibrosis severity is another predictive factor for SVR in TLV/PEG-IFN/RBV therapy[127]. For this reason, the FDA has recently approved genetic testing for *IL28B* rs12979860 genotyping prior to therapy with boceprevir, telaprevir, peginterferon α-2b and simeprevir. This recommendation has been included in the labels of the above mentioned drugs.

Ongoing studies argue against the importance of *IL28B* polymorphisms[169] since combination therapies with next-generation DAAs have shown very high SVR rates (90%) in PEG-IFN/RBV null responders[171]. IFN-free therapy is expected to become the future standard of care, especially in IFN-resistant patients. However, *IL28B* polymorphisms have been reported to affect viral kinetics even in the context of IFN-free regimens[172,173], suggesting that genotypes of *IL28B* polymorphisms will continue to affect treatment efficacy in future regimens[169].

IL28B IN HCV THERAPY

IL28B encodes IFN-λ3, which belongs to the type III IFN-λ family consisting of IFN-λ1, IFN-λ2 and IFN-λ3 (formerly IL29, IL28A and IL28B, respectively). Type-I (IFN-α and IFN-β) and type-III IFNs (IFN-λs) induce antiviral activity and suppress HCV replication *in vitro*[174,175] and *in vivo*[176] by activating the JAK/STAT pathway and upregulating the interferon-stimulated genes (ISGs)[174,177,178]. ISGs together with Toll-like receptor 3 and retinoic acid-inducible gene I signaling pathways of IFN-β induction of the immune system, represent the host antiviral defense against HCV infection[169]. Intrahepatic ISG stimulation by exogenous IFN during therapy is more pronounced in patients reaching SVR[179], whereas high expression before IFN administration correlates with poor response to PEG-IFN/RBV therapy[130,180]. Recent studies have revealed an association between the genotype of the two *IL28B* polymorphisms and expression levels of intrahepatic ISGs[181,182]. It was observed that HCV infection primarily induces type III IFNs (but not type I)[183] and that larger amounts of IFN-λs were produced in the liver of patients with a favorable *IL28B* genotype on treatment with IFN-α[184]. Moreover, IFN-α directly stimulated IFN-λ production by dendritic cells (DC)[185]. Therefore, a model explaining the role of the *IL28B* genotype in predicting the outcome of IFN-α therapy would have to consider the higher IFN-λ production by DC and/or HCV-infected hepatocytes induced by exogenous IFN-α in patients with the favorable *IL28B* genotype during PEG-IFN/RBV therapy[169] (Fig. 4).

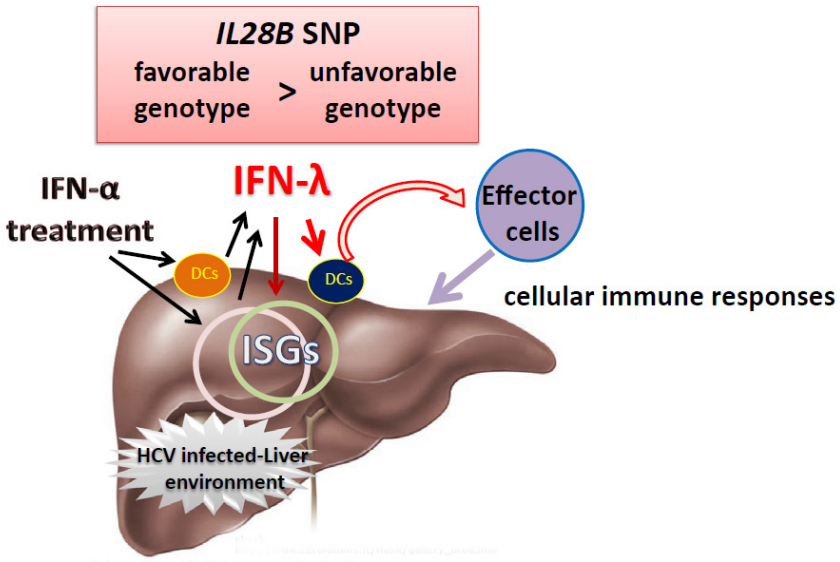


Fig. 4: Potential role of *IL28B* SNP in the response to IFN-α therapy. The favorable *IL28B* genotype induces DCs and infected hepatocytes to produce larger amounts of IFN-λ after IFN-α stimulation during therapy through upregulation of ISGs. (Source: [169])

IL28B, interleukin-28B; IFN, interferon; DCs, dendritic cells; ISGs, IFN-stimulated genes.

Recently, a new polymorphism (ss469415590) located upstream of *IL28B* and in linkage-disequilibrium with rs12979860 was found to be strongly associated with treatment-induced HCV clearance[186] in patients infected with HCV viral genotype 1/4 or 2/3[187]. The ΔG dinucleotide variant of ss469415590 is a frameshift that creates a novel gene, *IFNL4*, encoding the type III IFN-λ4 protein, similar to IFN-λ3 (Fig. 5). Compared to rs12979860, ss469415590 is more strongly associated with spontaneous and treatment-induced HCV clearance in patients of African ancestry[169], probably due to a lower level of linkage-disequilibrium between the two polymorphisms in this ethnicity[186].

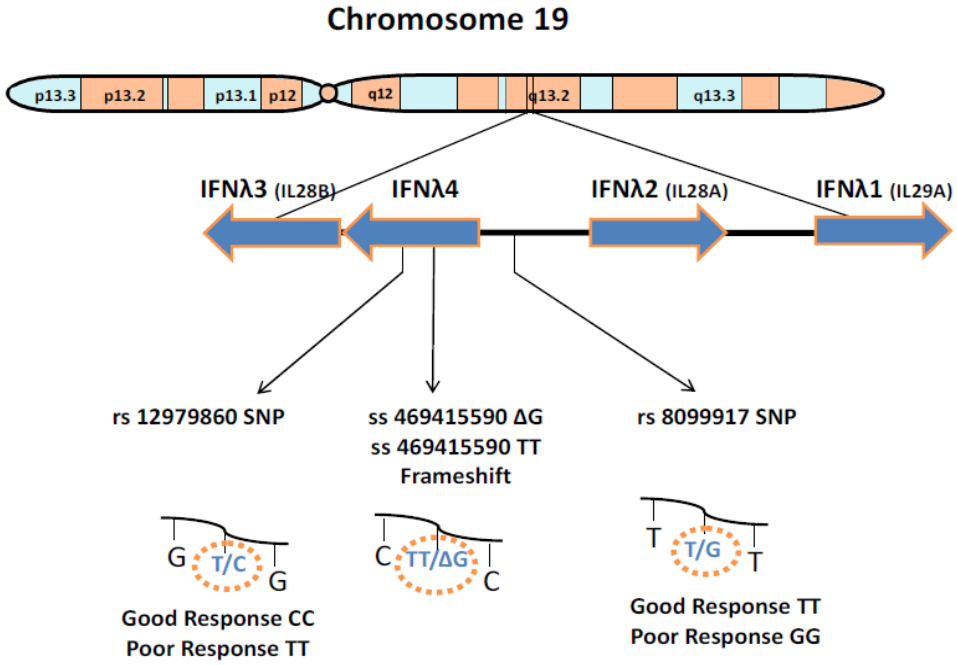


Fig. 5: *IFN-λ3* and *IFN-λ4* polymorphisms on chromosome 19 with predictive value of genotypes regarding treatment outcome.

3.3 MYELOPROLIFERATIVE NEOPLASMS (MPN)

The World Health Organization (WHO) classification system for hematopoietic tumors recognizes five categories of myeloid malignancies including acute myeloid leukemia (AML), myelodysplastic syndromes (MDS), myeloproliferative neoplasms (MPN), MDS/MPN overlap, and PDGFR/fibroblast growth factor receptor 1 (FGFR1)-rearranged myeloid/lymphoid neoplasm with eosinophilia. MPNs are further classified into eight subcategories, including the BCR-ABL1-defined chronic myeloid leukemia (CML) and the so-called "BCR-ABL1-negative MPN": polycythemia vera (PV), essential thrombocythemia (ET) and primary myelofibrosis (PMF). Belonging to the MPN category are also: chronic neutrophilic leukemia (CNL), systemic mastocytosis (SM), chronic eosinophilic leukemia, not otherwise specified (CEL-NOS) and unclassifiable MPN (MPN-U)[188] (Tab. 2).

1. Acute myeloid leukemia (AML) and related precursor neoplasms[a]
2. Myeloproliferative neoplasms (MPN)
2.1. Chronic myelogenous leukemia, *BCR-ABL1* positive (CML)
2.2. Polycythemia vera (PV)
2.3. Primary myelofibrosis (PMF)
2.4. Essential thrombocythemia (ET)
2.5. Chronic neutrophilic leukemia (CNL)
2.6. Chronic eosinophilic leukemia, not otherwise specified (CEL-NOS)
2.7. Mastocytosis
2.8. Myeloproliferative neoplasm, unclassifiable (MPN-U)
3. Myelodysplastic syndromes (MDS)
3.1. Refractory cytopenia[b] with unilineage dysplasia (RCUD)
3.1.1. Refractory anemia (ring sideroblasts <15% of erythroid precursors)
3.1.2. Refractory neutropenia
3.1.3. Refractory thrombocytopenia
3.2. Refractory anemia with ring sideroblasts (RARS; dysplasia limited to erythroid lineage and ring sideroblasts ⩾15% of bone marrow erythroid precursors)
3.3. Refractory cytopenia with multi-lineage dysplasia (RCMD; ring sideroblast count does not matter)
3.4. Refractory anemia with excess blasts (RAEB)
3.4.1. RAEB-1 (2–4% circulating or 5–9% marrow blasts)
3.4.2. RAEB-2 (5–19% circulating or 10–19% marrow blasts or Auer rods present)
3.5. MDS associated with isolated del(5q)
3.6. MDS, unclassifiable (MDS-U)
4. MDS/MPN overlap
4.1. Chronic myelomonocytic leukemia (CMML)
4.2. Atypical chronic myeloid leukemia, *BCR-ABL1* negative (aCML)
4.3. Juvenile myelomonocytic leukemia (JMML)
4.4. MDS/MPN, unclassifiable (MDS/MPN-U)
4.4.1. Provisional entity: Refractory anemia with ring sideroblasts associated with marked thrombocytosis (RARS-T)
5. Myeloid and lymphoid neoplasms with eosinophilia and abnormalities of *PDGFRA*, *PDGFRB* or *FGFR1*[c]
5.1. Myeloid and lymphoid neoplasms with *PDGFRA* rearrangement
5.2. Myeloid neoplasms with *PDGFRB* rearrangement
5.3. Myeloid and lymphoid neoplasms with *FGFR1* abnormalities

Tab. 2: WHO (World Health Organization) classification of myeloid malignancies. [a] AML-related precursor neoplasms include "therapy-related MDS" and "myeloid sarcoma"; [b] Either mono- or bi-cytopenia; [c] Genetic rearrangements involving platelet-derived growth factor receptor α/β (PDGFRA/PDGFRB) or fibroblast growth factor receptor 1 (FGFR1). (Source:[188])

A unifying model of the pathophysiology of BCR-ABL1-negative MPNs entails constitutive activation of the JAK-STAT pathway in megakaryocytes, resulting in initial thrombocytosis and bone marrow fibrosis in the long term[189]. Platelet formation is known to occur in the bone marrow environment following fragmentation of proplatelets, long filaments of megakaryocytes that cross the vascular endothelium and protrude into the sinusoid lumen[190–193]. Under normal conditions megakaryocytes contribute to the bone marrow-matrix environment by expressing fibronectin, type IV collagen and laminin[194]. In BCR-ABL1-negative MPNs driver mutations

appear to alter megakaryocyte differentiation, migratory ability and proplatelet formation, leading to increased platelet production[195]. The natural history of BCR-ABL1-negative myeloproliferative neoplasms is characterized not only by occurrence of thromboembolic complications but also by a tendency toward progression to more aggressive disease, including post–polycythemia vera myelofibrosis or post–essential thrombocythemia myelofibrosis and acute myeloid leukemia or blast-phase disease[196]. Disease progression is also associated with the acquisition of somatic mutations in driver genes responsible for subclonal evolution[197–200].

3.3.1 SOMATIC MUTATIONS IN MPN

In 2005, a unique "gain-of-function mutation" in exon 14 of the Janus kinase 2 gene (*JAK2*) was found in patients with PV, ET and PMF[201,202]. A valine-to-phenylalanine (V617F) alteration within the autoinhibitory pseudokinase domain of *JAK2* was shown to constitutively activate the JAK2 tyrosine kinase, resulting in increased phosphorylation of its substrates and heightened cytokine responsiveness of myeloid cells[201–203] (Fig. 6). The *JAK2* V617F mutation is present in approximately 95% of patients with PV and in 50 to 60% of those with ET or PMF[204]. Somatic mutations in exon 12 of *JAK2* were found in *JAK2* V617F-negative PV patients[205,206]. Soon after, in 2006, activating mutations in exon 10 of the thrombopoietin receptor gene *MPL* were reported in approximately 5% to 10% of ET and PMF patients with non-mutated *JAK2*[207,208]. Overexpression of the mutant *MPL* W515L allele in cell lines leads to cytokine-independent growth, thrombopoietin hypersensitivity, and activated JAK-STAT signaling[209] (Fig. 6). In a murine bone marrow transplant assay, expression of *MPL* W515L resulted in marked thrombocytosis, splenomegaly, and reticulin fibrosis, but not erythrocytosis[207]. *MPL* W515L/K mutations were not found in PV or other myeloid disorders such as MDS, chronic myelomonocytic leukemia, or AML[208].

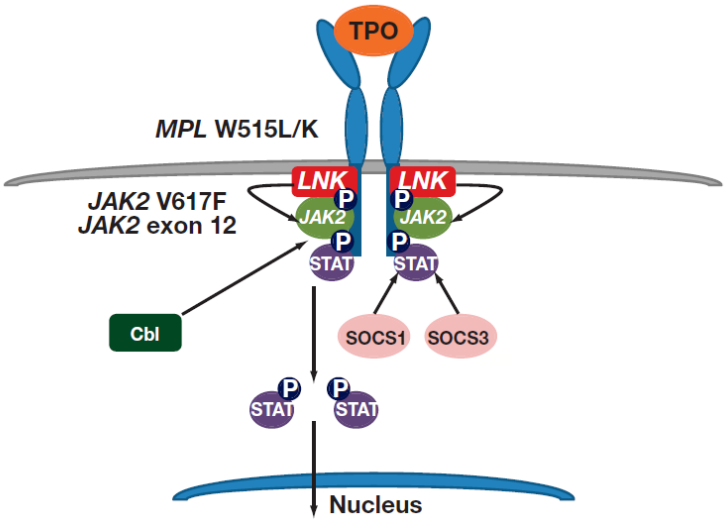


Fig. 6: The JAK-STAT pathway. Normally, cytokines bind to their receptors, which results in JAK2 phosphorylation and recruitment of STAT signaling proteins. JAK2 phosphorylates the STAT proteins and results in activation of downstream signaling pathways. The JAK2 V617F and JAK2 exon 12 mutant kinases bind cytokine receptors which are phosphorylated in the absence of cytokine ligand thus leading to cytokine-independent activation of downstream signaling pathways. The MPL W515L/K mutant thrombopoietin receptor is able to phosphorylate wild-type JAK2 in the absence of thrombopoietin and results in the activation of signaling pathways downstream of JAK2. JAK2 signaling is normally negatively regulated by SOCS proteins, most notably SOCS1 and SOCS3. LNK also serves as a negative regulator by inhibiting JAK-STAT signaling. Cbl proteins function as multifunctional adaptor proteins and ubiquitin ligases are involved in trafficking and degradation of tyrosine kinases. JAK2, Janus kinase 2; LNK, lymphocyte adaptor protein; SOCS, suppressor of cytokine signaling; STAT, signal transducer and activator of transcription; TPO, thrombopoietin. (Source:[209])

In December 2013, two groups independently reported the results of their whole-exome sequencing studies demonstrating the occurrence of somatic mutations in the *CALR* gene encoding for calreticulin in most of *JAK2/MPL*-unmutated PMF or ET[210,211] (Fig. 7 and Tab. 3). CALR is a multi-functional Ca^{2+}-binding protein chaperone that is mostly localized in the ER. Located on chromosome 19p13.2, the *CALR* gene contains nine exons. The protein consists of three domains: the N-terminal lectin-binding domain, the proline-rich P-domain and the unstructured C-terminal acidic domain containing multiple calcium-binding sites. All observed *CALR* mutations were somatic insertions and/or deletions in exon 9 of the gene (Fig. 7 and Tab. 3) that shift the reading frame by one base pair. In contrast to the wild-type protein, the mutant protein mostly contains positively charged amino acids in its C-terminal domain and lacks the KDEL motif. The KDEL amino acid sequence (Lys-Asp-Glu-Leu) is found in some resident ER proteins enabling retrieval of these proteins from the Golgi apparatus to the ER[211]. *CALR* mutations were found in hematopoietic progenitors and did not appear to affect the intracellular localization of the mutant protein[211]. Two variants are present in more than 80% of observed *CALR* mutations: Type 1 (p.L367fs*46), stemming from a 52 base pair (bp) deletion, and Type 2 (p.K385fs*47), which results from a 5-bp TTGTC insertion. Most of these mutations are present in a heterozygous state and only three cases (all of which carrying the Type 2 mutation) of homozygous *CALR* mutations were reported[210]. Both groups found the *CALR, JAK2* and *MPL*

mutations to be mutually exclusive. Since then, several papers reported mutational frequencies for *JAK2*, *CALR* and *MPL* and the percentage of "triple-negative" patients (64%, 15%, 4% and 16%, respectively, with *CALR* mutations accounting for 49% of *JAK2/MPL*-unmutated cases)[212–214]. Compared to patients with *JAK2*-mutations, *CALR* mutations were associated with male gender, younger age, lower hemoglobin level, lower leukocyte count, higher platelet count and higher chance of thrombosis-free survival[212,214].

Somatic mutations in other genes, such as *ASXL1, TET2, CBL, EZH2, IDH1/IDH2, TP53, DNMT3A* and *SRSF2* are present in a number of cases of myeloproliferative neoplasms, but may occur together with *JAK2, MPL* or *CALR* mutations and are not restricted to a certain type of myeloid cancers.

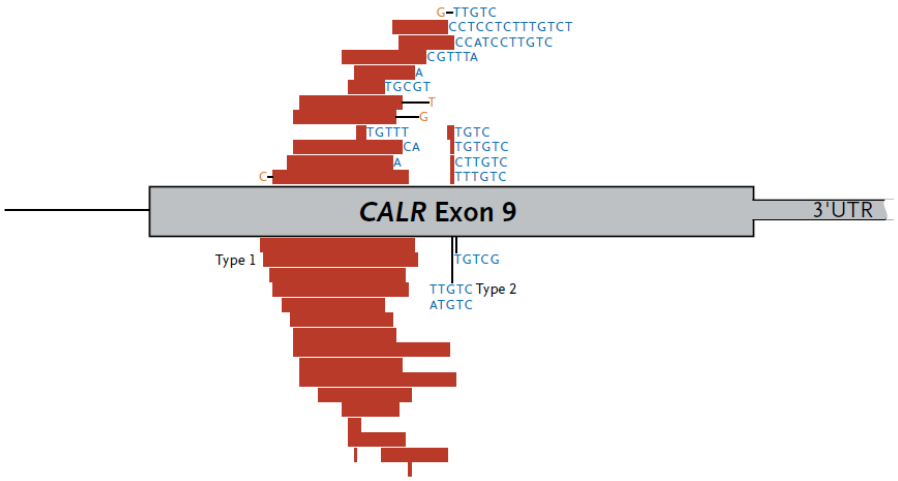


Fig. 7: Genomic positions of the 36 mutation types detected by Klampfl et al., 2013[210]. Red bars indicate deletions, blue letters inserted nucleotide sequences, and orange letters somatically acquired substitutions. Black lines connecting deletions or insertions with substitutions indicate their occurrence on the same allele. The most frequent mutations, Type 1 and Type 2, are marked. UTR, untranslated region. (Source:[210])

Somatic mutation type	Mutant *CALR* genomic annotation
Type 1	chr19:12915565_12915616del
Type 2	chr19:12915627_12915628insTTGTC
Type 3	chr19:12915568_12915613del
Type 4	chr19:12915575_12915608del
Type 5	chr19:12915564_12915615del
Type 6	chr19:12915567_12915612del
Type 7	chr19:12915575_12915626del
Type 8	chr19:12915577_12915610del
Type 9	chr19:12915613del
Type 10	chr19:12915627delinsTGTGTC
Type 11	chr19:12915565G>C; chr19:12915568_12915613del
Type 12	chr19:12915571_12915604del
Type 13	chr19:12915573_12915607delinsA
Type 14	chr19:12915574_12915607del
Type 15	chr19:12915575_12915608del; chr19:12915618C>G
Type 16	chr19:12915575_12915610delinsCA
Type 17	chr19:12915577_12915610del; chr19:12915621A>T
Type 18	chr19:12915577_12915628del
Type 19	chr19:12915583_12915613del
Type 20	chr19:12915591_12915609del
Type 21	chr19:12915591_12915618delinsCGTTTA
Type 22	chr19:12915593_12915596del
Type 23	chr19:12915593_12915604delinsTGCGT
Type 24	chr19:12915593_12915611del
Type 25	chr19:12915595_12915614delinsA
Type 26	chr19:12915595del
Type 27	chr19:12915596_12915598delinsTGTTT
Type 28	chr19:12915604_12915625del
Type 29	chr19:12915608_12915625delinsCCTCCTCTTTGTCT
Type 30	chr19:12915610_12915627delinsCCATCCTTGTC
Type 31	chr19:12915624A>G; chr19:12915627_12915628insTTGTC
Type 32	chr19:12915626_12915627delinsTGTC
Type 33	chr19:12915627_12915628insATGTC
Type 34	chr19:12915627delinsCTTGTC
Type 35	chr19:12915627delinsTTTGTC
Type 36	chr19:12915628_12915629insATGTC

Tab. 3: Genomic annotation of the 36 mutation types detected by Klampfl et al., 2013[210]. (Modified from [210])

3.3.2 DIAGNOSTIC CRITERIA FOR PV, ET AND PMF

After the discovery of *JAK2* and *MPL* mutations, the WHO (2008) officially included *JAK2* mutations as well as other clonal markers into their diagnostic criteria for BCR-ABL1-negative MPN. *JAK2* V617F was classified as major criterion for all three neoplasms. Beyond *JAK2* V617F, the presence of cytogenetic abnormalities or *MPL* mutations was considered for ET and PMF diagnosis[188]. Bone marrow (BM) morphology was included as a minor criterion in PV and a major criterion in both ET and PMF[188] (Tab. 4). Contextually, the WHO committee recognized the fact that neither *JAK2* nor *MPL* mutations were disease-specific and sufficient to distinguish between the three variants of BCR-ABL1-negative MPN.

	Polycythemia vera (PV)[a]	*Essential thrombocythemia (ET)*[a]	*Primary myelofibrosis (PMF)*[a]
Major criteria			
1	Hemoglobin (Hgb) >18.5 g/dl (men) >16.5 g/dl (women) or[b]	Platelet count $\geqslant 450 \times 10^9$/l	Megakaryocyte proliferation and atypia,[c] accompanied by either reticulin and/or collagen fibrosis, or[d]
2	Presence of *JAK2*V617F or *JAK2* exon 12 mutation	Megakaryocyte proliferation with large and mature morphology.	Not meeting WHO criteria for CML, PV, MDS, or other myeloid neoplasm
3		Not meeting WHO criteria for CML, PV, PMF, MDS or other myeloid neoplasm	Demonstration of *JAK2*V617F or other clonal marker or no evidence of reactive BM fibrosis
4		Demonstration of *JAK2*V617F or other clonal marker or no evidence of reactive thrombocytosis	
Minor criteria			
1	BM trilineage myeloproliferation	—	Leukoerythroblastosis
2	Subnormal serum erythropoietin level	—	Increased serum LDH level
3	Endogenous erythroid colony growth	—	Anemia
4		—	Palpable splenomegaly

Tab. 4: 2008 World Health Organization (WHO) diagnostic criteria for myeloproliferative neoplasms. [a] PV diagnosis requires meeting either both major criteria and one minor criterion or the first major criterion and two minor criteria. ET diagnosis requires meeting all four major criteria. PMF diagnosis requires meeting all three major criteria and two minor criteria. [b] Hgb or hematocrit >99th percentile of reference range for age, sex or altitude of residence or red cell mass >25% above mean normal predicted or Hgb >17 g/dL (men) and >15 g/dL (women), if associated with a sustained increase of ≥2 g/dL from baseline that cannot be attributed to correction of iron deficiency. [c] Small to large megakaryocytes with aberrant nuclear/cytoplasmic ratio and hyperchromatic and irregularly folded nuclei and dense clustering. [d] In the absence of reticulin fibrosis, the megakaryocyte changes must be accompanied by increased marrow cellularity, granulocytic proliferation and often decreased erythropoiesis (i.e., prefibrotic PMF). BM, bone marrow; CML, chronic myelogenous leukemia; LDH, lactate dehydrogenase; MDS, myelodysplastic syndrome (Source:[188]).

With the recent discovery that *CALR* mutations are present in about 49% and 74% of ET and PMF patients, respectively[212,213], a new revision of the WHO criteria has been proposed that includes the *CALR* mutation as a major diagnostic criterion for ET and PMF[188] (Tab. 5). However, since *CALR* mutations do not entirely fill the molecular gap in *JAK2/MPL*-unmutated disease and do not allow distinction between ET and PMF, the use of these clonal markers for diagnosis must be accompanied by assessment of BM morphology, which is another major diagnostic criterion for both ET and PMF[188] (Tab. 5). According to the proposed revision, for ET diagnosis is required the accomplishment of the four major diagnostic criteria or, in "triple negative" cases (*JAK2* V617F/*CALR*/*MPL*-negative), of the first three major criteria and the minor criterion. For PMF

diagnosis, it is necessary to meet the three major criteria - 1) typical BM morphology; 2) absence of evidence for another myeloid malignancy as CML, PV and ET; 3) presence of *JAK2*, *CALR* or *MPL* mutations. In triple-negative PMF cases, meeting the first two major criteria together with the three minor criteria is required - 1) exclusion of reactive bone marrow fibrosis; 2) presence of clinical and 3) laboratory features that are typical of PMF[188].

Concerning the diagnosis of PV and its discrimination from "masked PV", which is characterized by *JAK2* mutations as well as hemoglobin levels of 16-18.5 g/dL for men and 15-16.5 g/dL for women, it has been suggested to lower the hemoglobin level (16.5 g/dL in men and 16 g/dL in women) compared to the 2008 criteria (>18.5 g/dL in men and >16.5 g/dL in women)[188]. However, morphologic confirmation through determination of BM morphology is required as a second major criterion for PV diagnosis, except for clinically overt cases (i.e. *JAK2*-mutated patients with hemoglobin levels of >18.5 g/dL in men and >16.5 g/dL in women). In rare *JAK2*-unmutated cases, fulfillment of the minor criterion – subnormal erythropoietin level – is required for diagnosis[188].

	Polycythemia vera (PV)[a]	*Essential thrombocythemia (ET)*[b]	*Primary myelofibrosis (PMF)*[c]
Major criteria			
1	Hemoglobin >16.5 g/dl (men) >16 g/dl (women) or hematocrit >49% (men) >48% (women)	Platelet count $\geqslant 450 \times 10^9$/l	Megakaryocyte proliferation and atypia[d], accompanied by either reticulin and/or collagen fibrosis or[e]
2	BM trilineage myeloproliferation with pleomorphic megakaryocytes	Megakaryocyte proliferation with large and mature morphology	Not meeting WHO criteria for CML, PV, ET, MDS or other myeloid neoplasm
3	Presence of *JAK2* mutation	Not meeting WHO criteria for CML, PV, PMF, MDS or other myeloid neoplasm	Presence of *JAK2*, *CALR* or *MPL* mutation
4		Presence of *JAK2*, *CALR* or *MPL* mutation	
Minor criteria			
1	Subnormal serum erythropoietin level	Presence of a clonal marker (e.g. abnormal karyotype) or absence of evidence for reactive thrombocytosis	Presence of a clonal marker (e.g. abnormal karyotype) or absence of evidence for reactive bone marrow fibrosis
2			Presence of anemia or palpable splenomegaly
3			Presence of leukoerythroblastosis[f], or increased lactate dehydrogenase[f]

Tab. 5: 2014 proposed revision for WHO diagnostic criteria for BCR-ABL1-negative myeloproliferative neoplasms. [a] PV diagnosis requires meeting either all three major criteria or the first two major criteria and one minor criterion. [b] ET diagnosis requires meeting all four major criteria or first three major criteria and one minor criterion. [c] PMF diagnosis requires meeting all three major criteria or the first two major criteria and all three minor criteria. [d] Small-to-large megakaryocytes with aberrant nuclear/cytoplasmic ratio and hyperchromatic and irregularly folded nuclei and dense clustering. [e] In the absence of reticulin fibrosis, the megakaryocyte changes must be accompanied by increased marrow cellularity, granulocytic proliferation and often decreased erythropoiesis (that is, prefibrotic PMF). [f] Degree of abnormality can be borderline or marked and institutional reference range should be used for lactate dehydrogenase level. BM, bone marrow; CML, chronic myelogenous leukemia; MDS, myelodysplastic syndrome (Source:[188]).

In summary, for clinical practice, the *iter* for PV, ET or PMF diagnosis must include the following steps to assess MPN type:

1) Detection of *JAK2* V617F mutation: indicates PV.
2) Detection of *JAK2* exon 12 mutations in the absence of *JAK2* V617F mutation and in the presence of subnormal serum erythropoietin level: indicates PV.
3) Detection of *CALR* mutations in *JAK2*-unmutated cases: indicates ET or PMF.

4) Detection of *MPL* mutations in *JAK2*- and *CALR*-unmutated cases: indicates ET or PMF.

3.3.3 RISK STRATIFICATION AND THERAPY

Stratification of patients according to the prognostic model of the International Working Group for MPN Research and Treatment (IWG-MRT) is necessary for decision-making in therapy. DIPSS-plus is the current prognostic model that follows DIPSS (Dynamic International Prognostic Scoring System) which, in turn, is an evolution of ISSP (International Prognostic Scoring System). ISSP, developed in 2009 by the IWG-MRT[215], was the first prognostic model that included five independent predictors of reduced survival: age >65 years, hemoglobin <10 g/dL, leukocyte count >25 × 10^9/L, circulating blasts ≥1%, and presence of constitutional symptoms[215]. The presence of 0, 1, 2, and ≥3 adverse factors defines low, intermediate-1, intermediate-2, and high-risk disease. DIPSS uses the same prognostic variables but can be applied at any time during the course of the disease[216]. In addition, DIPSS-plus introduced three independent risk factors: platelet count <100 × 10^9/L, need of red cell transfusion and unfavorable karyotype[217]. The four DIPSS-plus risk categories are: low (no risk factors), intermediate-1 (one risk factor), intermediate-2 (two or three risk factors), and high (four or more risk factors)[218]. Recent studies showed that *CALR/ASXL1* mutational status is predictive of leukemic transformation, with *CALR*$^+$*/ASXL1*$^-$ being associated with the longest survival and *CALR*$^-$*/ASXL1*$^+$ with the shortest[219,220].

At present there is no curative therapy for PMF or post-PV/ET MF that can prolong survival. Allogenic stem cell transplant (ASCT) represents the only potential curative treatment but is very dangerous, with transplant-related deaths or severe morbidity occurring in about half of the cases[221]. DIPSS-plus risk categories must be taken into account before administration of therapy. For example, patients with low or intermediate-1 risk disease or low risk molecular profile (*CALR+/ASXL1-*) can be observed without any therapeutic intervention. Specific therapy is considered only in the presence of symptoms (e.g. androgens for anemia, hydroxyurea for symptomatic splenomegaly)[218]. For the management of patients with intermediate-2 or high-risk disease, investigational drug therapy and ASCT are considered. The latter is also indicated in the presence of *CALR*$^-$*/ASXL1*$^+$ mutational status[218]. Investigational drugs currently include pomalidomide, JAK-inhibiting ATP mimetics and mTOR inhibitors, while JAK2 inhibitors have not shown to display disease-modifying activity, including reversal of bone marrow fibrosis or induction of remission[222,223]. Pomalidomide is a second generation immunemodulatory drug that was shown to induce platelet response in a phase-3 study[224].

JAK inhibitor ATP mimetics include:

- ruxolitinib, a JAK1/JAK2 inhibitor, which was seen to reduce spleen size, improve constitutional symptoms through reduction in proinflammatory cytokines and weight

gain, now FDA approved for use in MF[225]. Adverse events include thrombocytopenia, anemia, and a "cytokine rebound reaction" upon drug discontinuation, characterized by acute relapse of symptoms and splenomegaly[226].

- fedratinib, a selective JAK2 inhibitor, which was seen to reduce spleen size, improve constitutional symptoms, normalize platelet and leukocyte count, decrease *JAK2* V617F allele burden[227]. Adverse events are encephalopathy (that caused its withdrawal from further development), anemia, thrombocytopenia and diarrhea[228].
- momelotinib, a JAK1/JAK2 inhibitor, which was seen to induce anemia and spleen response and to improve constitutional symptoms[229]. Adverse events include thrombocytopenia, hyperlipasemia, elevated liver transaminase levels, headache, peripheral neuropathy.
- pacritinib, a JAK2/FLT3 inhibitor, which was seen to induce spleen, symptoms and anemia responses[230]. The most common adverse event is diarrhea.

Since mTOR is part of the JAK-STAT pathway, it is tested as a potential drug target. Up to now, mTOR inhibitors have been demonstrated to reduce splenomegaly, to induce constitutional symptoms response and to lead to pruritus resolution[231].

3.4 ASSAY METHODS

The most widely used molecular techniques for genetic investigations include:

1) Conventional (end-point) PCR
2) Real-time PCR
3) Sanger sequencing
4) Next-generation sequencing

3.4.1 CONVENTIONAL (END-POINT) PCR

The PCR (polymerase chain reaction) method was developed by Kary Mullis in the 1980s and is based on the ability of the DNA polymerase enzyme to synthesize a new strand of DNA that is complementary to a template strand. The DNA polymerase adds nucleotides to the free 3'-OH group of a primer, a short DNA fragment that is complementary to a short stretch of the target region. As the PCR progresses, with each new replication cycle the generated DNA itself is used as a template for replication, resulting in a chain of replications during which the DNA template sequence is exponentially amplified.

Amplification products can be analyzed with different visualization methods:

- **Agarose gel electrophoresis** is a process which enables the sorting of molecules based on size. Using an electric field, DNA molecules move through an agarose gel containing a DNA intercalating agent (e.g. ethidium bromide). After the electrophoresis step, the DNA can be visualized using ultraviolet light that causes the intercalating agent that is bound to the DNA to fluoresce.
- **Reverse allele-specific hybridization** is an oligonucleotide-based DNA array method. Most commonly, a biotinylated DNA amplification product is incubated with specific oligonucleotide probes immobilized on a membrane (often in parallel lines on a membrane strip). DNA fragments that contain sequences complementary to a probe remain bound to the membrane. After this hybridization step, streptavidine-labeled alkaline phosphatase is added and binds to any hybrid of probe and biotinylated DNA. After washing, the membranes are incubated with the chromogenic substrate BCIP/NBT, which is converted into a purple/brown precipitate where alkaline phosphatase is bound.

LIMITATIONS OF CONVENTIONAL PCR

The exponential phase of an end-point PCR is cycle-limited due to several factors like PCR inhibitors in the sample, reagent limitation, pyrophosphate accumulation, and self-annealing of the accumulating PCR product. At this point the polymerase chain reaction ceases to amplify the

target sequence at an exponential rate and enters the "plateau phase", which makes end-point quantification of PCR products unreliable.

3.4.2 REAL-TIME PCR

Real-time PCR allows an accurate quantification of the starting amounts of target molecules. Fluorescence is measured during each cycle, which greatly increases the dynamic range of the reaction, since the amount of fluorescence is proportional to the amount of PCR product. PCR products can be detected using either fluorescent dyes that bind to any double-stranded DNA (dsDNA) or fluorescently labeled sequence-specific probes.

DNA BINDING DYES

DNA binding dyes are fluorescent dyes that bind dsDNA molecules. When stimulated by light, these dyes emit a fluorescent signal of a defined wavelength. Detection takes place during the extension step and the intensity of the fluorescent signal is correlated to the cycle number due to the accumulation of PCR product. The use of dsDNA dyes allows analysis of many different targets without having to synthesize target-specific labeled probes. However, non-specific PCR products and primer–dimers also contribute to the fluorescent signal. Therefore, high PCR specificity is required when using dsDNA dyes.

FLUORESCENTLY LABELED PROBE

Fluorescently labeled probes provide a highly specific method of detection, as only the desired PCR product is detected. However, PCR specificity is also important when using sequence-specific probes. Amplification artifacts such as non-specific PCR products and primer–dimers may result in reduced yields of the desired PCR product. Competition for reaction components between the specific product and PCR artifacts can compromise assay sensitivity and efficiency.
The following probe chemistries are frequently used:

Hydrolysis (TaqMan®) probes

Hydrolysis assays (TaqMan or 5' nuclease assays) include a sequence-specific, fluorescently labeled oligonucleotide probe in addition to a sequence-specific PCR primer. Hydrolysis assays exploit the 5' to 3' exonuclease activity of Taq polymerase. The hydrolysis probe is labeled with a fluorescent reporter at the 5' end and a quencher at the 3' end. When the hydrolysis probe is intact, the reporter fluorescence is quenched due to its proximity to the quencher (Fig. 8). The amplification reaction includes a combined annealing/extension step during which the probe hybridizes to the target, and the dsDNA-specific 5' to 3' exonuclease activity of Taq cleaves off the reporter. As a result, the reporter is separated from the quencher, resulting in a fluorescence signal that is proportional to the amount of amplified product in the sample.

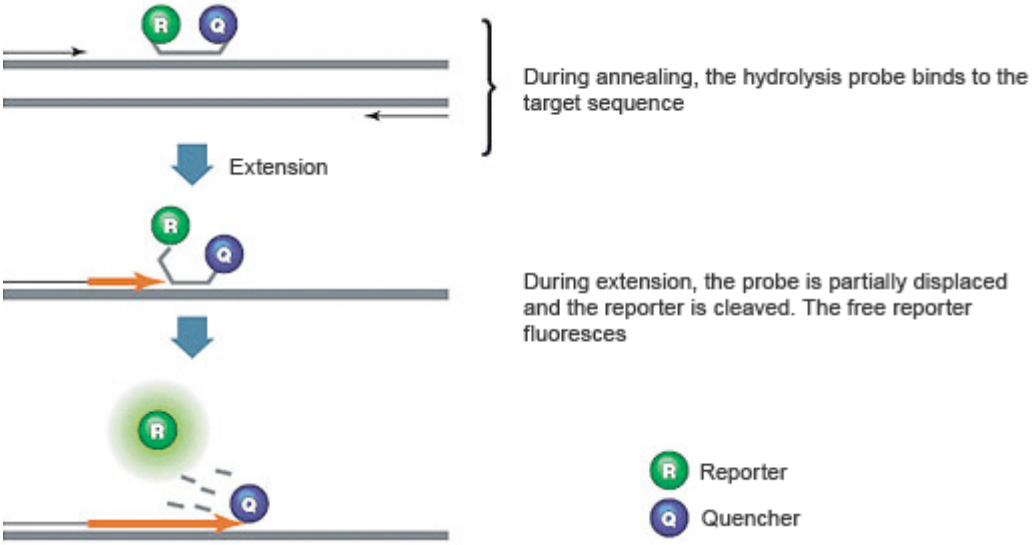


Fig. 8: Mechanism of action of hydrolysis (TaqMan®) probes. (Source:[232])

MGB Eclipse® and TaqMan® MGB® probes

These assays use two DNA primers and a sequence-specific oligonucleotide probe. The Eclipse® probe (Fig. 9) has a fluorescent reporter molecule at the 3' end combined with a quencher and a DNA minor-groove binder (MGB®) at the 5' end. The TaqMan® probe has an opposite orientation to the Eclipse® probe, with the MGB® moiety at the 3’ end and the reporter at the 5' end. With the Eclipse® technology, quenching occurs as the nonhybridized probe adopts a random coil conformation that brings the reporter and quencher together. During annealing of the DNA primers, the probe also hybridizes to the target with the help of the minor-groove binder. The probe thus becomes linearized, separating the reporter and quencher. The resulting fluorescence signal is proportional to the amount of amplified product in the sample. With the TaqMan® technology, fluorescence is emitted after cleavage of the probe by the 5’ exonuclease activity of Taq polymerase during the primer extension step.

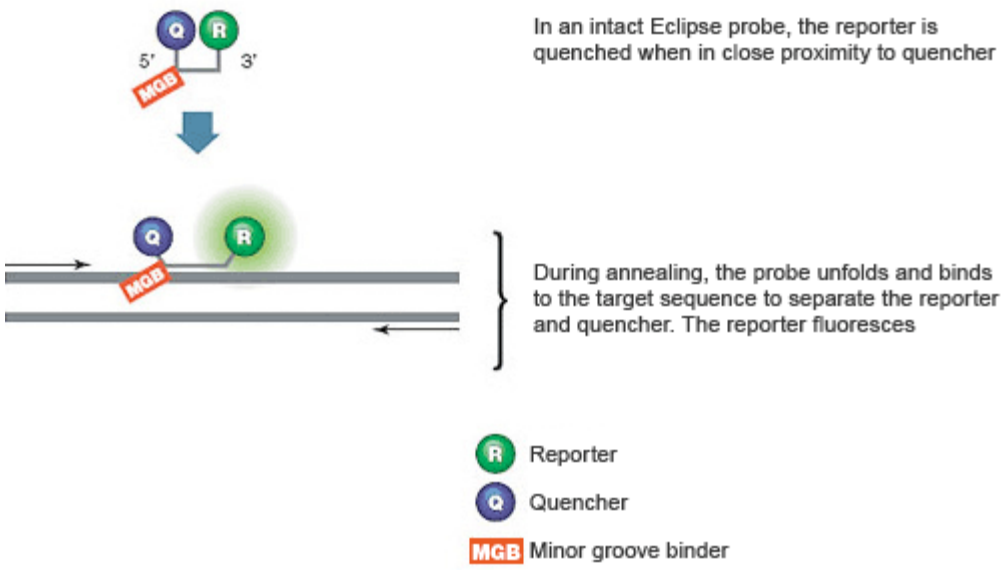


Fig. 9: Mechanism of action of MGB Eclipse® probes. (Source:[232])

Molecular Beacons

Molecular beacon assays include a sequence-specific, fluorescently labeled oligonucleotide probe called a molecular beacon. A molecular beacon is a dye-labeled oligonucleotide (25–40 nucleotides [nt]) that forms a hairpin structure with a stem and a loop (Fig. 10). The 5' and 3' ends of the probe have complementary sequences of 5–6 nt that form the stem structure. The loop portion of the hairpin is designed to specifically hybridize to a 15–30 nt section of the target sequence. A fluorescent reporter molecule is attached to the 5' end of the molecular beacon, and a quencher is attached to the 3' end. Formation of the hairpin therefore brings the reporter and quencher together, hence no fluorescence is emitted. During the annealing step of the amplification reaction, the loop portion of the molecular beacon binds to its target sequence, causing the stem to denature. The reporter and quencher are thus separated, quenching is abolished, and the reporter fluorescence is detectable. Because fluorescence is emitted from the probe only when it is bound to the target, the amount of fluorescence detected is proportional to the amount of target in the reaction.

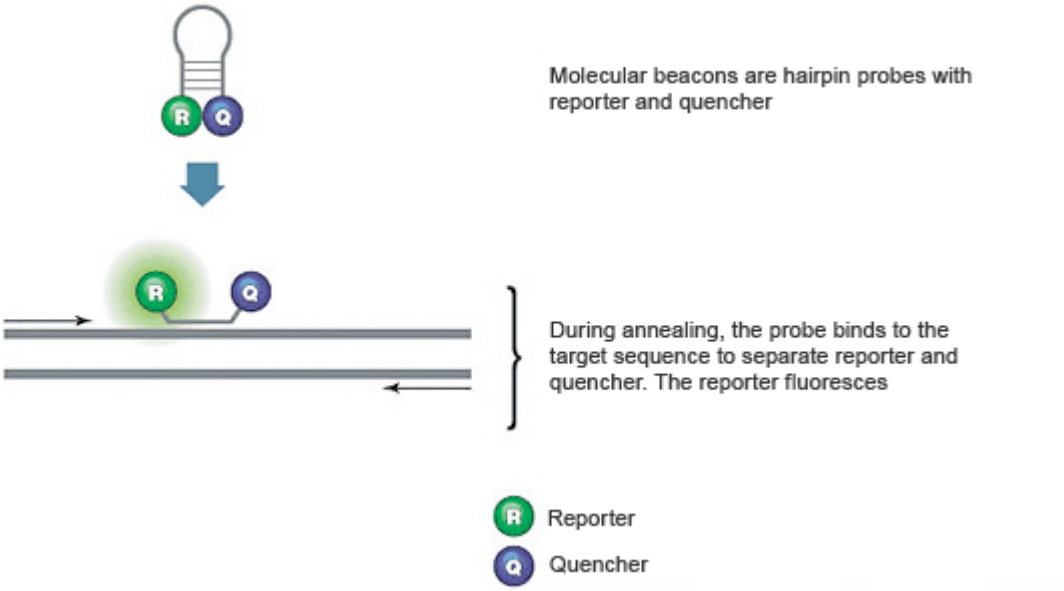


Fig. 10: Mechanism of action of Molecular Beacons. (Source:[232])

Scorpions® PCR primers

In these assays one of the PCR primers also serves as a probe containing a stem-loop structure with a 5' fluorescent reporter and 3' quencher. The loop of the Scorpions® probe includes a sequence that is complementary to an internal portion of the target sequence (Fig. 11). During the first amplification cycle, the Scorpions® PCR primer is extended also generating the sequence complementary to the loop sequence (internal target sequence). After subsequent denaturation and annealing, the loop of the Scorpions® probe hybridizes to the internal target sequence, and the reporter is separated from the quencher. The resulting fluorescence signal is proportional to the amount of amplified product in the sample. The Scorpions® probe contains a PCR blocker, just downstream (in the 3' direction) of the quencher, to prevent read-through during the extension of the opposite strand.

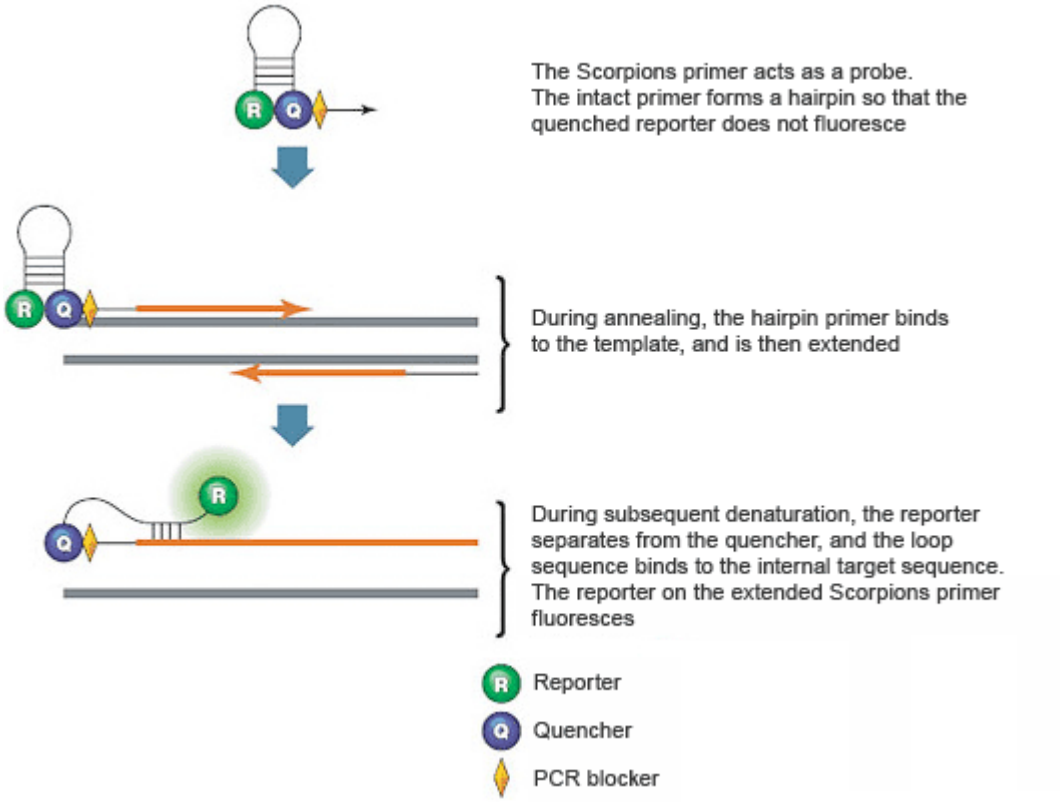


Fig. 11: Mechanism of action of Scorpions® PCR primers. (Source:[232])

Dual hybridization probes

These assays use two sequence-specific oligonucleotide probes in addition to two sequence-specific DNA primers. The two probes are designed to bind to adjacent sequences in the target (Fig. 12). The probes are labeled with a dye pair performing fluorescence resonance energy transfer (FRET). The donor dye is attached to the 3' end of the first probe, while the acceptor dye is attached to the 5' end of the second probe. During real-time PCR, excitation is performed at a wavelength specific to the donor dye, and the reaction is monitored at the emission wavelength of the acceptor dye. In the annealing step, the probes hybridize to their target sequences in a head-to-tail arrangement. This annealing brings the donor and acceptor dyes into proximity, allowing FRET to occur, resulting in fluorescent emission by the acceptor. The increasing amount of acceptor fluorescence is proportional to the amount of PCR product present.

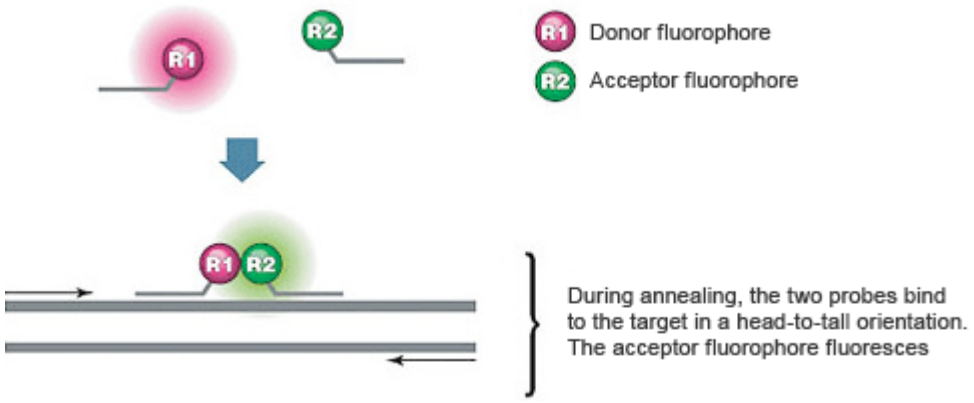


Fig. 12: Mechanism of action of dual hybridization probes. (Source:[232])

Various fluorescent dyes are available for real-time PCR with sequence-specific probes, each with its own excitation and emission maxima. The wide variety of dyes makes multiplex real-time PCR possible (detection of 2 or more targets in the same reaction), provided the dyes are compatible with the excitation and detection capabilities of the real-time cycler, and the emission spectra of the chosen dyes are sufficiently distinct from one another. Therefore, when carrying out multiplex PCR, it is best practice to use dyes with the widest channel separation possible to avoid any signal crosstalk.

PCR QUANTIFICATION

Nucleic acids can be quantified using either absolute quantification or relative quantification. Absolute quantification determines the absolute amount of target (expressed as copy number or concentration), whereas relative quantification determines the ratio between the amounts of target and a control (e.g. an endogenous reference, usually a suitable housekeeping gene). Subsequently, this normalized value can then be used to compare, for example, differential gene expression in different samples.

Absolute quantification

Use of external standards enables the amount or concentration of a target to be given based on the absolute copy number. A standard curve (plot of C_T values or crossing points of different standard dilutions against the log value of the copy numbers contained in the standards) is generated using a dilution series with different known concentrations (standards). For precise quantification, the standards should cover the range of expected values (copy numbers) for test samples. Since the C_T values of standard samples are known, the amount of target in the sample can be determined comparing the C_T value of the sample to the standard curve.

Standards should have the following characteristics:

- primer and probe binding sites identical to the target to be quantified;
- sequence between primer binding sites identical or highly similar to the target sequence;
- sequences upstream and downstream from the amplified sequence identical or similar to the "natural" target;
- equivalent amplification efficiencies of standard and target molecules.

DNA standards for absolute quantification

Plasmid DNA: The most convenient way to create a DNA standard is to clone a PCR product into a standard vector. Advantages of this method are that large amounts of standard can be produced, its identity can be verified by sequencing and DNA can easily be quantified by spectrophotometry. Plasmid standards should be linearized upstream or downstream of the target sequence, rather than using supercoiled plasmid for amplification. This is because the amplification efficiency of a linearized plasmid often differs from that of the supercoiled conformation and more closely mimics the amplification efficiency of genomic DNA or cDNA.
After spectrophotometric determination of the concentration of plasmid DNA, the copy number of DNA molecules in the standard can be calculated using the following formula:
(X g/µL DNA / [plasmid length in base pairs × 660]) × 6.022×10^{23} = Y molecules/µL
PCR fragment: A PCR product containing the target sequence can also be used as a DNA standard. The copy number is calculated using the formula for plasmid DNA (see above), replacing "plasmid length" with the length of the PCR product.

Relative quantification

For relative quantification, the ratio between the amounts of a target and a control gene (e.g., an endogenous reference gene present in all samples) is determined. This normalized value is determined for each sample and can be used, for example, to compare differential expression of a gene in different tissues.

REAL-TIME PCR GLOSSARY

Background: non-specific fluorescence (noise) in the reaction, for example, due to inefficient quenching of the fluorophore or the presence of large amounts of double-stranded DNA template when using DNA binding dyes. The background component of the signal is subtracted by the software algorithm of the real-time cycler (see Baseline below).
Baseline: noise level (background) measured in early cycles, typically between cycles 3 and 15, where there is no detectable increase in fluorescence by amplification products. The average fluorescence of these early cycles is subtracted from the fluorescence value obtained for the amplification products.

Reporter signal: fluorescent signal that is generated during the real-time PCR by either a DNA binding dye or a fluorescently labeled sequence-specific probe.

Passive reference dye: on some real-time cyclers, the fluorescent dye (e.g., ROX) serves as an internal reference for normalization of the fluorescent signal. It allows correction of well-to-well variation due to pipetting inaccuracies, well position, and fluorescence fluctuations.

dUTP/UNG system: a common method to minimize cross-contamination that is based on incorporation of deoxyuridine triphosphate (dUTP) into any PCR product instead of deoxythimidine triphosphate (dTTP). Before the amplification, uracil-containing DNA is selectively degraded by uracil-DNA N-glycosylase (UNG).

Normalized reporter signal (Rn): emission intensity of the reporter dye divided by the emission intensity of the passive reference dye measured in each cycle.

Delta Rn (ΔRn): the normalization of Rn obtained by subtracting the baseline (ΔRn = Rn – baseline).

Threshold: adjusted to a value above the background and significantly below the plateau of an amplification plot. It must be placed within the region of the amplification curve that represents the phase of exponential amplification (linear range, if viewed in logarithmic scale).

Threshold cycle (C_T): the cycle at which the amplification plot crosses the threshold (i.e., there is a significant detectable increase in fluorescence). C_T allows calculation of the starting template amount.

Positive control: control reaction using a known template. A positive control is usually used to check whether the primer set or primer–probe set works and the reaction has been set up correctly.

No template control (NTC): control reaction that contains all essential components of the amplification reaction except the template. This enables detection of contaminations in the PCR reagents or by foreign DNA.

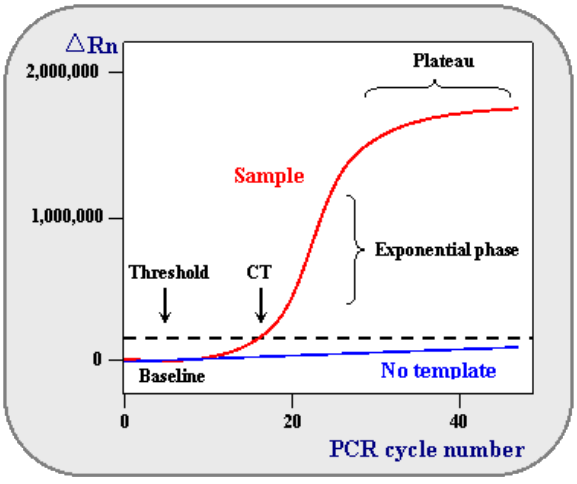


Fig. 13: Real-time PCR amplification plot in linear scale. Some basic concepts of the real-time PCR method are represented.

3.4.3 SANGER SEQUENCING

Sanger sequencing is a method of DNA sequencing developed by Frederick Sanger and colleagues in 1977[233]. The method is based on selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during *in vitro* DNA replication, and has been the most widely used sequencing method for about 25 years. During Sanger sequencing, DNA polymerase copies single-stranded DNA templates by adding nucleotides to a growing chain (extension product) by the formation of a phosphodiester bridge between the 3'-hydroxyl group of a primer and the 5'-phosphate group of the incoming deoxynucleotide in 5' to 3' direction. In addition to normal deoxynucleosidetriphosphates (dNTPs), the reaction also contains modified di-deoxynucleotidetriphosphates (ddNTPs) that lack the 3'-OH group required for the formation of the phosphodiester bond between two nucleotides, causing the DNA polymerase to cease extension of DNA when incorporated. The ddNTPs may be radioactively or fluorescently labeled for detection in automated sequencing machines.

SEQUENCING WITH DYE PRIMERS

When using dye primer-based sequencing chemistry, four separate reactions are performed. Each reaction contains a primer labeled at its 5' end with one of 4 different fluorescently-labeled dyes corresponding to each of the 4 nucleotides–A, C, G or T. ddNTPs are also present in each reaction mix and randomly terminate DNA synthesis, creating DNA fragments of varying lengths. Since a fluorescently-labeled primer is used for extension, all terminated fragments are fluorescently labeled. Following a sufficient number of cycles to allow for optimal generation of extended products, the products of the four reactions are pooled and analyzed on a capillary electrophoresis-based genetic analyzer (Fig. 14).

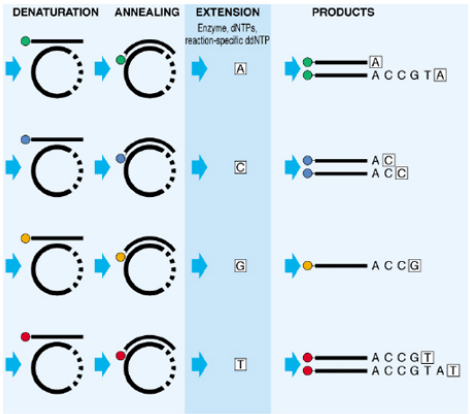


Fig. 14: One cycle of dye primer cycle sequencing. (Source:[234])

SEQUENCING WITH DYE TERMINATORS

Fluorescent DNA sequencing can also be performed by directly attaching a different dye to each of the four ddNTPs, thereby requiring only one reaction tube per sample instead of four. DNA template, unlabeled primer, buffer, the four dNTPs, the four fluorescently-labeled ddNTPs and DNA polymerase are added to the reaction tube. Fluorescently-labeled fragments are generated by incorporation of the dye-labeled ddNTPs that stop DNA strand elongation. Therefore, all terminated fragments contain a dye at their 3' end and are then separated by capillary electrophoresis (Fig. 15).

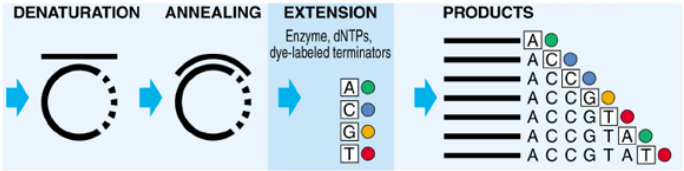


Fig. 15: One cycle of dye terminator cycle sequencing. (Source:[234])

CAPILLARY ELECTROPHORESIS

During capillary electrophoresis, the products of the cycle sequencing reaction are injected electrokinetically into capillaries filled with polymer. High voltage is applied so that the negatively-charged DNA fragments move through the polymer in the capillaries toward the positive electrode. Capillary electrophoresis can resolve DNA molecules that differ in molecular weight by only one nucleotide. Shortly before reaching the positive electrode, the fluorescently-labeled DNA fragments, separated by size, move through the path of a laser beam that causes the dyes on the fragments to fluoresce. An optical detection device on a genetic analyzer detects the fluorescence signal (Fig. 16). The data collection software converts the fluorescence signal to digital data that are recorded in a file. Because each dye emits light at a different wavelength when excited by the laser, all four colors, and therefore, all four bases, can be detected and distinguished in one capillary injection (Fig. 16).

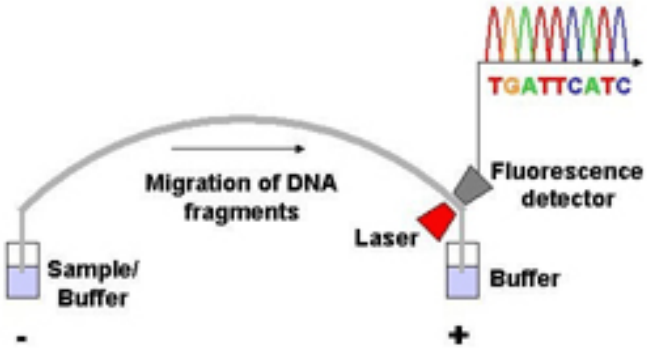


Fig. 16: DNA fragments pass through a laser beam; emitted fluorescence signals are detected by an optical detector and converted to digital data by software. (Source:[234])

3.4.4 NEXT-GENERATION SEQUENCING

Since completion of the first human genome sequence in 2003, demand for cheaper and faster sequencing methods has greatly increased. This demand has driven the development of second-generation or next-generation sequencing (NGS) methods. NGS platforms perform massively parallel sequencing, during which millions of DNA fragments from a single sample are sequenced. Massively parallel sequencing technology facilitates high-throughput sequencing, which allows an entire genome to be sequenced in less than one day. In the past decade, several NGS platforms were developed and provided low-cost, high-throughput sequencing. The creation of such platforms has made sequencing accessible to more labs, thus rapidly increasing the amount of research and clinical diagnostics being performed with nucleic acid sequencing[235]. Three major NGS systems are routinely used in many laboratories today. In 2005, the first system to become commercially available was the Genome Sequencer 454 from Life Sciences (later acquired by Roche), in 2006 Illumina launched the Genome Analyzer and in 2007, the SOLiD system was introduced by Applied Biosystems (now part of Life Technologies). The key steps of a sequencing workflow are the same for all of these technologies, i.e. template preparation (DNA library preparation, amplification of DNA fragments, distribution of templates on a solid support), sequencing and imaging, and data analysis (quality control, base calling, alignment with a reference sequence for resequencing applications) (Fig. 17).

Template preparation

Template preparation consists of building a nucleic acid (DNA or complementary DNA [cDNA]) library and amplifying that library on a solid support. Sequencing libraries are constructed by fragmenting the DNA (or cDNA) sample and ligating adapter sequences (synthetic oligonucleotides of a known sequence) onto DNA fragments ends. Once constructed, libraries are clonally amplified using two different methods: a water-in-oil emulsion PCR after hybridization of library fragments onto microbeads (454 and SOLiD) or a bridge amplification that forms template clusters on a flow cell (Illumina)[236,237].

Sequencing and imaging

The Illumina and 454 technology is based on a sequencing-by-synthesis principle where a DNA polymerase is used to extend a sequencing primer by incorporating nucleotides that form a growing sequence complementary to the template DNA. As nucleotides incorporate into the growing DNA strand, they are digitally recorded as sequences. Detection of nucleotide sequence information may be performed through detection of pH changes induced by the release of a hydrogen ion upon the incorporation of a nucleotide into a growing strand of DNA (454) or detecting fluorescence that is generated by the incorporation of fluorescently labeled nucleotides (Illumina)[236,237]. The SOLiD technology is based on sequencing-by-ligation, where a

DNA ligase is used to add fluorescently labeled probes to a growing oligonucleotide chain. Probes are labeled according to the first two bases using a scheme for two-base encoding with four fluorophores. After imaging, the fluorescent label and the three universal bases of the probe are cleaved off, and a new set of probes is added[236].

Data analysis

Once sequencing is complete, raw sequence data must undergo several analysis steps. A generalized data analysis pipeline for NGS data includes preprocessing the data to remove adapter sequences and low-quality reads, mapping of the data to a reference genome or *de novo* alignment of the sequence reads, and analysis of the compiled sequence. Sequence analysis includes a wide variety of bioinformatics assessments, including genetic variant calling for detection of SNPs or indels (i.e., the insertion or deletion of bases), detection of novel genes or regulatory elements, and assessment of transcript expression levels[238].

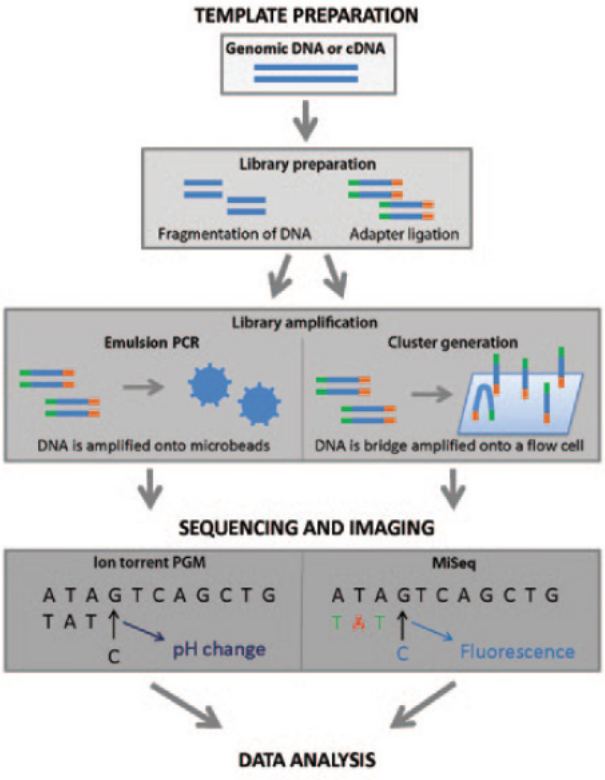


Fig. 17: Workflow of next-generation sequencing with Ion torrent (Life Technologies), which is similar to the 454, and MiSeq (Illumina) Genome Analyzers. (Source: [235]).

3.5 GUIDELINES FOR GENETIC TESTS (CE-IVD)

> *"'in vitro diagnostic medical device' means any medical device which is a reagent, reagent product, calibrator, control material, kit, instrument, apparatus, equipment, or system, whether used alone or in combination, intended by the manufacturer to be used in vitro for the examination of specimens, including blood and tissue donations, derived from the human body, solely or principally for the purpose of providing information:*
> *— concerning a physiological or pathological state, or*
> *— concerning a congenital abnormality, or to determine the safety and compatibility with potential recipients, or*
> *— to monitor therapeutic measures."*

Cited from Article 1 of Directive 98/79/EC[239]

Development and commercialization of a genetic test as IVD requires the following steps:

1. Determination of clinical utility
2. Development using harmonized standards
3. Evaluation of conformity to essential requirements of Directive 98/79/EC
4. CE marking and placing on the market
5. Registration and notification with the Competent Authority

3.5.1 DETERMINATION OF CLINICAL UTILITY

Before a genetic test can be generally accepted in clinical practice, data must be collected to demonstrate the benefits and risks that accrue from both positive and negative results. The balance between benefits and risks of a genetic test determine its clinical utility. In its narrower sense, clinical utility is the ability of a screening or diagnostic test to prevent or ameliorate adverse health outcomes such as mortality, morbidity, or disability through the adoption of efficacious treatments on the basis of test results. In its broader sense, it refers to any use of test results to influence clinical decision-making[240].

In oncogenomics, a genetic test contributes to the diagnosis and the prognosis of a disease, often indicating the best-therapeutic strategy to adopt, and helps monitoring the progression of the disease during/after treatment. In pharmacogenomics, a genetic test provides pre-therapeutic guidance for drug and dose selection and in post-therapeutic monitoring allows dose optimization and evaluation of possible adverse drug reactions or therapeutic failure. Pharmacogenetic tests are usually indicated in the labels of FDA-approved drugs.

3.5.2 DEVELOPMENT USING HARMONIZED STANDARD

A harmonized standard is a European standard that provides guidelines to meet the essential requirements of the European Union legislation (Directive 98/79/EC[239], transposed by the GD 332/2000 in Italy). Manufacturers and conformity assessment bodies can use harmonized standards to demonstrate that products comply with relevant EU legislation. The use of these standards remains voluntary. Manufacturers are free to choose any other technical solution that provides compliance with the mandatory legal requirements.

ISO 13485:2012 is the harmonized standard that applies to the "*Quality management systems for medical devices — Requirements for regulatory purposes*". Among others, it defines the quality management system, management responsibilities and all parts of product realization.

In Directive 98/79/EC[239] (Annex I "Essential requirements") is stated that:

> "*The devices... must achieve the performances, in particular, where appropriate, in terms of analytical sensitivity, diagnostic sensitivity, analytical specificity, diagnostic specificity, accuracy, repeatability, reproducibility, including control of known relevant interference, and limits of detection, stated by the manufacturer*".

Therefore, it is of relevant importance to establish the performances of the genetic test in terms of analytical and diagnostic specificity and sensitivity. Analytical specificity is the capacity of an assay to exclusively detect the marker. Analytical sensitivity corresponds to the limit of detection of the assay, which is the minimum quantity of the marker the assay is able to precisely detect. Diagnostic specificity is the number of samples that resulted "negative" with the assay (no marker detected) on the total number of samples not presenting the marker; high diagnostic specificity of a test corresponds to low risk of "false positive" results. Diagnostic sensitivity is the number of samples that resulted "positive" with the assay (marker detected) on the total number of samples presenting the marker; high diagnostic sensitivity of a test corresponds to low risk of "false negative" results. Analytical validation of a genetic test always requires testing the assay performances with a positive control (synthetic construct or sample containing the marker), a negative control (synthetic construct or sample not containing the marker) and a reaction control (no sample). Diagnostic performances must be established with clinical samples that have already been analyzed with a reference method for the test marker. The reference method is the gold standard method currently used for the investigational purpose of the assay and should be referred to in any phase of the analytical validation. Blind repeat testing and comparison of results with other laboratories is required to evaluate repeatability and robustness of the assay.

3.5.3 EVALUATION OF CONFORMITY TO ESSENTIAL REQUIREMENTS OF DIRECTIVE 98/79/EC

"Essential requirements" (Annex I of Directive 98/79/EC[239]) establish the general requirements of a device and the specific requirements for design and manufacturing.

General requirements concern:

- safety and health of patients and users;
- risk analysis and solutions to reduce the risk associated with the use of the device;
- performances evaluation in terms of analytical sensitivity, diagnostic sensitivity, analytical specificity, diagnostic specificity, accuracy, repeatability, reproducibility, including control of known relevant interference and detection limits;
- maintenance of the performances during the whole device lifetime as indicated by the manufacturer;
- maintenance of the performances and characteristics during transport and under the recommended storage conditions.

Design and manufacturing requirements include:

- chemical and physical properties (concerning quality of materials used to produce the device);
- infection and microbial contamination (concerning sterile devices and the risk of contamination during selection and handling of raw materials, manufacture, storage and distribution);
- manufacturing and environmental properties (concerning the risk derived from intrinsic or environmental physical features);
- devices which are instruments or apparatus with a measuring function (concerning precision and accuracy of devices having a primary analytical measuring function);
- protection against radiation (concerning devices that emit potentially hazardous, visible and/or invisible radiation);
- requirements for medical devices connected to or equipped with an energy source (concerning devices incorporating electronic systems, including software to reduce electromagnetic, electric, mechanical and thermal risks);
- requirements for devices for self-testing (to ensure that the device is easy to use by lay users and reduce the risk of user error);
- information supplied by the manufacturer (concerning the data on the label and in the instruction for use). Information supplied with the device must include: the name and address of the manufacturer; identification of both the device and the contents of the packaging; batch code or serial number; expiration date; where appropriate, a

statement indicating the *in vitro* use of the device; storage and/or handling conditions; warning and/or precautions to take.

Directive 98/79/EC[239] groups *in vitro* diagnostic medical devices into two main product classes, denominated List A and List B. To List A belong reagents and reagent products, including related calibrators and control materials, for determining the blood groups ABO system, rhesus (C, c, D, E, e) and anti-Kell, and for the detection, confirmation and quantification in human specimens of markers of HIV infection (HIV 1 and 2), HTLV I and II, and hepatitis B, C and D. To List B belong reagents and reagent products, including related calibrators and control materials, for determining: the blood groups anti-Duffy and anti-Kidd, irregular anti-erythrocytic antibodies, the human infections cytomegalovirus and chlamydia, the HLA tissue groups DR, A, B, the tumor marker PSA, the risk of trisomy 21. In this list are also included reagents and reagent products for the detection and quantification in human samples of the congenital infections rubella and toxoplasmosis, for diagnosing the hereditary disease phenylketonuria and the self-diagnosing device for the measurement of blood sugar.
Since the large majority of *in vitro* diagnostic medical devices do not belong to the above mentioned lists, do not constitute a direct risk to patients and provide results that can often be confirmed by similar methods, relative conformity assessment procedures can be carried out under the sole responsibility of the manufacturer. According to Directive 98/34/EC and national regulations, the intervention of notified bodies in this phase is necessary only for the devices, of which the correct performance is essential to medical practice and the failure can cause a serious risk to health (List A and List B).

3.5.4 CE MARKING AND PLACING ON THE MARKET

In order to affix the CE marking and before placing the device on the market, the manufacturer draws up the EC declaration of conformity (Annex III of Directive 98/79/CE[239]). It states that manufacturer must prepare the technical documentation that includes:

- product description (intended use, functional properties, typology, variants planned, eventual combination with other products or optional);
- documentation of the quality system (regulation adopted by the manufacturer to guarantee the quality system in the form of industrial strategies and written procedures like programs, schemes, manuals, reports);
- design information (the characteristics of the basic materials, characteristics and limitations of the performance of the devices, methods of manufacture);
- information about the manufacturing method (schemes and descriptions of used technologies, product components and quality control procedures);

- risk analysis and solution adopted to reduce the risk;
- adequate performance evaluation data (performances claimed by the manufacturer and supported by reference measurement systems, the reference methods, the reference materials, the known reference values, the accuracy and measurement units used; such data should originate from studies in a clinical or other appropriate environment or result from relevant biographical references);
- design calculations and of the inspections carried out;
- test reports;
- labels and instructions for use;
- stability studies.

3.5.5 REGISTRATION AND NOTIFICATION WITH THE COMPETENT AUTHORITY

For placing a medical device on the European Economic Area (EEA) market, Directive 98/79/EC[239] requires manufacturers to provide certain information to the Competent Authorities in the EEA Member State where they have a registered place of business. These requirements have been transposed into national laws of the EEA Member States (GD 332/2000 in Italy). After the receipt of the notification/registration, the Competent Authorities shall process the data and inform the Commission of the European Communities and the other States Party to the Agreement on the EEA, upon request. Data required by law are made available to the Competent Authorities and stored in the European database. The following data shall be submitted:

- on registration of manufacturers
- relating to certificates issued, amended, suspended, withdrawn or refused
- obtained in accordance with the vigilance procedure
- on clinical investigations.

4 AIM OF THE STUDY

Aim of this doctorate project was the development of CE-IVD diagnostic kits for:

1. Detection and genotyping of the *IL28B* polymorphisms rs12979860 and rs8099917
2. Detection and, if possible, semi-quantification of *CALR* mutations and the *MPL* W515L/K mutations.

The field of this project was personalized medicine with real-world applications. The final objective was to promote the use of personalized medicine by the National Healthcare System by reducing the costs of diagnostic tests, which are the main limiting factor to the widespread use of personalized medicine in routine clinical practice.

To optimize the test cost-effectiveness, the following issues should be considered:

- the test should be completely informative for the management of a disease, that explains the increasing diffusion of gene panel testing in the last years;
- the technology used in the test should allow customization of reagents and sample number processed in a run.

On these grounds, we aimed to develop fully informative tests supporting clinical decision-making in chronic hepatitis C infection and BCR-ABL1-negative myeloproliferative neoplasms. For the latter, we did not investigate the most common *JAK2* V617F mutation because tests for the detection, semi-quantification and absolute quantification of this mutation are already available from AB ANALITICA.

The process of assay development applied during this study is summarized in the flowchart below (Fig. 18). Development of an IVD test includes the following steps:

1) Definition of objectives and requirements
2) Selection of assay method
3) Acquisition of clinical samples, preparation of positive controls, design of components
4) Optimization of protocol via testing different reaction conditions and preparation of a prototype assay
5) Validation on different instruments (for instrument type-dependent assays)
6) Determining and validating the performance of the assay
7) Notification with the Ministry of Health and marketing as CE-IVD diagnostic kit.

Each step of this process was subjected to accurate verification in order to assess whether the established requirements were met. External validation of the prototype in a partner laboratory is optional but can be useful to obtain information on the use of the assay in the “real practice” e.g. in terms of comparison with other methods in use and different operators performing the test.

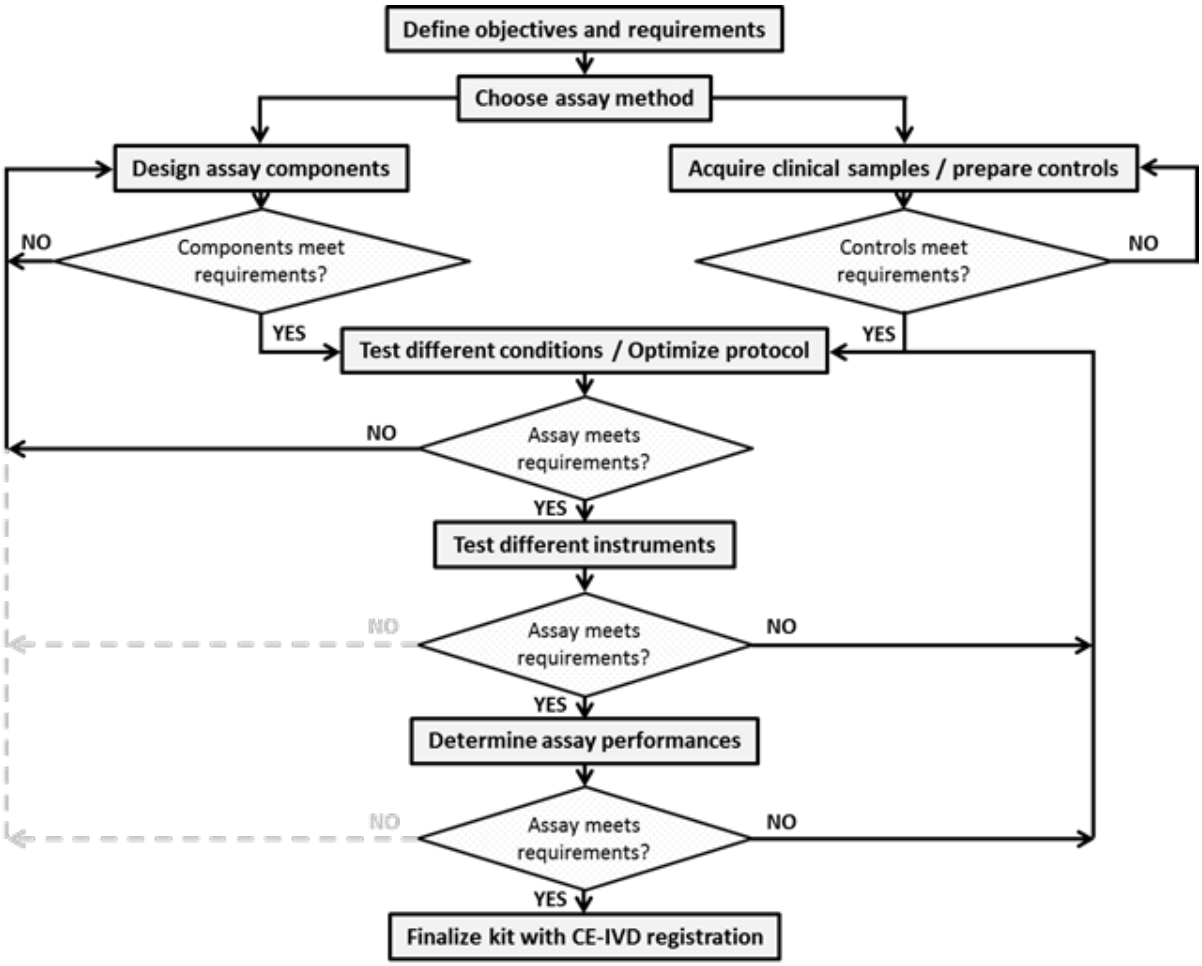


Fig. 18: Flowchart for IVD test development. Each step of the process must be verified and meet the specific requirements in order to pass to the next. Dashed lines indicate an alternative course to follow in case a problem persists even after several rounds of optimization (solid line).

5 MATERIALS AND METHODS

5.1 SAMPLES

5.1.1 IL28B ASSAYS

We used samples of genomic DNA provided by different Italian institutes and analysis laboratories (Tab. 6). Genomic DNA was extracted from peripheral whole blood with the systems routinely used by the providing laboratory. We isolated DNA from blood samples that had been treated with an anticoagulant using the automated extraction system *EZ1 Advanced XL System* (Qiagen) with the *EZ1 Advanced DSP DNA Blood Card* (Qiagen). All samples were genotyped with a reference IVD system (*GENEQUALITY IL28B-ITPA TYPE*, AB ANALITICA). *GENEQUALITY IL28B-ITPA TYPE* is an *in vitro* diagnostic tool based on Reverse Line Blot (RLB) that allows genotyping by allele-specific hybridization of the PCR product to a membrane and colorimetric analysis ("strip"). This system permits simultaneous detection and genotyping of the polymorphisms rs12979860 and rs8099917 in the *IL28B* gene and genotyping of rs7270101 and rs1127354 in the *ITPA* gene (Inosine Triphosphatase). Genotyping data provided from our two main partner laboratories (Department of Medicine of the University of Padua, DIMED and Laboratory of Molecular Hepatology of the Venetian Institute of Molecular Medicine of Padua, VIMM) respectively using *TaqMan Allelic Discrimination Assay* (Life Technologies) and *LightMix Kit IL28B* (Roche), fully confirmed the results obtained with the above mentioned reference IVD system.

Sample number	Extraction systems	Genotyping reference systems
4	*MagNA Pure* (Roche)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA)
3	*QuickGene-Mini80 and QuickGene-810* (Fujifilm)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA)
4	BioRobot EZ1 with *EZ1 DSP Blood* (Qiagen)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA)
5	*Maxwell 16 Blood DNA Purification Kit* (Promega)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA)
4	*NucleoSpin® Blood* (Mackerey-Nagel)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA)
3	*chemagic Prepito®-D* (PerkinElmer)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA)
106	*QIAamp DNA Blood Mini Kit* (Qiagen)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA) and *TaqMan Allelic Discrimination Assay* (Life Technologies)
71	*High Pure PCR Template Preparation Kit* (Roche)	*GENEQUALITY IL28B-ITPA TYPE* (AB ANALITICA) and *LightMix Kit IL28B* (Roche)

Tab. 6: List of samples used for development of IL28B assays.

5.1.2 CALR MUTATION AND MPL W515L/K ASSAYS

All mutated samples for *CALR* and *MPL* genes and part of the wild-type samples used in this study were provided by the partner laboratory of the Department of Experimental and Clinical Medicine, University of Florence (Tab. 7 and Tab. 8). In this laboratory, genomic DNA was extracted from peripheral whole blood or isolated granulocytes using *QIAsymphony DSP DNA Midi Kit* (Qiagen) or *ReliaPrep™ Blood gDNA Miniprep System* (Promega). To increase the panel of tested extraction systems, we also isolated genomic DNA from known wild-type blood samples using different commercial extraction systems and following the manufacturer's instructions. All samples were sequenced with the Sanger method, which is the gold standard for determination of individual mutational profile.

CALR MUTATION

Sample number	Extraction systems	Reference system
1	*QIAamp DNA Blood Mini Kit* (Qiagen)	Bidirectional sequencing
6	BioRobot EZ1 with *EZ1 DSP Blood* (Qiagen)	Bidirectional sequencing
3	*MagNA Pure* (Roche)	Bidirectional sequencing
6	*Mag maxi Kit* (LGC)	Bidirectional sequencing
49	*QIAsymphony DSP DNA Midi Kit* (Qiagen) and *ReliaPrep™ Blood gDNA Miniprep System* (Promega)	Bidirectional sequencing

Tab. 7: List of samples used for CALR MUTATION assay.

MPL W515L/K

Sample number	Extraction systems	Reference system
11	BioRobot EZ1 with *EZ1 DSP Blood* (Qiagen)	Bidirectional sequencing
6	*QIAsymphony DSP DNA Midi Kit* (Qiagen)	Bidirectional sequencing
4	*ReliaPrep™ Blood gDNA Miniprep System* (Promega)	Bidirectional sequencing

Tab. 8: List of samples used for MPL W515L/K assay.

5.2 PREPARATION OF CONTROLS

5.2.1 IL28B rs12979860 ASSAY

As positive controls for the assay for analysis of *IL28B* polymorphism rs12979860 we designed two *MiniGene™ Synthetic Genes* (IDT). A minigene is a compact version of a gene in which regions that do not affect protein function have been removed. IDT minigenes are usually 50 to 500 bp long and are inserted into a *pIDTBlue vector* ampicillin resistant (Fig. 19). This high copy vector is based on the pBluescript II SK (+) phagemid. The plasmid contains the thef1 intergenic region that allows cis-activated single-strand DNA replication of the sense strand and packaging when co-expressed with a helper phage. T3 and T7 promoters that flank the insertion site enable *in vitro* transcription on either strand. The pUC-based, double-stranded origin of replication begins at base 1143.

pIDTBlue vectors containing the two sequences were resuspended (see "resuspension protocol" below) and chemically transformed into *One Shot® TOP10 Chemically Competent E. coli* (Life Technologies; see "chemical transformation protocol" below) . Plasmid DNA was then isolated with the *QIAprep Spin Miniprep Kit* (Qiagen), quantified with the NanoDrop 2000 Spectrophotometer (Thermo Scientific) and diluted in UltraPure™ Salmon Sperm DNA Solution of 20 ng/µL (Life Technologies).

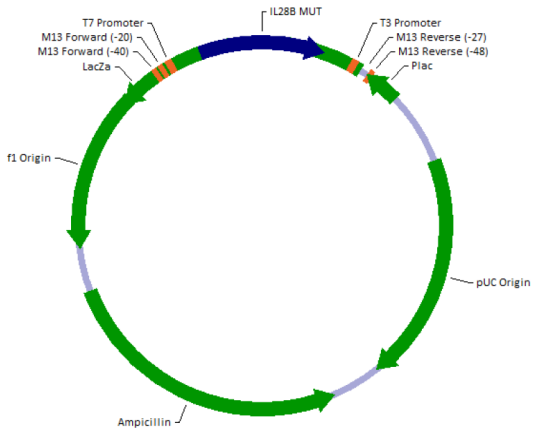


Fig. 19: *pIDTBlue vector*.

RESUSPENSION PROTOCOL

- Centrifuge the tube containing the *pIDTBlue vector* for few seconds to pellet the material
- Add 40 µL of Tris-EDTA (TE)
- Vortex for 20 seconds
- Incubate for 30 minutes at room temperature
- Centrifuge for 1 minute

CHEMICAL TRANSFORMATION PROTOCOL

- Add 2 µL of resuspended *pIDTBlue vector* into a vial of *One Shot® Chemically Competent E. coli* and mix gently
- Incubate on ice for 15 minutes
- Heat-shock the cells for 30 seconds at 42°C without shaking
- Immediately transfer the tubes to ice
- Add 250 µL of room temperature *S.O.C. Medium*
- Shake the tubes horizontally at 37°C for 1 hour
- Spread 25 µL and 50 µL from each transformation on a pre-warmed selective plate with ampicillin and incubate over-night (ON) at 37°C
- Pickup 2-6 white colonies and culture them overnight (ON) in lysogeny broth (LB) medium containing 50 µg/mL ampicillin.

5.2.2 IL28B rs8099917 ASSAY

As positive controls for the assay for analysis of *IL28B* polymorphism rs8099917 we cloned two clinical samples carrying the two polymorphic variants using the *pGC™ Blue Cloning & Amplification Kits* (Lucigen). The *pGC Blue vector* contains a high-copy replication origin and the kanamycin resistance gene. Strong transcription terminators flank the lacZ gene to protect the vector from fortuitous transcription from cloned inserts. The vector is supplied pre-cut with single 3'-C overhangs and dephosphorylated 5' ends. The copy number is similar to that of pUC plasmids (~300 copies/cell). Insert DNA that contains 3'-G tails and 5'-phosphate groups was ligated to the *pGC Blue vector*, transformed into competent cells, and spread on plates containing kanamycin plus XGAL and IPTG[241] (Fig. 20).

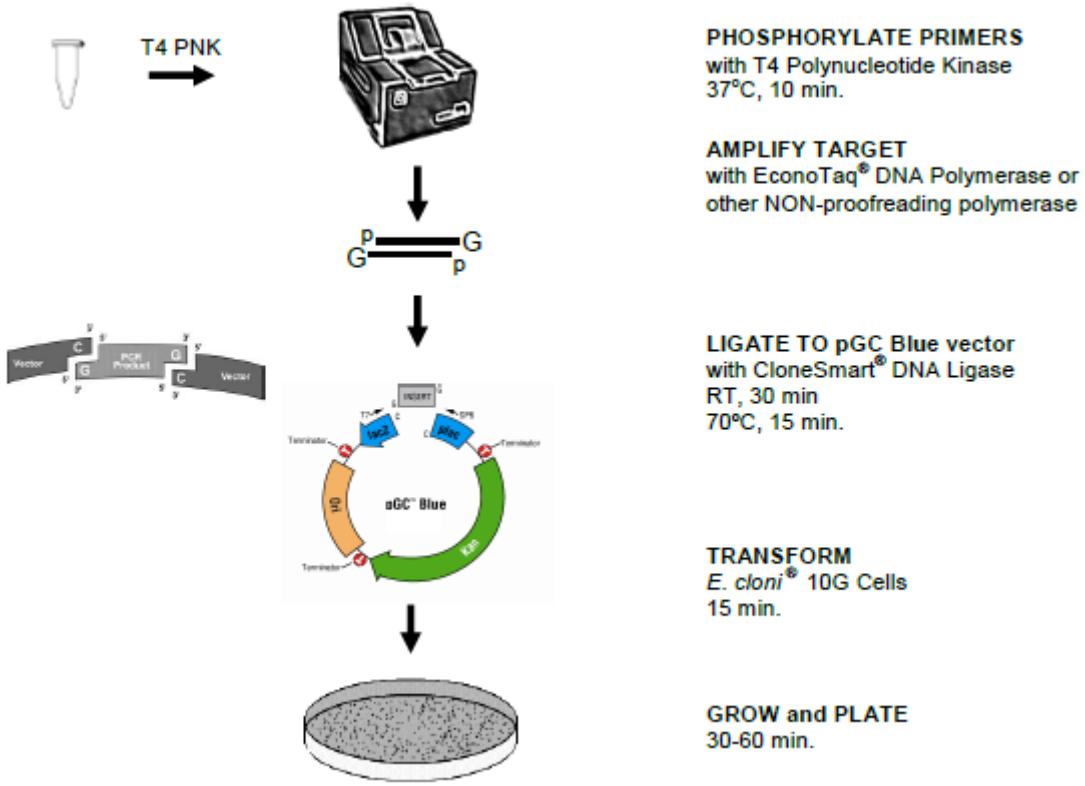


Fig. 20: Protocol of *pGC™ Blue Cloning & Amplification Kits.* (Source:[241])

1. PRIMER PHOSPHORYLATION

   Use *T4 Polynucleotide Kinase* (*PNK*) to add 5´ phosphates to PCR primers before performing the PCR reaction.

   Primer Kinase Reaction:

   2 µL of Forward primer (100µM)

   2 µL of Reverse primer (100µM)

   1 µL of *Primer Kinase Buffer*

   1 µL of *T4 PNK* (10U/µL)

   4 µL of H_2O

   10 µL total

   Incubate at 37°C for 10 minutes.

2. PCR

   A standard PCR using a non-proofreading polymerase was performed. The non-proofreading polymerase adds a single G nucleotide to the 3' end of the PCR product that is necessary for the ligation.

3. QUANTIFICATION OF DNA INSERT

   Approximate quantification of the DNA insert was performed using the *DNA-Marker pUC19/Msp I* (Carl Roth) on an agarose gel.

   Comparing the intensity of the amplicon band with the marker bands, it is possible to approximate the amount of PCR product with the following equation:

   Y : 2686 bp = X : 200 ng

   Y is the size of the marker band closest to the intensity of the amplicon band; X is the quantity (in ng) of the DNA used for electrophoresis; 2686 bp is the total length of the marker; 200 ng is total quantity of marker DNA.

4. LIGATION

   Ligation reaction:

   x µL of Insert DNA (10-400 ng)

   2.5 µL of *4X pGC Blue Vector Premix*

   1 µL of *CloneSmart® DNA Ligase* (2 U/µL)

   y µL of H_2O

   10 µL total

   The reaction mixture was incubated at room temperature for 90 minutes.

The ligation reaction was heat denatured at 70°C for 15 minutes.

5. TRANSFORMATION

Protocol:

- Add 4 µL of denatured product of the ligation reaction to a vial of competent *E. coli* cells on ice and incubate for 30 minutes
- Heat shock cells by placing them in a 42°C water bath for 45 seconds
- Return the cells to ice for 2 minutes
- Add 260 µL of room temperature *Recovery Medium* to the cells in the culture tube
- Place the tubes in a shaking incubator at 37°C for 1 hour
- Plate 50-100 µL of transformed cells on nutrient agar plates containing 30 µg/mL kanamycin plus XGAL and IPTG and incubate the plates ON at 37°C.

6. COLONY SCREENING AND EXPANSION

The *pGC Blue vector* uses the standard blue/white colony screen based on lacZα complementation. Up to 10 white colonies were picked and checked for the presence of the DNA insert using a standard PCR. Transformants were grown ON in LB medium containing 30 µg/mL kanamycin.

7. DNA ISOLATION AND DILUTION OF CONTROLS

Plasmid DNA was isolated using *QIAprep Spin Miniprep Kit* (Qiagen). After quantification of the plasmid solution, the copy number concentration [c/µL] was determined (see "Conversion of molar concentration to copy number concentration" below) and the plasmid diluted in *UltraPure™ Salmon Sperm DNA Solution* of 20 ng/µL (Life Technologies).

CONVERSION OF MOLAR CONCENTRATION TO COPY NUMBER CONCENTRATION

Molar concentration of the isolated DNA plasmid (ng/µL) was converted into copy number concentration [c/µL] using the formulas:

$$C_i = c_i \times N_A \quad (1)$$

$$c_i = n_i/V \quad (2)$$

$$n_i = g/MW \quad (3)$$

n_i is the amount [mol]; V is the volume [L]; C_i is the copy number concentration; c_i is the molar concentration; N_A is the Avogadro constant; MW is the molar weight of double-stranded DNA, which can be approximated by 650 Da × number of base pairs (Da = Dalton).

5.2.3 CALR MUTATION AND MPL W515L/K ASSAYS

As positive controls for the assays for detection of *CALR* and *MPL* mutations we cloned clinical samples with wild-type and mutated sequences using the *TOPO TA Cloning*®[242] (Life Technologies).

TOPO® *Cloning* is a one-step cloning strategy for the direct insertion of Taq polymerase-amplified PCR products into a plasmid vector (Fig. 21).

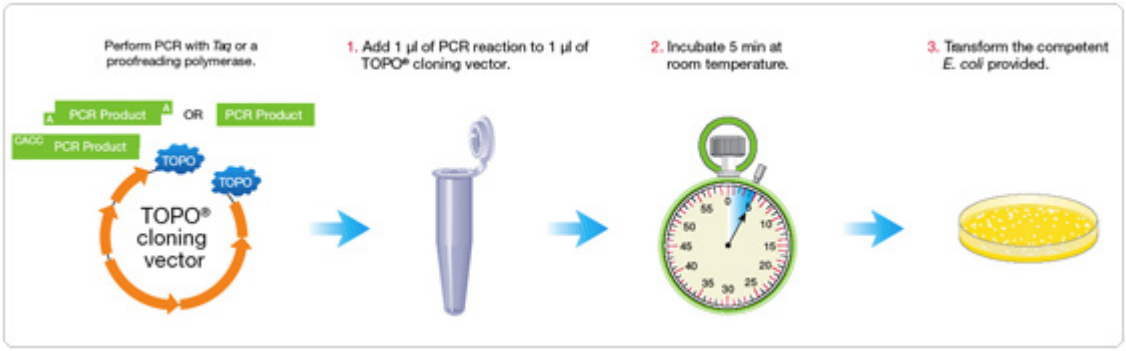


Fig. 21: Protocol of TOPO TA Cloning®. (Source:[243])

The plasmid vector is supplied in linearized form with:

– Single 3´-thymidine (T) overhangs for *TA Cloning*®
– Topoisomerase I covalently bound to the vector ("activated" vector, Fig. 22)

Taq polymerase has a non-template-dependent terminal transferase activity that adds a single deoxyadenosine (A) to the 3´ ends of PCR products. The linearized vector has single, overhanging 3´ deoxythymidine (T) residues. This allows PCR inserts to ligate efficiently with the vector. Topoisomerase I from Vaccinia virus binds to duplex DNA at specific sites and cleaves the phosphodiester backbone after 5'-CCCTT in one strand[244]. The energy from the broken phosphodiester backbone is conserved by formation of a covalent bond between the 3' phosphate of the cleaved strand and a tyrosyl residue (Tyr-274) of topoisomerase I. The phospho-tyrosyl bond between the DNA and enzyme can subsequently be attacked by the 5' hydroxyl of the original cleaved strand, reversing the reaction and releasing topoisomerase[245].

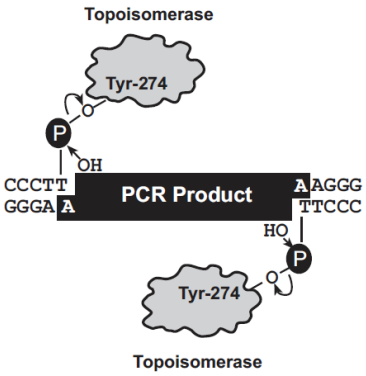


Fig. 22: Topoisomerase I of *TOPO TA Cloning*®. (Source:[242])

1. PCR

   An amplification reaction of the samples containing the target regions was set up with a master mix based on Taq polymerase.

   Amplification reaction:

   x µL of DNA template (10-100 ng)

   12.5 µL of Master Mix

   250 nM primers

   y µL of H_2O

   50 µL total

   Cycling parameters suitable for this reaction mix were used, which included a final extension step of 10 minutes at 72°C to ensure that all PCR products were full length and 3'-adenylated. The PCR product was checked by agarose gel electrophoresis.

2. LIGATION

   TOPO® Cloning Reaction:

   4 µL of fresh PCR product

   1 µL of Salt Solution

   1 µL of *TOPO® vector*

   6 µL total

   The *TOPO® Cloning Reaction* was incubated for 5 minutes at room temperature.

3. TRANSFORMATION

   2 µL of the *TOPO® Cloning Reaction* were transformed into *One Shot® TOP10 Chemically Competent E. coli* (see "Chemical transformation protocol" of *IL28B* rs12979860 assay).

4. COLONY SCREENING AND EXPANSION

   TOPO TA Cloning® uses the standard blue/white colony screen based on lacZα complementation. Up to 10 white colonies were picked and checked for the presence of the DNA insert using a standard PCR. Transformants were grown ON in LB medium containing 50 µg/mL ampicillin.

5. DNA ISOLATION AND DILUTION OF CONTROLS

   Plasmid DNA was isolated using *QIAprep Spin Miniprep Kit* (Qiagen). After quantification of the plasmid solution, the copy number concentration [c/µL] was determined (see

"Conversion of molar concentration to copy number concentration" above) and the plasmid diluted in *Lambda DNA* (Promega) of 10 ng/µL.

SEQUENCING

The positive controls were checked with Sanger sequencing to confirm that the target sequences were cloned in the correct orientation. Sanger sequencing by capillary electrophoresis is the gold-standard technique that is used in a number of experimental workflows for DNA sequence determination. We also used Sanger sequencing as reference method for determination of the performance of the MPL W515L/K and CALR MUTATION assays and confirmation of sample genotypes in case of non-corresponding results (IL28B assays). Samples were prepared for sequencing through amplification of the target region with specific primers. A total of 10 µL of PCR product and the same volume of *DNA-Marker pUC19/Msp I* (Carl-Roth) were loaded into an agarose gel. The fluorescence intensity of the PCR product bands and the DNA-ladder (200 ng/µL) were compared to approximate the amount of DNA in the sample. A quantity of PCR product corresponding to 2 to 4 ng per 100 bp for each sample and 12.8 pmol of primer were dried at 65°C for 40 min. The dried PCR products and primers were delivered to a sequencing service (BMR Genomics, Padua).

5.3 DESIGN AND ANALYSIS OF PRIMERS AND PROBES

Design and analysis of primers and probes can be divided into three phases:

1. Selection of the target region
2. Design of primer pairs and probes
3. Analysis of primers specificity

5.3.1 SELECTION OF THE TARGET REGION

The genomic target sequence was identified using GenBank and dbSNP. These are open access databases of the NCBI (National Center for Biotechnology Information) containing a collection of publicly available nucleic acid and protein sequences and their genetic variations within and across different species.

The following NCBI genomic reference sequences were used: NG_007525.1 for *MPL*, NG_029662.1 for *CALR* and NC_000019.10 for *IL28B*. "NG" refers to an incomplete genomic region while "NC" refers to a complete genomic region, usually a reference assembly.

The UCSC Genome Browser is a web-based graphical viewer that allows locating DNA sequences in the whole genome (Fig. 23). It provides a fast display of any requested portion of the genome at any scale, together with dozens of aligned annotation tracks (known genes, predicted genes, ESTs, mRNAs, CpG islands, assembly gaps and coverage, chromosomal bands, mouse homologies, and more) to recognize other possibly relevant features present in such a region.

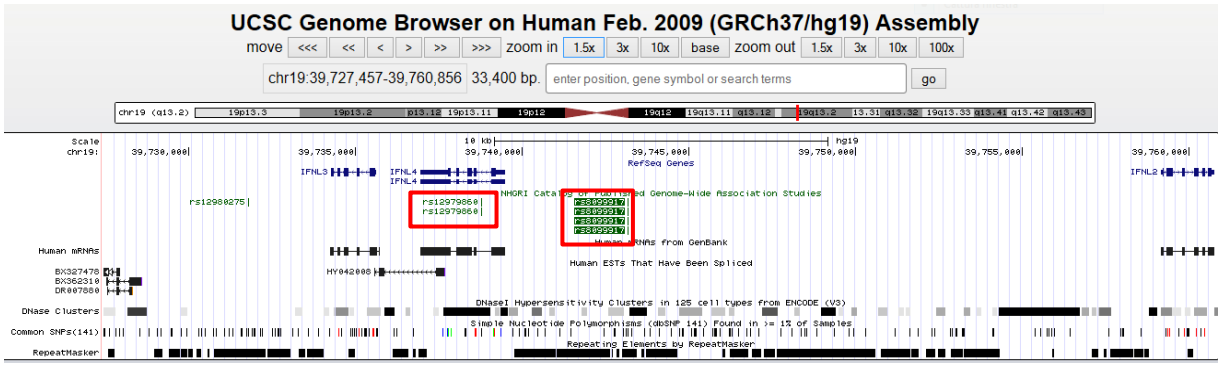


Fig. 23: UCSC screenshot of the chromosome 19 containing rs12979860 and rs8099917 SNPs.

5.3.2 DESIGN OF PRIMER PAIRS AND PROBES

The primers for the qualitative assay were designed using Primer-BLAST, a web software for primer design that combines features of both Primer3 and BLAST.

General guidelines for primer design are:

- primer length: 18-30 bp
- amplicon length: 100-150 bp
- no more than 3 nucleotides, dinucleotide repeats or long stretches of guanosine (G)
- less than 3 bp difference in length between primer pairs
- primer melting temperatures (Tm): 65-75°C
- difference in melting temperature of primers less than 5°C
- GC content between 40 and 60% with the 3' of a primer ending in C or G to promote binding
- balanced distribution of GC-rich and AT-rich domains
- no intra-primer homology (more than 3 bases that complement within the primer) or inter-primer homology (forward and reverse primers having complementary sequences) to avoid self-dimers or primer-dimer formation
- primers designed and ordered in the correct orientation (5' to 3')
- primers for cloning outside of the amplified sequence.

The primers and the probes for the real-time PCR based assays were designed using the *Primer Express® Software v3.0.1* (Life Technologies), an oligo designing tool for gene quantitation and allelic discrimination. This software supports the following platforms for real-time PCR: StepOne™, StepOnePlus™, 7300, 7500, 7500 Fast, 7900HT by Applied Biosystems (Life Technologies) and ViiA™ 7 and QuantStudio™ by Life Technologies.

We designed oligos according to the recommended guidelines listed in Tab. 9. The sequences of the primers and probes designed are a proprietary of AB ANALITICA and therefore cannot be reported here.

	Quantification Assay		Allelic Discrimination Assay	
	Primer	**Probe**	**Primer**	**Probe**
Primer Length (bases)	20	13-30	20	13-30
Amplicon length (bases)	50-150	50-150	50-150	50-150
Tm	58-60°C	68-70°C	58-60°C	65-67°C
GC content	30-80%	30-80%	30-80%	30-80%
Other features:				
No G residue at the 5′ and 3’ ends		X		X
No repeated oligonucleotides and long stretches of G	X	X	X	X
No more than two CC dinucleotides in the middle of the probe		X		X
No G in the second position on the 5′ end of FAM™ dye-labeled probes		X		X
No more than two G + C residues at the 3’ end	X		X	
Tm difference between probes not greater than 1°C				X
SNP site in the middle third of sequence or toward 3′ end but not in the last two bases of 3′ end (Fig. 24)				X
No consecutive A residues		X		X

Tab. 9: Guidelines for primer and probe design in real-time PCR assays.

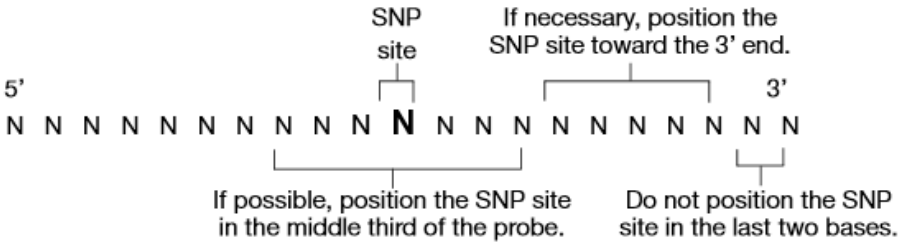


Fig. 24: Guidelines for probe design in assays based on SNPs detection.

5.3.3 ANALYSIS OF PRIMER SPECIFICITY

After designing primers, a BLAST search was performed to determine their specificity. Both the forward and the reverse primer of a primer pair need to match to get a positive result.

We tested for secondary structures using Mfold. This is a versatile folding program for use in analysis of amplification products, linear probes over 30 nucleotides, and Molecular Beacons.

The assessment of secondary structures is important as they could affect the efficiency of the reaction.

For any real-time PCR application it is desirable to obtain a 100% amplification efficiency, which means that each time a cycle is completed, the amount of amplicons has exactly doubled. If a secondary structure is thermodynamically more stable than the primer-target hybrid, formation of the latter will be disfavored. In addition, secondary structures can also prevent a read through of the polymerase.

To determine whether certain structures are favored, the change in Gibbs Free Energy (ΔG) must be calculated. Structures with negative ΔG are favored (ΔG= –10 are strongly favored), whereas a positive ΔG indicates disfavored structures (ΔG=10 strongly disfavored). For each assay the primer pair producing the amplicon with minimal secondary structure was chosen. In cases that secondary structures could not be avoided, the annealing temperature was increased.

In thermodynamics, ΔG is a parameter related to the stability of a system. The transition from one state to another results in a change of the energy of the system. This can be represented by the formula:

[Oligo1] + [Oligo2] ↔ [Duplex]

[Oligo1] and [Oligo2] represent the concentration of the two single stranded oligos in the system and [Duplex] represents the concentration of the hybridization product (duplex) formed by the two oligos.

The change in Gibbs Free Energy (ΔG, [kcal/mol]) is the net exchange of energy between the system and its environment and can be described by the equation:

$$\Delta G = \Delta H - T \times \Delta S$$

ΔH (Enthalpy difference) represents the total energy exchange between the system and its surrounding environment [kcal/mol], ΔS (Entropy difference) represents the energy spent by the system to organize itself [cal/K × mol] and T represents the absolute temperature of the system.

This equation indicates that ΔG is temperature dependent. So, at a given temperature, a positive ΔG value indicates that the system will go in the direction of denaturation, while a negative value indicates that the system will go in the direction of hybridization.

6 RESULTS

6.1 METHOD

6.1.1 IL28B ASSAYS

For the identification of allelic variants of the two *IL28B* polymorphisms rs12979860 and rs8099917 we chose to develop multiplexed (more than one target assayed in the same reaction) end-point (data is collected at the end of a PCR) assays for allelic discrimination (AD).

For each assay we used two primer pairs and specific probes that allow recognition of the two most common variants of each SNP (single-nucleic polymorphism) site. The actual quantity of the target sequence is not determined. Each assay employed a dye-labeled TaqMan® probe that was a perfect match for the major allele of the SNP (allele 1) and one dye-labeled TaqMan® probe that was a perfect match for the minor allele of the SNP (allele 2).

The allelic discrimination assay classifies unknown samples as:

- Homozygous (samples having only allele 1 or only allele 2)
- Heterozygous (samples having both allele 1 and allele 2)

The allelic discrimination assay measures the change in fluorescence of the dye-labeled probes.

In both assays and for all validated PCR instruments we used the FAM™-dye for the major allele and the JOE™-dye for the minor allele.

Fig. 25 illustrates the mechanism of identifying different SNP alleles through specific probes in TaqMan® SNP genotyping assays.

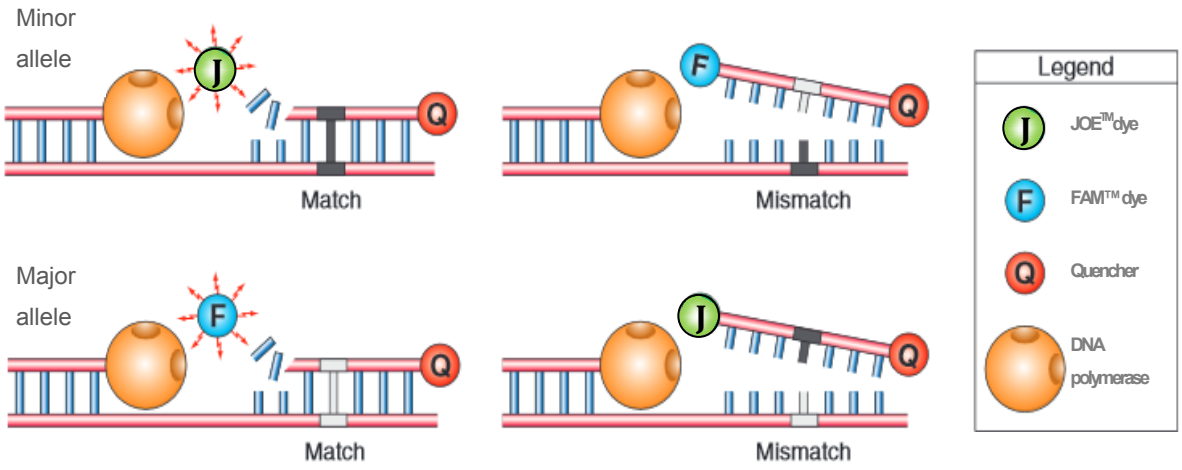


Fig. 25: Mechanism for identification of the allelic variant of the two *IL28B* SNPs through specific TaqMan® probes. (Modified from [246])

Tab. 10 shows the correlation between fluorescence signals and the genotype.

A substantial increase in…	**The sample is…**
JOE™ fluorescence only	Homozygous for the minor allele
FAM™ fluorescence only	Homozygous for the major allele
Both fluorescence signals	Heterozygous

Tab. 10: Correlation between fluorescence signals and the genotype.

6.1.2 CALR MUTATION ASSAY

For the detection of *CALR* mutations we used a conventional multiplex end-point PCR. This method was chosen due to the fact that, up to date, all described *CALR* mutations are insertions, deletions or complex indels found in a relatively short sequence in exon 9 of *CALR*. All known *CALR* mutations generate sequences that differ in length from the wild-type sequence and, therefore, can be detected using conventional PCR and visualization of the amplification products by agarose gel electrophoresis.

6.1.3 MPL W515L/K ASSAY

For assessment of the mutant allele burden concerning the *MPL* gene we chose to develop a semi-quantitative assay based on a multiplexed real-time PCR that specifically detects variants of a single-nucleotide mutation site. In contrast to the allelic discrimination assays utilized for the *IL28B* analysis, the semi-quantitative assay allows to determine the quantity ratio between the mutated and the wild-type allele. However, the absolute quantity of target sequences cannot be obtained. The assay utilized three specific TaqMan® MGB (minor groove binder) probes labeled with fluorescent dyes. The probe specific for the wild-type *MPL* allele was labeled with the VIC®-dye, whereas the probes specific for the mutated alleles (*MPL* W515K and *MPL* W515L) were labeled with the FAM™-dye. Each probe was designed to be a perfect match for its target sequence. The protocol of the assay would utilize two different reaction mixes for analysis of the two mutations (*MPL* W515L and W515K). The mixes would share the PCR primers and the probe that recognizes the wild-type *MPL*, but each mix would only contain one of the mutation-specific probes. Each sample would be tested with both mixes. As controls, we prepared serial dilutions of the mutated allele on a wild-type allele background[247]. Semi-quantitative analysis was performed comparing the amplification curves of the samples with the amplification curves obtained for the positive controls. The ratio between the ΔRn values (ΔRn = fluorescence intensity corrected for the background) of the targets (mutated and wild-type allele) was used to determine the cut-off of the assay (cut-off = threshold value below which samples are considered negative for the analyzed mutation).

6.2 OPTIMIZATION

6.2.1 IL28B ASSAYS

Amplification mixes of IL28B assays were optimized testing different conditions (master mixes and concentrations of primers and probes) with samples of known genotype.

We optimized the amplification reaction step by step, evaluating several parameters of the amplification plot in the linear scale. The requirements for the assay were:

- high separation of the target specific fluorescence signals in homozygous samples (high ΔRn difference) and curves overlap (similar C_T and ΔRn) in heterozygous samples
- balanced fluorescence signals for the two probes in terms of C_T and ΔRn (fluorescence intensity)
- correct genotyping with the automatic call of the instrument

Since the output signals may be substantially different depending on the cycler that is used, optimization of both IL28B assays was performed separately on different Real-Time PCR Systems: Applied Biosystems 7300 (ABI 7300), Applied Biosystems 7500 Fast Dx (ABI 7500 Fast Dx), Applied Biosystems 7500 Fast (ABI 7500 Fast), Applied Biosystems StepOnePlus™ (ABI StepOnePlus™), Applied Biosystems StepOne™ (ABI StepOne™), Dx Real-Time System (Bio-Rad Dx), CFX96™ Real-Time PCR Detection System (Bio-Rad CFX96™).

Fig. 26 and Fig. 28 show two examples of the optimization process for the IL28B rs12979860 assay on ABI 7500 Fast/Fast Dx (Fig. 26) and for the IL28B rs8099917 assay on ABI StepOne™/StepOnePlus™ (Fig. 28).

For optimization of the IL28B rs12979860 assay we first compared two different master mixes with the same concentration of primers and probes. Both mixes gave similar results (ratio of fluorescence intensity) but amplification signals differed in fluorescence intensity by one log unit (Fig. 26 a, b). The mix exhibiting the higher fluorescence intensity was selected for further optimization. In the next step, we aimed to balance the fluorescence intensity of the two probes by increasing the concentration of the probe labeled with JOE™ (Fig. 26 c). Although this led to more similar specific amplification signals of the two probes, it also increased the background of the JOE™ signal. To this point, the amplification curves did not show the ideal S-shape, indicating an inhibition of the PCR. Since too high primer concentrations may inhibit the PCR, we reduced the total primer concentration in the amplification mix. As expected, by this we obtained curves that were closer to the ideal S-shape (Fig. 26 d). The final amplification mix showed balanced fluorescence signals for the two probes in terms of C_T, well separated curves in homozygous samples and similar curves (similar C_T and ΔRn) in heterozygous samples (Fig. 26 e).

Fig. 27 (A and B) shows the results of optimization of IL28B rs12979860 assay on ABI StepOne™/StepOnePlus™, Bio-Rad Dx/CFX96™ and ABI 7300.

For optimization of the IL28B rs8099917 assay we proceeded in a similar manner, except for testing different master mixes due to the similar nature of both IL28B assays. The initial condition we tested showed lower intensity of JOE™ signal compared to the FAM™ signal. This phenomenon was most probably due to a higher base fluorescence of the JOE™ dye (Fig. 28 a). We tried to increase the fluorescence intensity of the JOE™ labeled probe using a higher probe concentration. Since results were not satisfactory (Fig. 28 b), we also tested the mix with a lower concentration of the FAM™ probe. In this case the fluorescent signals of the two probes were more balanced in terms of C_T and ΔRn (Fig. 28 c). Increasing primer concentration led to inhibition of the PCR (Fig. 28 d). For the final mix we selected the conditions that best balanced the signals of the two probes (Fig. 28 e).

Fig. 29 (A and B) shows the results of optimization of IL28B rs8099917 assay on ABI 7500 Fast/Fast Dx, Bio-Rad Dx/CFX96™ and ABI 7300.

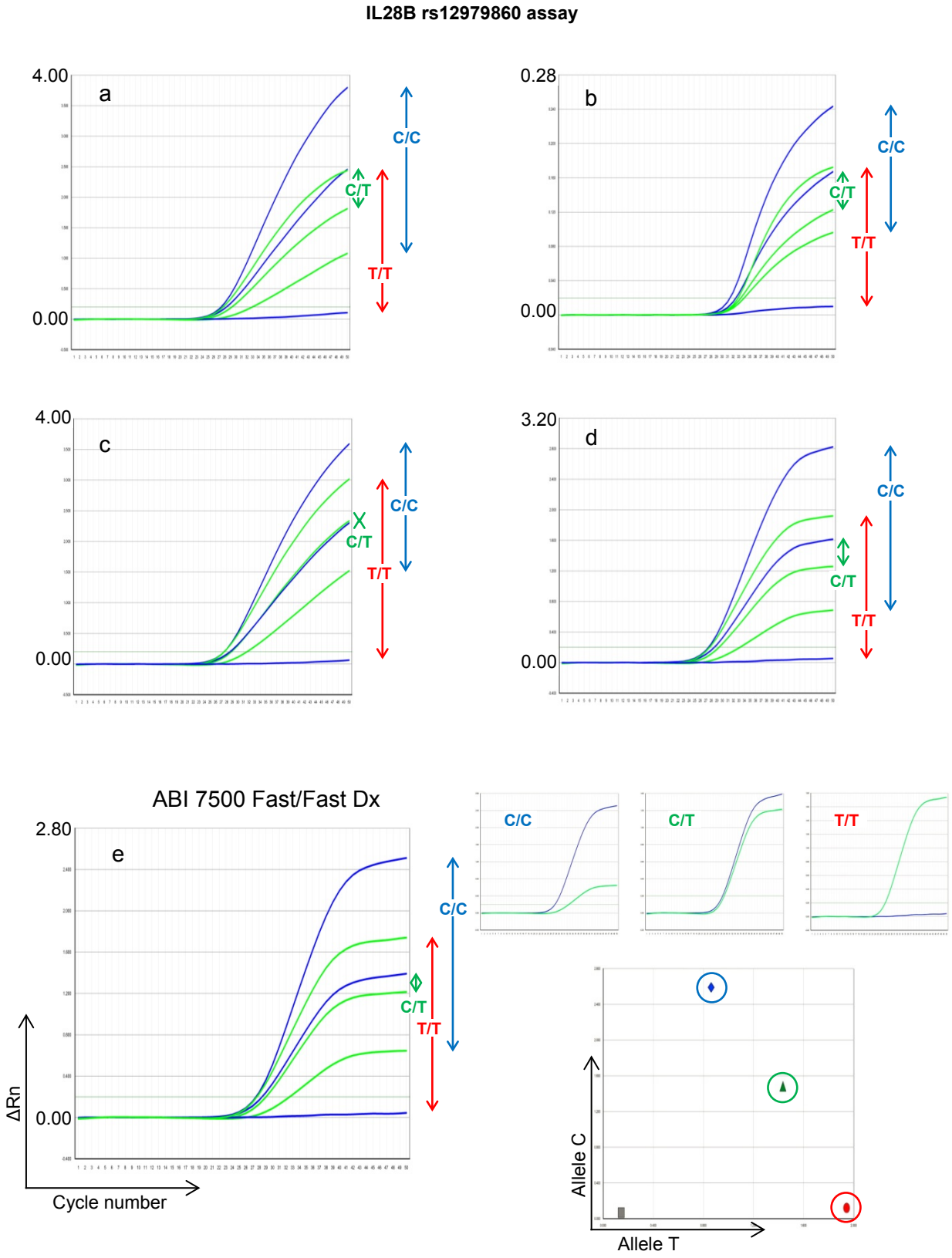


Fig. 26: Steps for development of IL28B rs12979860 assay on ABI 7500 Fast/Fast Dx. Double arrows show ΔRn difference between curves of each sample. a) initial mix, b) different master mix, c) condition "a" with higher concentration of JOE™ probe, d) condition "a" with lower primers concentration, e) IL28B rs12979860 assay on ABI 7500 Fast/Fast Dx. At the bottom right corner are shown the details of amplification plot "e" divided by genotype and the respective allelic discrimination scatter plot.

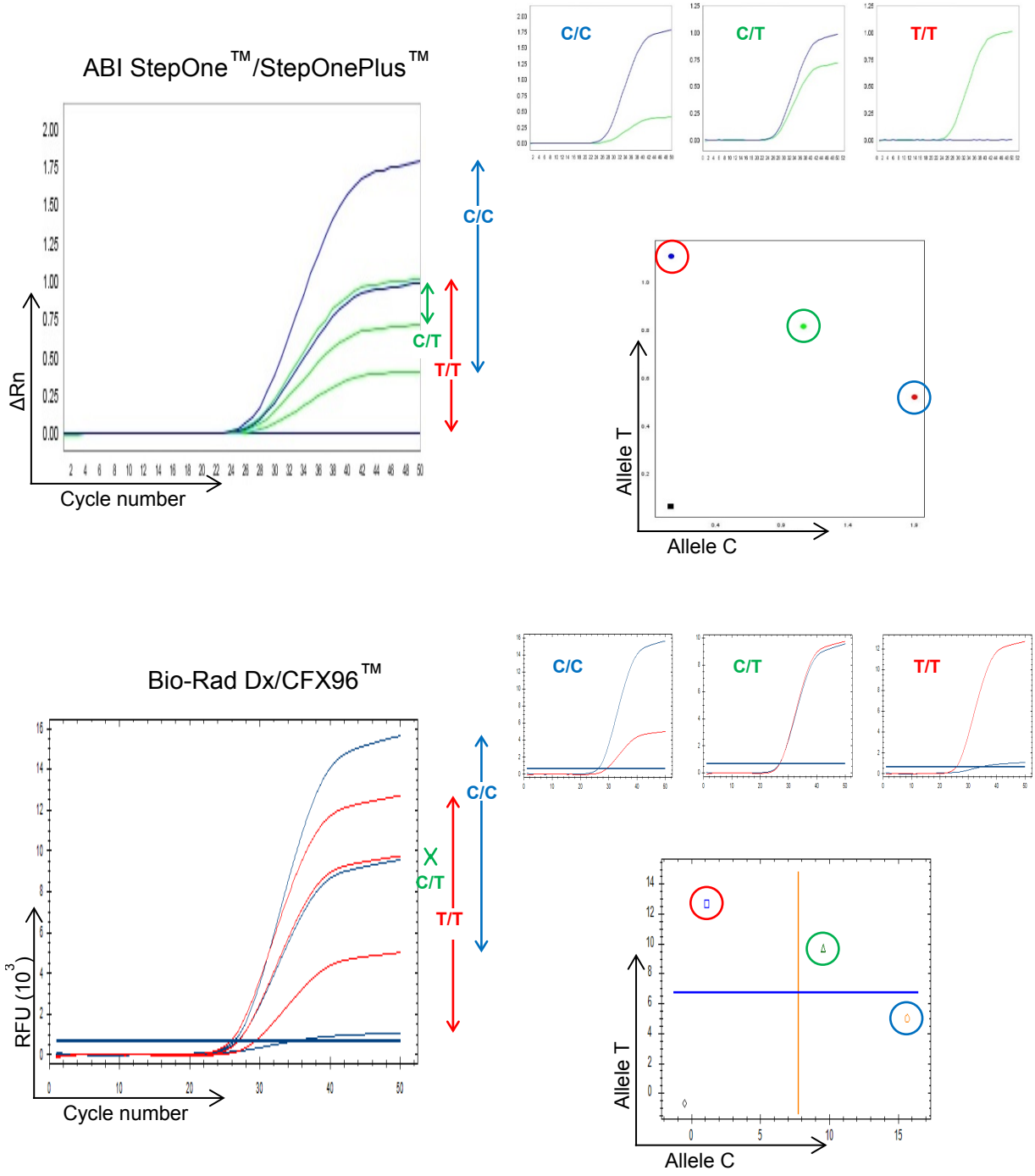


Fig. 27A: IL28B rs12979860 assay on ABI StepOne™/StepOnePlus™ and Bio-Rad Dx/CFX96™. Left: amplification curves of three samples with the different genotypes. ΔRn difference between curves of each sample is shown. Top right corner: details of amplification plot divided by genotype. Bottom right corner: allelic discrimination scatter plot of the same samples.

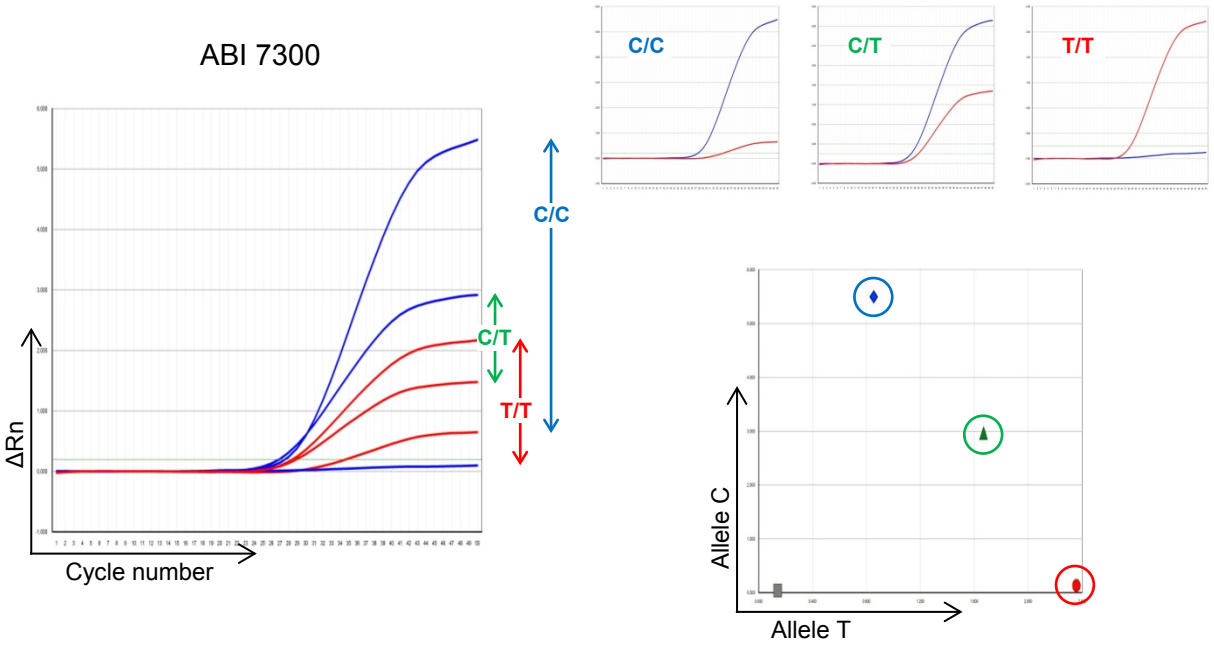


Fig. 27B: IL28B rs12979860 assay on ABI 7300. Left: amplification curves of three samples with the different genotypes. ΔRn difference between curves of each sample is shown. Top right corner: details of amplification plot divided by genotype. Bottom right corner: allelic discrimination scatter plot of the same samples.

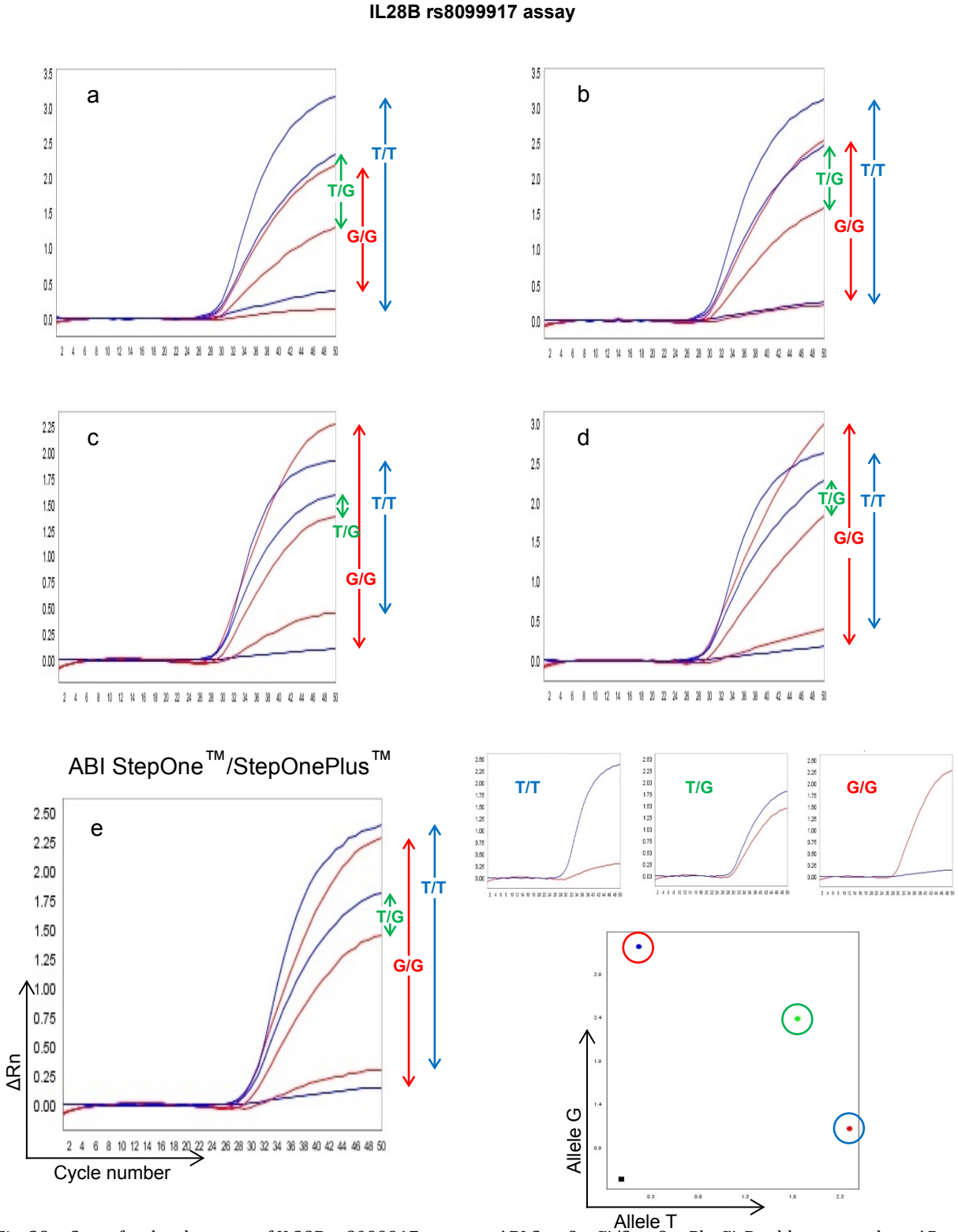


Fig. 28: Steps for development of IL28B rs8099917 assay on ABI StepOne™/StepOnePlus™. Double arrows show ΔRn difference between curves of each sample. a) initial mix, b) condition "a" with higher concentration of JOE™ probe, c) condition "b" with lower concentration of FAM™ probe, d) condition "c" with higher primers concentration, e) IL28B rs8099917 assay on ABI StepOne™/StepOnePlus™. In the bottom right corner are shown the details of amplification plot "e" divided by genotype and the corresponding allelic discrimination scatter plot.

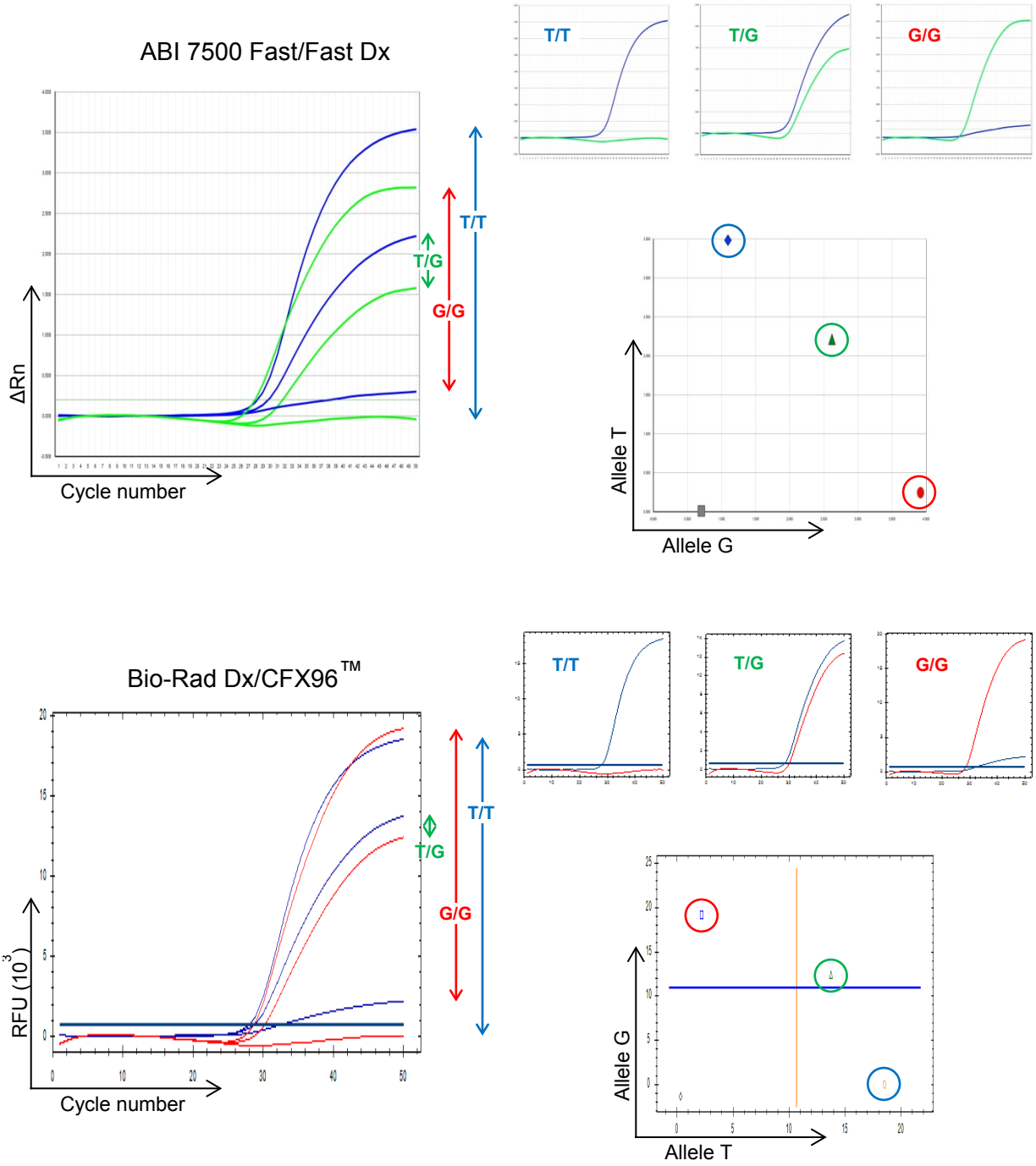


Fig. 29A: IL28B rs8099917 assay on ABI 7500 Fast/Fast Dx and Bio-Rad Dx/CFX96™. Left: amplification curves of three samples with the different genotypes. ΔRn difference between curves of each sample is shown. Top right corner: details of amplification plot divided by genotype. Bottom right corner: the allelic discrimination scatter plot of the same samples.

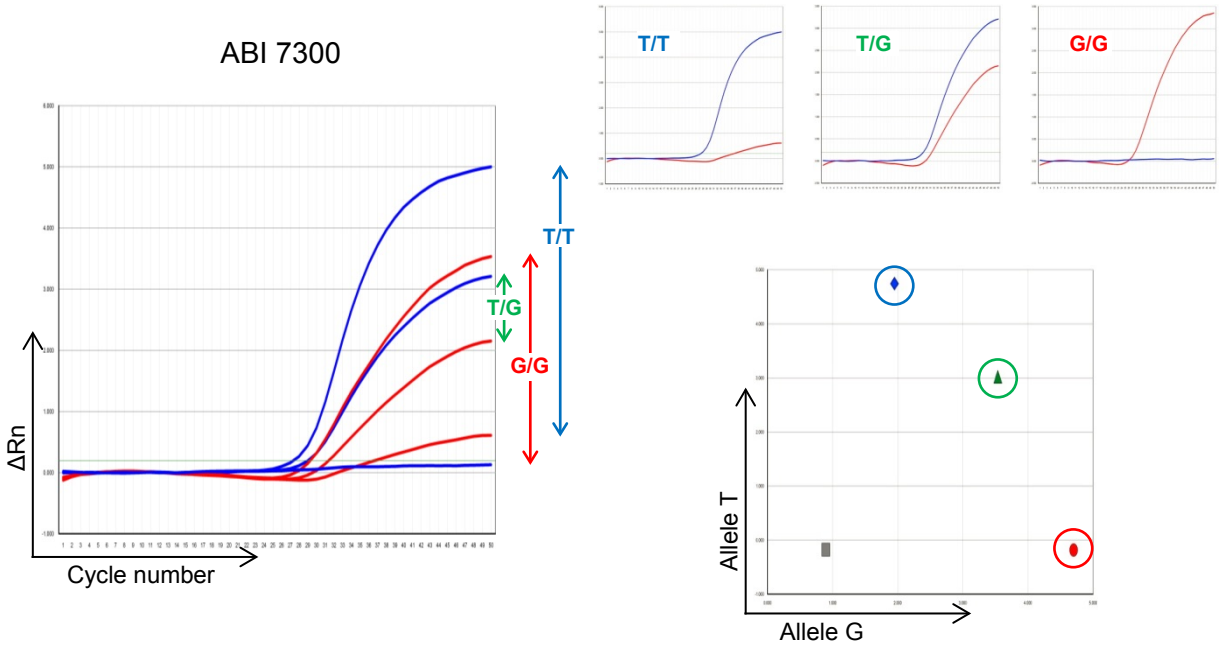


Fig. 29B: IL28B rs8099917 assay on ABI 7300. Left: amplification curves of three samples with the different genotypes. ΔRn difference between curves of each sample is shown. Top right corner: details of amplification plot divided by genotype. Bottom right corner: the allelic discrimination scatter plot of the same samples.

To summarize, IL28B rs12979860 and IL28B rs8099917 mixes conditions were different for each instrument. The exact composition of the mixes cannot be reported here, as they constitute proprietary reagents of AB ANALITICA. Both IL28B assays have the same amplification protocol (Tab. 11) and thermal profile (Tab. 12) for all real-time PCR systems used in this study.

Amplification protocol

Amplification reagents	Volume
IL28B amplification mix	24 µL
Positive controls/ DNA/ no template control	1 µL
Total	25 µL

Tab. 11: Amplification protocol of IL28B rs12979860 and IL28B rs8099917 assays.

Thermal profile

PCR phase	Temperature	Time	Cycles
UNG activation	50°C	2 minutes	1
Taq activation	95°C	10 minutes	1
Denaturation	95°C	10 seconds	50
Annealing/Extension	60°C	45 seconds	

Tab. 12: Thermal profile of the amplification reaction of IL28B rs12979860 and IL28B rs8099917 assays.

POSITIVE CONTROLS

For each assay, several clones of the three control plasmids were prepared. The heterozygous control was prepared as a 1:1 mixture of the two homozygous controls. Different dilutions of these controls were tested with samples on the real-time PCR systems. Since positive controls must mimic clinical samples, we selected as controls those plasmids that produced signals that were closest (similar C_T and ΔRn) to the signals of clinical samples. Two examples are shown in Fig. 30 for IL28B rs12979860 assay and in Fig. 31 for IL28B rs8099917 assay, respectively.

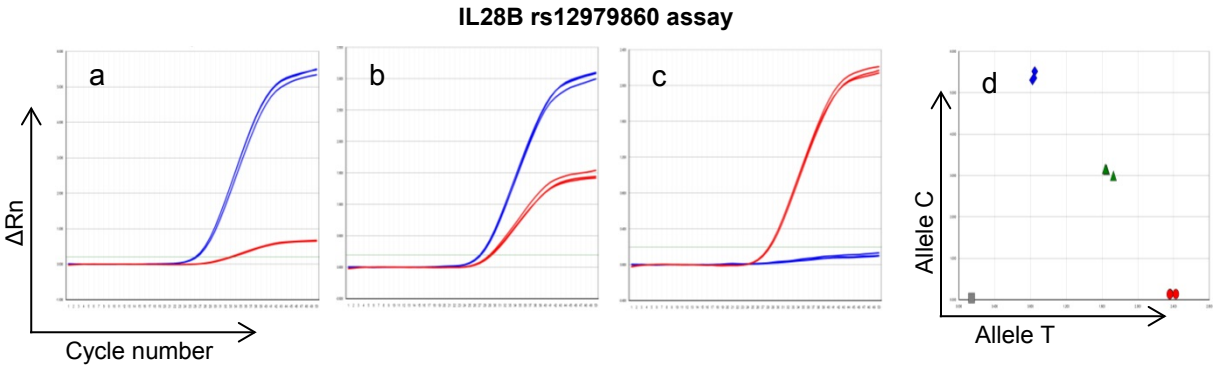


Fig. 30: Amplification and allelic discrimination plots of two plasmid clones and one clinical sample on ABI 7300. Amplification plots show very similar C_T and ΔRn values of controls and sample curves for the three genotypes: C/C (a), C/T (b), T/T (c). Scatter plot (d) shows very similar genotyping results of controls and sample for three genotype groups.

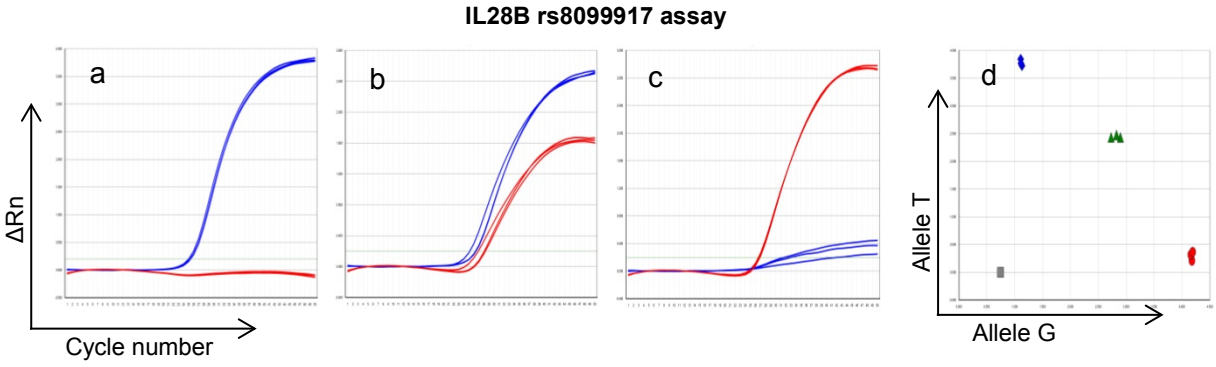


Fig. 31: Amplification and allelic discrimination plots of two plasmid clones and one clinical sample on ABI 7500 Fast/Fast Dx. Amplification plots show very similar C_T and ΔRn values of controls and sample curves for the three genotypes: T/T (a), T/G (b), G/G (c). Scatter plot (d) shows very similar genotyping results of controls and sample for three genotype groups.

6.2.2 CALR MUTATION ASSAY

We tested different experimental conditions to optimize the CALR MUTATION assay. We tested samples of known mutational status with different master mixes (Fig. 32 a, b, c), different concentrations of primers (Fig. 32 d, e, f) and different number of amplification cycles (Fig. 32 g, h). The visualization of the PCR products by gel electrophoresis was optimized testing different types of agarose (Fig. 32 i, j, k), gels of varying densities (Fig. 32 l, m), different DNA-intercalating agents (Fig. 32 n, o) and loading different volumes of PCR product into the gel (Fig. 32 p, q). Step by step, we compared the different conditions considering both brightness and resolution of bands and selected those allowing higher discrimination of the specific bands (orange rectangles in Fig. 32).

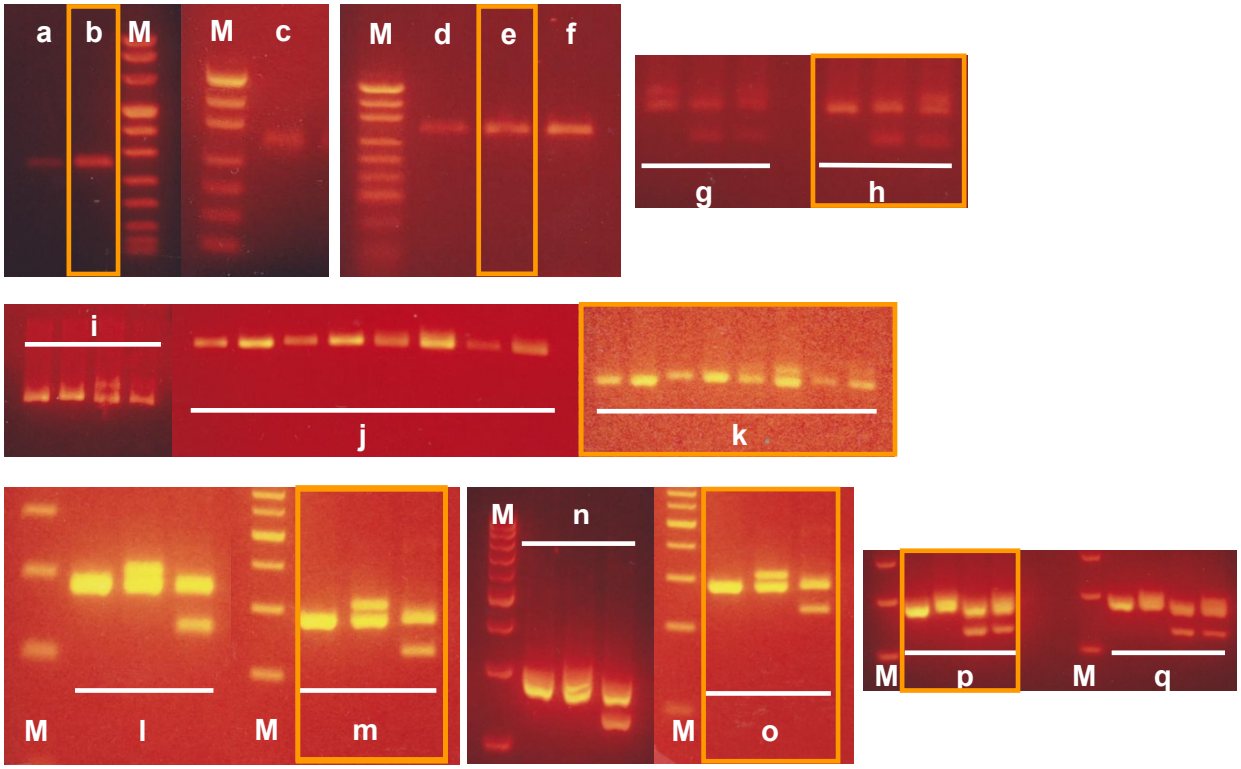


Fig. 32: Gels showing steps of CALR MUTATION assay optimization. Step 1: testing different master mixes (a, b, c); Step 2: testing different primers concentrations (d, e, f); Step 3: testing different number of amplification cycles (g, h); Step 4: testing different types of agarose (i, j, k); Step 5: testing gels of varying densities (l, m); Step 6: testing different DNA-intercalating agents (n, o); Step 7: testing different volumes of PCR product to load in the gel (p, q). Orange rectangles indicate the condition chosen step by step. M, DNA marker.

The positive controls were optimized testing different dilutions of the DNA plasmids (Fig. 33). We selected the higher dilution of plasmids because it allowed the discrimination of the two bands specific for Type-2 mutation (Fig. 33 b).

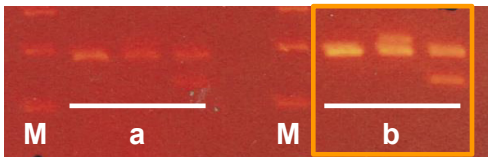


Fig. 33: Two dilutions of plasmids with the wt, Type-2 and Type-1 sequences. The orange square indicates the condition chosen. M, DNA marker.

Our final amplification protocol uses 20 µL of reaction mix and 5 µL of DNA (Tab. 13). For the thermal profile of the amplification reaction see Tab. 14. The final composition of the amplification mix is proprietary of AB ANALITICA and cannot be reported here.

Amplification protocol

Amplification reagents	Volume
CALR Master Mix	20 µL
Positive controls / DNA / no template control	5 µL
Total	25 µL

Tab. 13: Amplification protocol of CALR MUTATION assay.

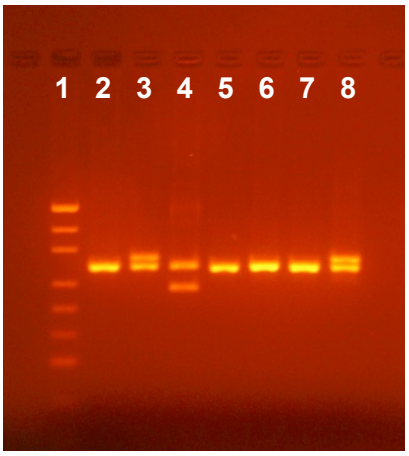
Thermal profile

Steps	Temperature	Time	Cycles
Taq activation	95°C	2 minutes	1
Denaturation	95°C	30 seconds	35
Annealing	60°C	30 seconds	
Extension	72°C	30 seconds	
Final extension	72°C	10 minutes	1

Tab. 14: Thermal profile of the amplification reaction of CALR MUTATION assay.

To increase specificity we established an additional step (post-PCR treatment) that has to be performed after the PCR. After the amplification reaction, each PCR product was divided into two aliquots. One aliquot of each sample was treated with a post-PCR reagent and denatured at 95°C for 5 minutes. The composition and mechanism of action of this post-PCR reagent are proprietary of AB ANALITICA and for this reason cannot be discussed here. For the visualization of PCR products, we prepared a 4% agarose with 1X TAE and 0.5 µg/mL of ethidium bromide. A volume of 5 µL of the treated and untreated PCR products was loaded into the gel, together with 2 µL of a loading buffer containing bromophenol blue. As molecular weight reference we loaded 10 µL of a DNA ladder. Selected visualization conditions and post-PCR treatment allowed identification and discrimination of mutated samples with a sequence length very close to the

wild-type (wt) sequence (e.g., insertions or deletions of only 4 bp, data not shown) and analysis of homozygous mutated samples (100% of mutant allele, Fig. 34).



1) DNA Marker
2) wt control
3) Type-2 control
4) Type-1 control
5) wt sample
6) Type-2 sample (100%)
7) Type-2 sample (100%) with post-PCR reagent without denaturation
8) Type-2 sample (100%) with post-PCR reagent with denaturation

Fig. 34: Gel visualization of positive controls (2, 3, 4), a wt sample (5) and a Type-2 mutated sample (6, 7, 8). Treatment of PCR product with post-PCR reagent and denaturation at 95°C for 5 minutes allow identification of 100% mutated samples with a sequence length ± 5 bp (8) compared to the wt sequence.

6.2.3 MPL W515L/K ASSAY

The MPL W515L/K assay was based on two multiplex mixes for the specific amplification of the wild-type *MPL* sequence and the *MPL* sequence containing the mutations W515L or W515K. We carried out a preliminary optimization on the Applied Biosystems StepOnePlus™ Real-Time PCR System. To optimize the amplification mix, we started testing different master mixes for the *MPL* W515L target (Fig. 35 a, b, c, d). We selected the master mix that showed curves with higher similarity in terms of C_T and ΔRn between the wild-type control (0%) and a wild-type sample and between the wild-type and mutated (100%) controls (Fig. 35 d). Using the selected master mix, we tested different primer and probe concentrations with the 100%, 1% and 0% controls for both the MPL W515L (Fig. 35 e, f, g) and the MPL W515K (Fig. 35 h, i, j) specific mixes. We selected the conditions that produced the highest difference in ΔRn ratios ($\Delta Rn_{MUT}/\Delta Rn_{WT}$) between the 0% and 1% controls (Fig. 35 g, j). Furthermore, the selected condition for the MPL W515K amplification mix (Fig. 35 j) did not show inhibited curves compared to the other conditions tested (Fig. 35 h, i).

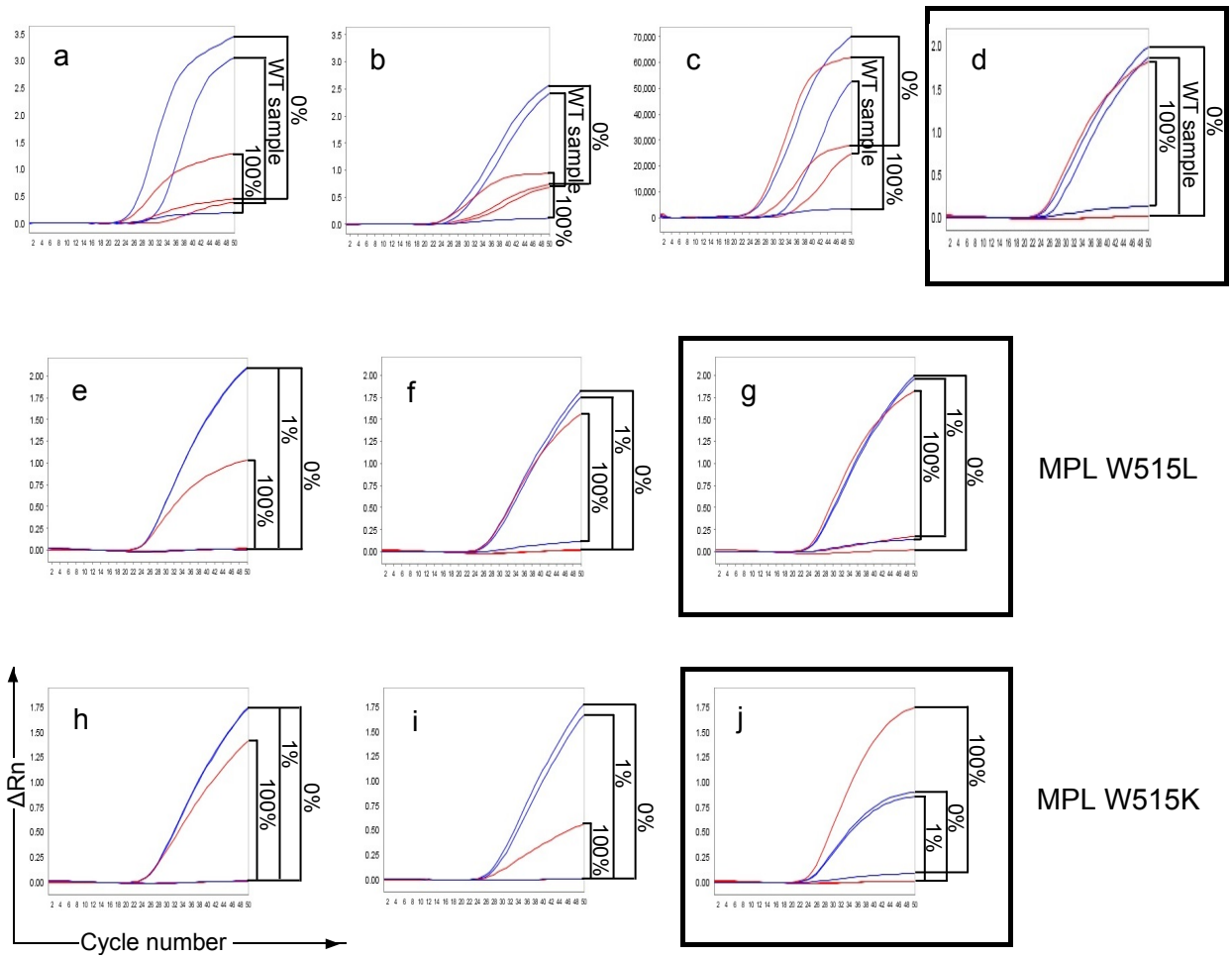


Fig. 35: Amplification plots showing steps for optimization of the MPL W515L/K amplification mixes. Step 1: testing different master mixes (a, b, c, d); Step 2: testing different primer and probe concentrations for MPL W515L mix (e, f, g) and MPL W515K mix (h, i, j). Amplification plots with conditions chosen step by step are highlighted in black squares.

We prepared serial dilutions of the 100% MPL W515L and MPL W515K controls on background of the wild-type control to obtain positive controls with 70%, 50%, 30%, 10%, 1.5%, 1% of the mutated allele positive controls. In order to determine the appropriate plasmid concentration for controls, we tested the 100%, 1.5%, 1% and 0% controls at two different concentrations and calculated the ΔRn ratio of the two targets in the reaction ($\Delta Rn_{MUT}/\Delta Rn_{WT}$). We decided to use the higher concentration for the positive controls as it showed a larger difference ($\Delta(\Delta R_n)_{X-0\%}$) between the ΔRn ratios of the the low positive controls (1% and 1.5%) and the 0% (wild-type) control for MPL W515K (orange rectangle in Tab. 15). The two concentrations of controls showed comparable results with the MPL W515L mix (Tab. 15).

	ΔRn ratio W515K/WT	W515K $\Delta(\Delta R_n)_{X-0\%}$	ΔRn ratio W515L/WT	W515L $\Delta(\Delta R_n)_{X-0\%}$
100%	11.165		3.093	
1.5% higher concentration	0.903	0.174	0.123	0.027
1% higher concentration	0.841	0.112	0.110	0.014
1.5% lower concentration	0.817	0.088	0.120	0.024
1% lower concentration	0.776	0.047	0.113	0.017
0%	0.729		0.096	

Tab. 15: Comparison between two concentrations of positive controls (higher and lower) calculated using the low positive controls 1.5% and 1% (possible cut-offs of the assay) for both target mixes. $\Delta(\Delta Rn)_{X-0\%}$ is the difference between ΔRn of the positive control and ΔRn of the wild-type control (0%). An assay with high $\Delta(\Delta Rn)_{CUTOFF-0\%}$ values presumably has high diagnostic specificity, too. The highest values were obtained with controls at higher concentrations using the MPL W515K mix (orange rectangle).

We developed a prototype of the MPL W515L/K assay based on amplification of 5 µL of DNA in a total reaction volume of 25 µL (Tab. 16). The amplification mix is a proprietary reagent of AB ANALITICA and therefore cannot be discussed here. The thermal profile of the amplification reaction is reported in Tab. 17.

Amplification protocol

Amplification reagents	Volume
MPL W515L/K amplification mix	20 µL
Positive controls/ DNA/ no template control	5 µL
Total	25 µL

Tab. 16: Amplification protocol of MPL W515L/K prototype.

Thermal profile

Steps	Temperature	Time	Cycles
UNG activation	50°C	2 minutes	1
Taq activation	95°C	10 minutes	1
Denaturation	95°C	15 seconds	50
Annealing/Extension	60°C	1 minute	

Tab. 17: Thermal profile of the amplification reaction of the MPL W515L/K prototype.

The prototype showed ΔRn values proportional to the dilution of the positive controls (100%, 70%, 50%, 30%, 10%, 1.5%, 1% and 0%). Larger intervals between the ΔRn values were obtained for the controls with higher percentages of the mutated allele (Fig. 36). This translates into a higher resolution in the range of 10 to 100% compared to the range of 0 to 1.5% of mutated allele.

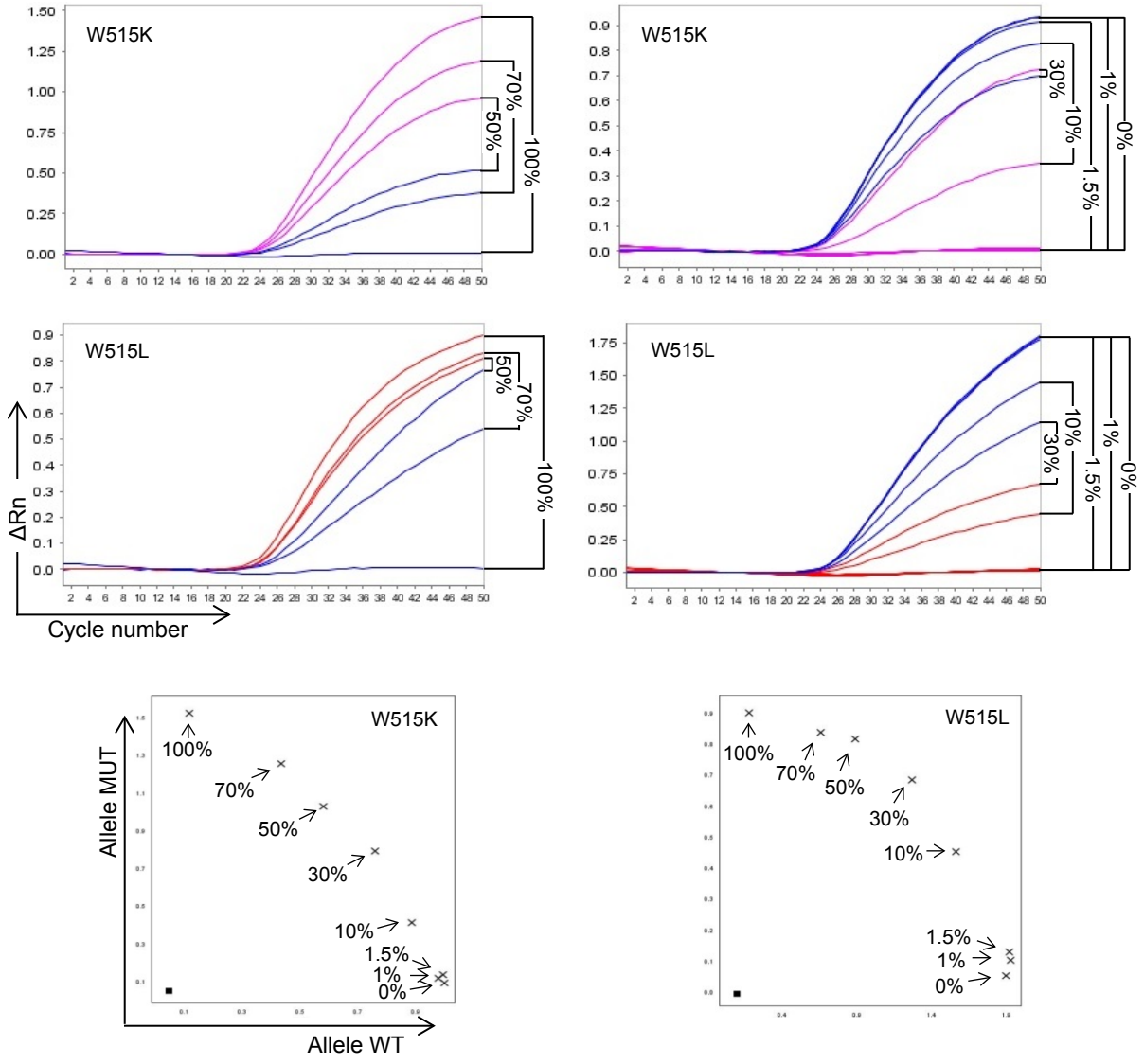


Fig. 36: Amplification and allelic discrimination plots of the 100%, 70%, 50%, 30%, 10%, 1.5%, 1%, 0% positive controls with the MPL W515L/K prototype.

6.3 PERFORMANCES

6.3.1 IL28B ASSAYS

ANALYTICAL SPECIFICITY

To ensure a high specificity of the assays for analysis of the *IL28B* polymorphisms rs12979860 and rs8099917, highly specific amplification primers and probes were selected and the protocol established with stringent reaction conditions. To test the specificity, the primer and probe sequences were aligned in public genome databases (NCBI Blast). The sequences showed no non-specific sequence pairing. Primer specificity was demonstrated by real-time PCR using a DNA binding dye that showed no unspecific amplification signals. The specific amplification was additionally confirmed by Sanger sequencing (an example is given in Fig. 43). Amplicons exclusively contained the target sequence and the sequencing produced a well-defined electropherogram with sharp peaks separated by even spaces and little background interference at the peak baseline. This indicates the absence of competing amplification products and a specific amplification of the target sequence.

ANALYTICAL SENSITIVITY

Analytical sensitivity of the IL28B rs12979860 and IL28B rs8099917 assays was tested using two clinical samples for each genotype at a DNA concentration of 2 ng/µL on ABI StepOne™/StepOnePlus™, ABI 7500 Fast/Fast Dx and Bio-Rad Dx/CFX96™. In all cases the samples were correctly genotyped (Tab. 18).

			Real-Time PCR System		
	Sample (2 ng DNA)	Genotype	ABI StepOne™/ StepOnePlus™	ABI 7500 Fast/Fast Dx	Bio-Rad Dx/CFX96™
IL28B rs12979860	RQ00081	C/C	✓	✓	✓
	RQ00086	C/C	✓	✓	✓
	RQ00104	C/T	✓	✓	✓
	RQ00301	C/T	✓	✓	✓
	RQ00079	T/T	✓	✓	✓
	RQ00084	T/T	✓	✓	✓
IL 28B rs8099917	RQ00086	T/T	✓	✓	✓
	RQ00104	T/T	✓	✓	✓
	RQ00100	T/G	✓	✓	✓
	RQ00101	T/G	✓	✓	✓
	RQ00084	G/G	✓	✓	✓
	RQ00594	G/G	✓	✓	✓

Tab. 18: Both IL28B assays correctly detected 2 ng/ µL DNA in samples with the three genotypes on ABI StepOne™/StepOnePlus™, ABI 7500 Fast/Fast Dx and Bio-Rad Dx/CFX96™.

With both assays, 2 ng of DNA per reaction (ng/rx) could still be analyzed (curves above threshold) and correctly genotyped by automatic genotype call on instruments that support this function (Fig. 37 and Fig. 38).

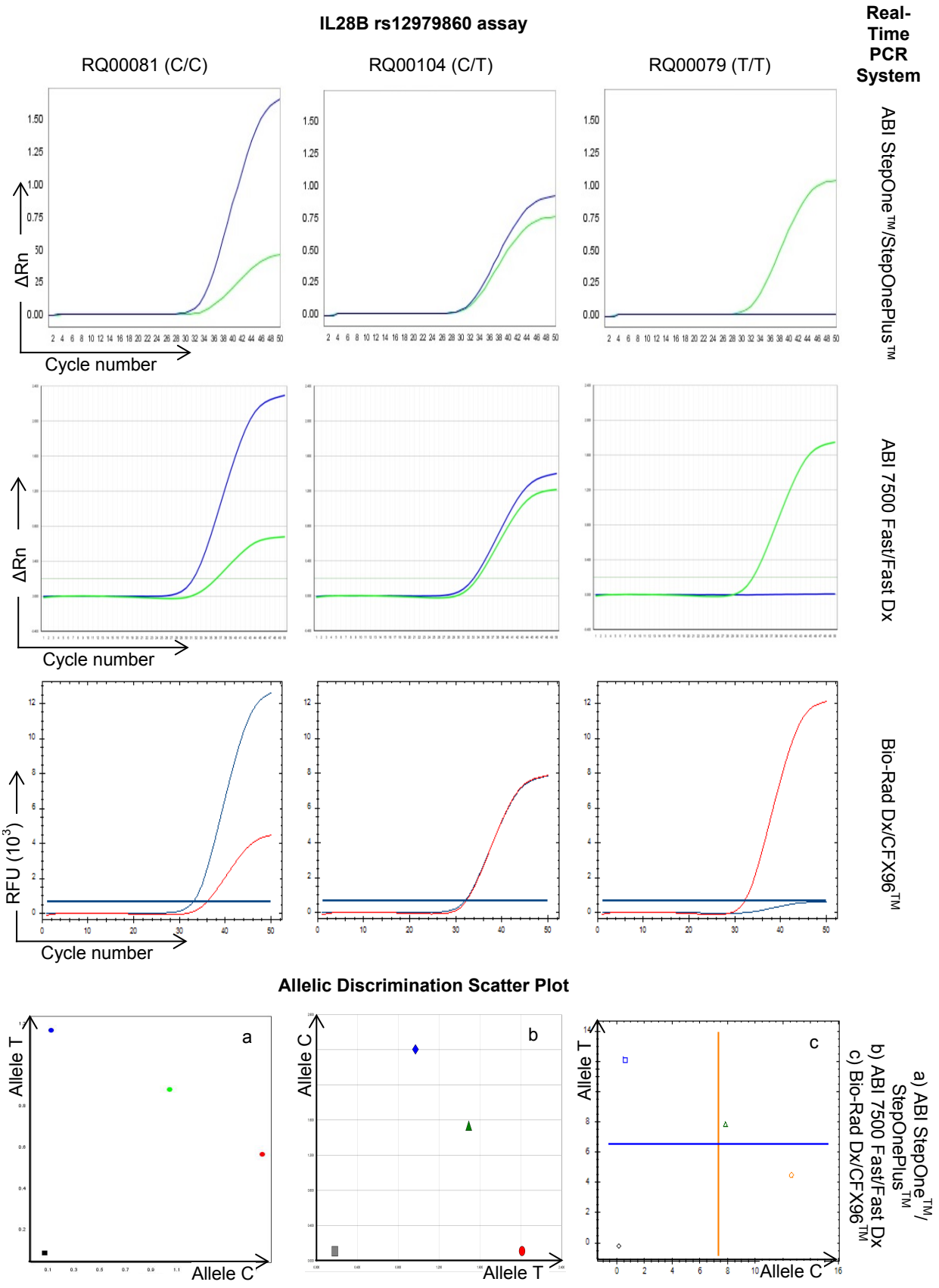


Fig. 37: Example of detection and allelic discrimination of 2 ng/rx DNA samples with the 3 rs12979860 genotypes.

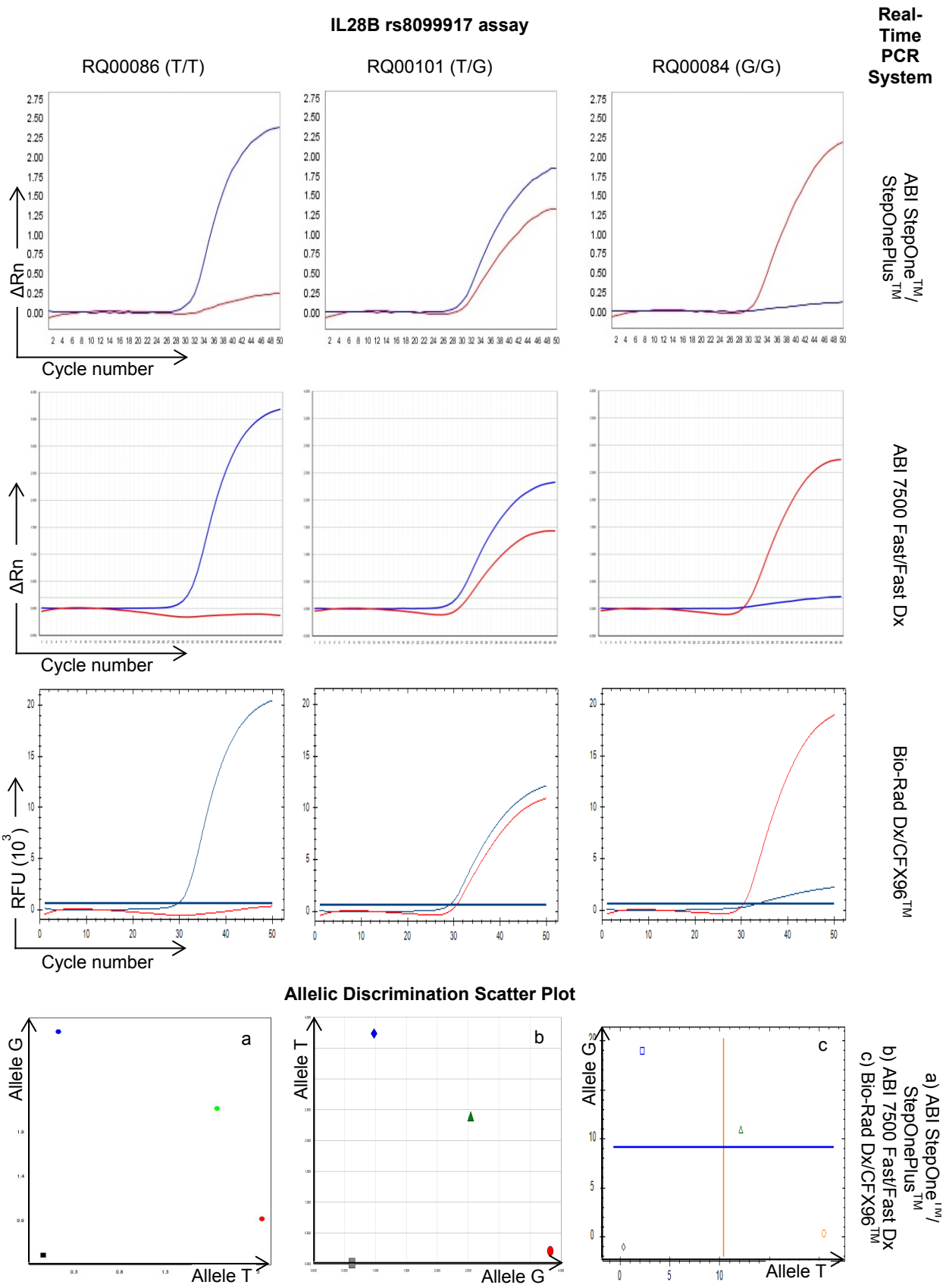


Fig. 38: Example of detection and allelic discrimination of 2 ng/rx DNA samples with the 3 rs8099917 genotypes.

RANGE OF DNA QUANTITY

To assess the compatibility of the assays with extraction systems that produce high DNA yields and to evaluate assay performances when using high concentrations of DNA, we tested clinical samples at a concentration of 250 ng. Both the IL28B rs12979860 and the IL28B rs8099917 assays were able to correctly assign the genotype of these samples.

DIAGNOSTIC SENSITIVITY AND SPECIFICITY

We determined the diagnostic sensitivity and specificity of the assays testing 200 (IL28B rs12979860) and 193 (IL28B rs8099917) clinical samples on Bio-Rad Dx/CFX96™. Overall, 199 of 200 samples (IL28B rs12979860) and 192 of 193 samples (IL28B rs8099917) were correctly genotyped. All correctly genotyped samples showed well-separated curves ("bundle of curves", Fig. 39 and Fig. 41) and could unequivocally be assigned to the correct genotype group by automatic call (Fig. 40 and Fig. 42). Based on these data, diagnostic sensitivity and diagnostic specificity were calculated (IL28B rs12979860: 99.50%; IL28B rs8099917: 99.48%). The assays failed to correctly analyze samples RQ00616 (rs12979860 C/T, identified as T/T) and RQ00613 (rs8099917 T/G, identified as T/T) probably due to sample degradation. The DNA sequence of these samples was checked by Sanger sequencing (RQ00613 sequence is reported in Fig. 43).
At the same time, we validated the assays on ABI 7300, ABI 7500 Fast/Fast Dx, ABI StepOne™/StepOnePlus™ using around 50 clinical samples (Fig. 39, Fig. 40, Fig. 41, Fig. 42).

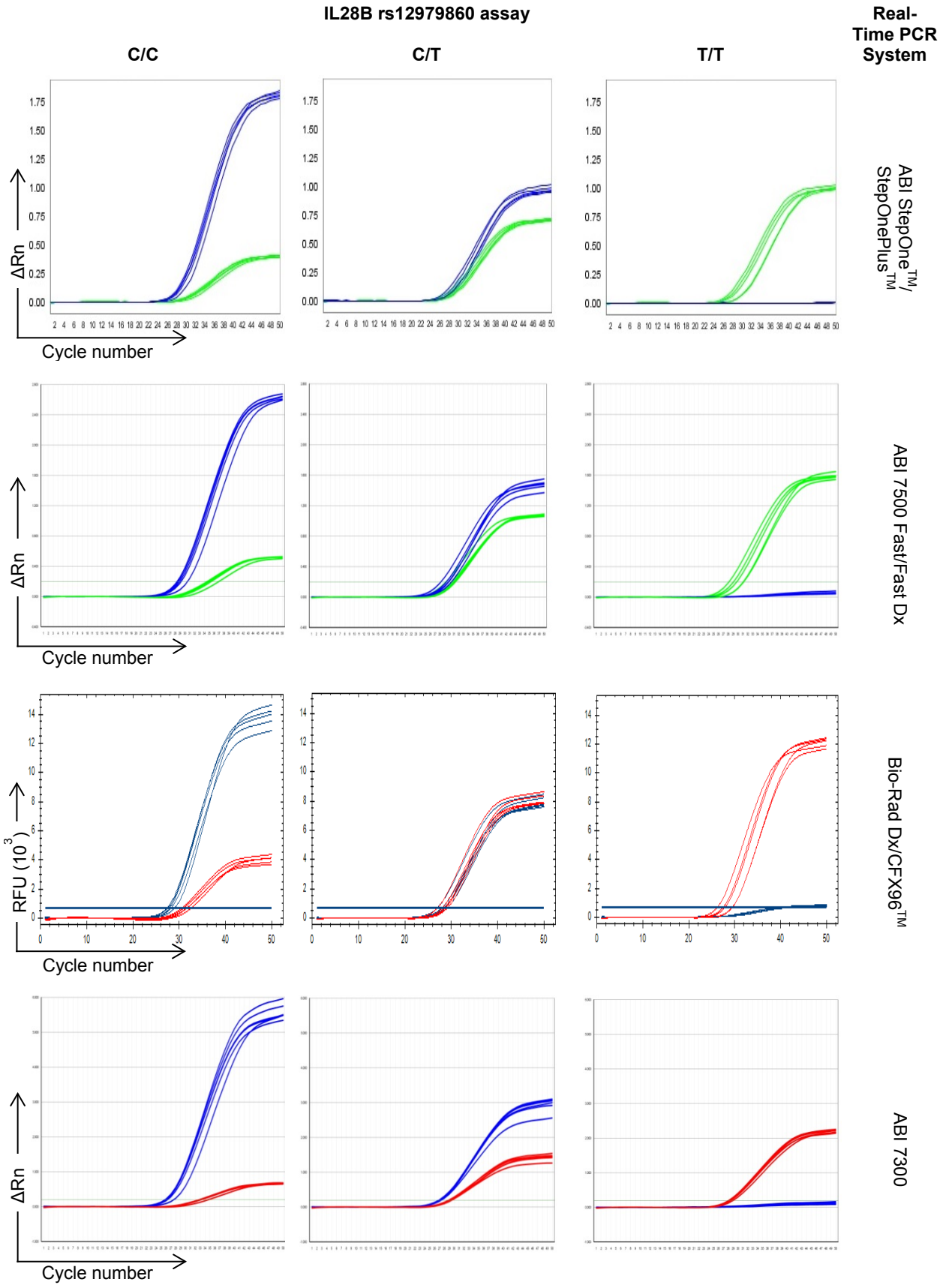


Fig. 39: Example of amplification plots of clinical samples with the three genotypes of rs12979860 on ABI StepOne™/StepOnePlus™, ABI 7500 Fast/Fast Dx, Bio-Rad Dx/CFX96™ and ABI 7300. Samples with the same genotype show the characteristic bundle of curves.

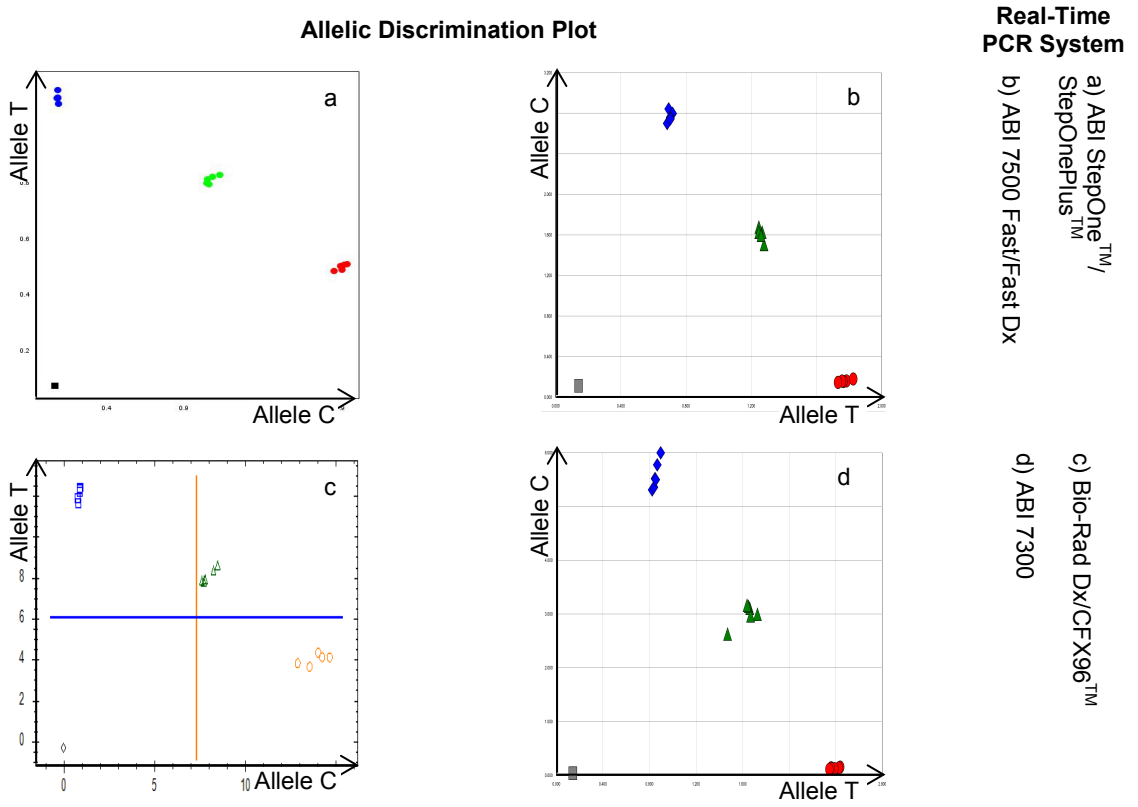


Fig. 40: Allelic discrimination plots of the same samples of Fig. 39 (rs12979860) on ABI StepOne™/StepOnePlus™ (a), ABI 7500 Fast/Fast Dx (b), Bio-Rad Dx/CFX96™ (c) and ABI 7300 (d). Genotype was correctly assigned by the automatic call of the instrument that formed three well-separated groups.

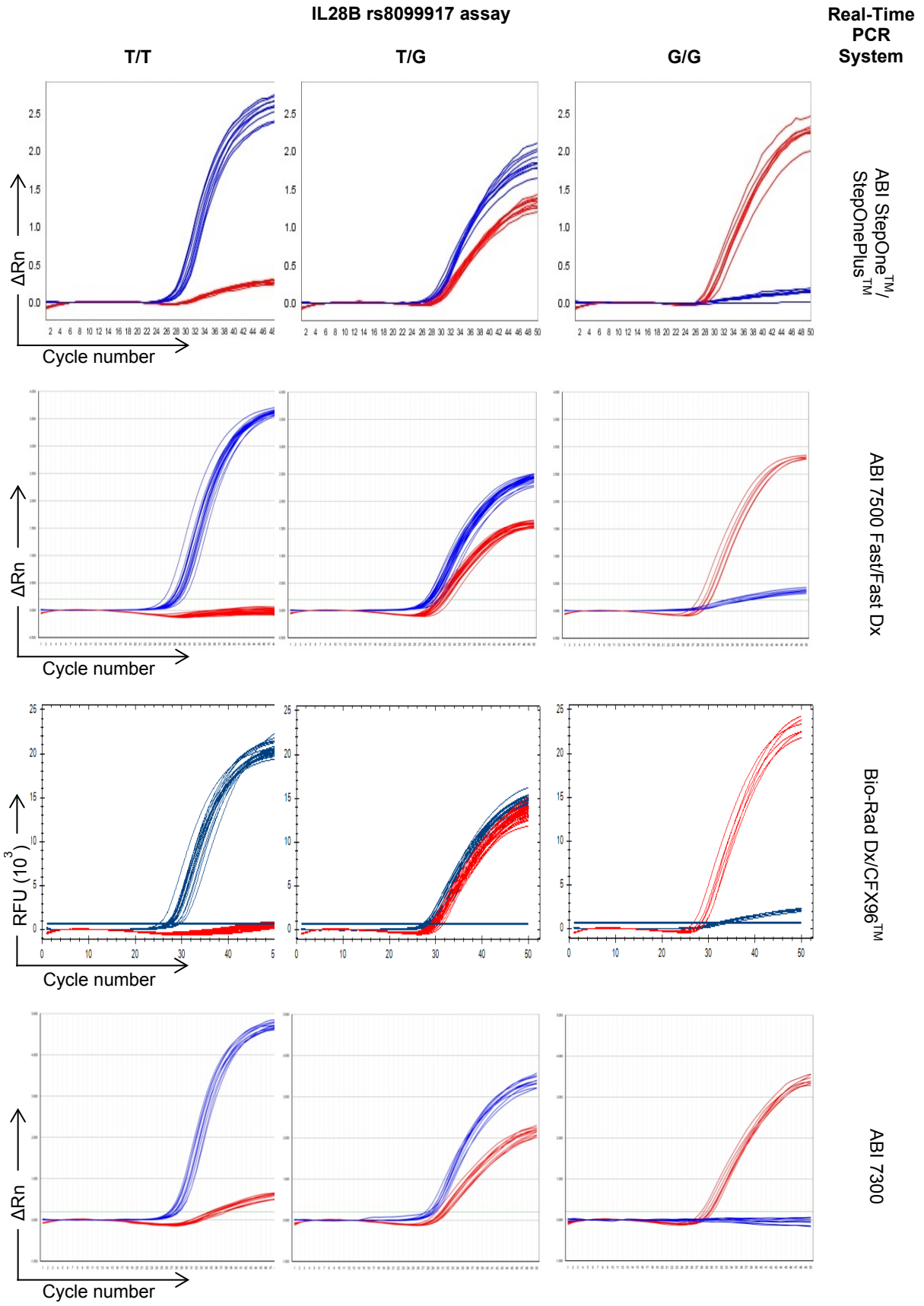


Fig. 41: Example of amplification plots of clinical samples with the three genotypes of rs8099917 on ABI StepOne™/StepOnePlus™, ABI 7500 Fast/Fast Dx, Bio-Rad Dx/CFX96™ and ABI 7300. Samples with the same genotype show the characteristic bundle of curves.

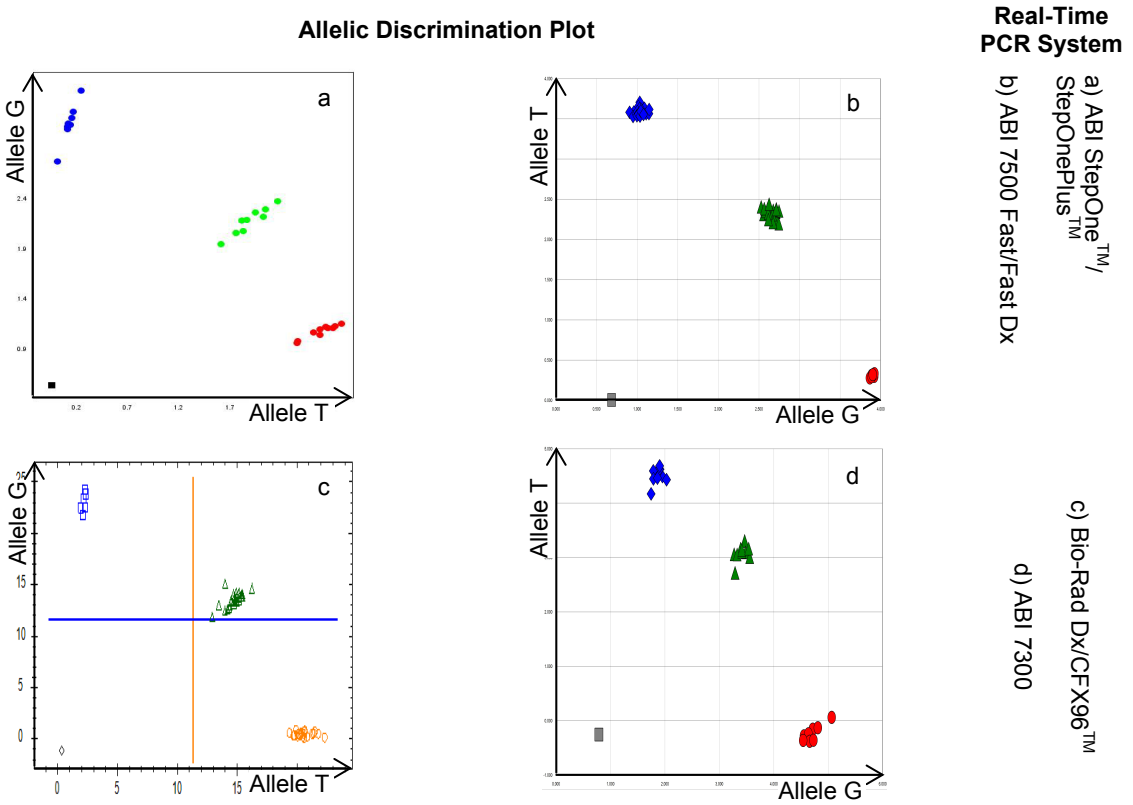


Fig. 42: Allelic discrimination plots of the same samples of Fig. 41 (rs8099917) on ABI StepOne™/StepOnePlus™ (a), ABI 7500 Fast/Fast Dx (b), Bio-Rad Dx/CFX96™ (c) and ABI 7300 (d). Genotype of samples was correctly assigned by the automatic call of the instrument that formed three well-separated groups.

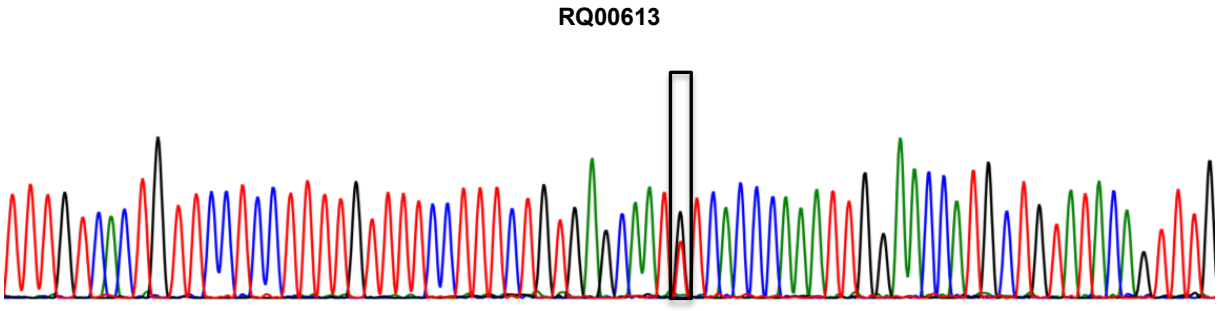


Fig. 43: Electropherogram of RQ00613 by Sanger sequencing. The rs8099917 locus clearly presents two overlapping peaks (in the rectangle).

STABILITY STUDY

We performed a stability study to determine the shelf-life of the reagents of the IL28B rs12979860 and IL28B rs8099917 assays. During this study aliquots of freshly prepared reagents were stored at the recommended storing conditions and were/will be tested at different points in time. Results were/will be compared to the data obtained at the starting point (T0). We prepared three batches of amplification mixes with reagents from different stock solutions and the positive controls for each genotype. We divided the mixes and the controls into aliquots in order to avoid that any reagent undergoes more than two freeze-thaw cycles, which might lead to degradation of the components due to excessive temperature variations. The aliquots were stored at -20°C/-30°C. Each batch of the amplification mixes was tested on the Dx Real-Time System (Bio-Rad) directly after preparation (T0) and at the following points in time: T3, T8, T12, T14, T16, T18 and T20 (months after preparation) (IL28B rs12979860); T6, T10 and T14 (months after preparation) (IL28B rs8099917). The stability study of the IL28B rs8099917 assay is in progress (T18 and T20 still need to be tested). We included larger time intervals for the IL28B rs8099917 assay evaluation, owing to the robustness observed for the IL28B rs12979860 reagents. At each data point we evaluated the obtained C_T and RFU values based on our experience with similar IVDs according to the European Standard EN 13640:2002[1]. We did not observe any significant variation of the C_T values at T20 and T14 (months) for the IL28B rs12979860 assay and the IL28B rs8099917 assay, respectively (Fig. 44). Although we detected fluctuations in the RFU values (Fig. 45), these are unlikely to result from mix degradation. A degradation of reagents would cause the RFU values to gradually decrease over time. In contrast, the observed variations include decreased and increased RFU values. This, in combination with the fact that all batches and both assays show the same course of variation, indicates that the fluctuations are probably caused by varying performance of the real-time PCR machine over time. Therefore, we assigned both assays, IL28B rs12979860 and IL28B rs8099917, a shelf-life of 1 year.

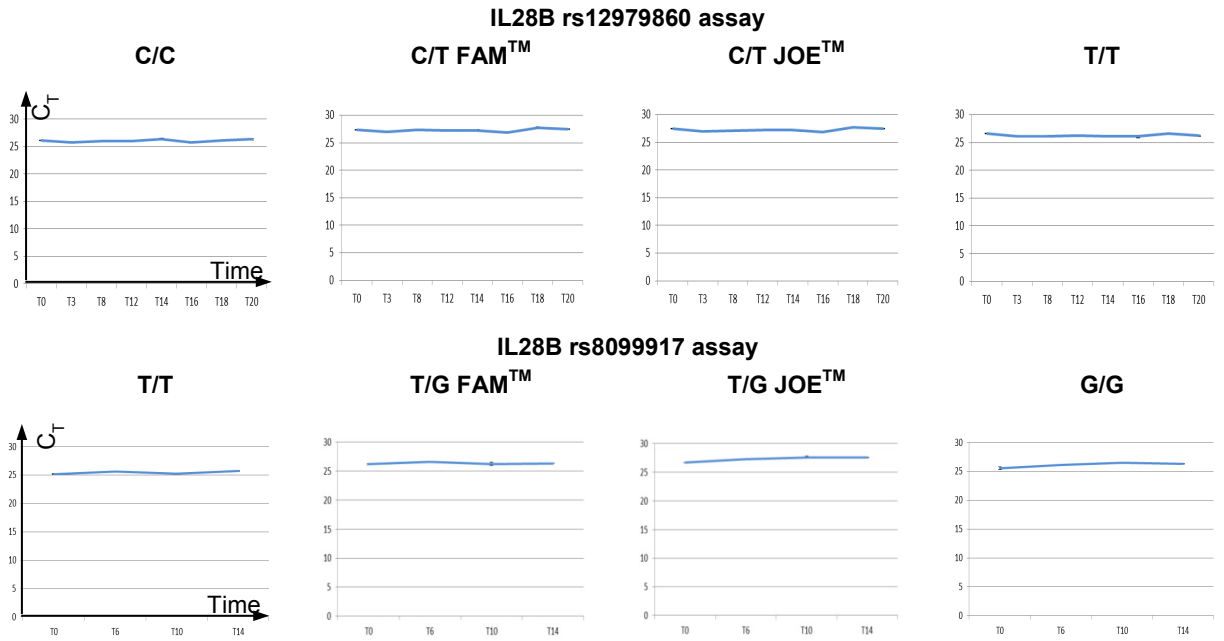


Fig. 44: C_T values of fluorescence signals of IL28B rs12979860 and IL28B rs8099917 assays at T0, T3, T8, T12, T14, T16, T18, T20 (rs12979860) and at T0, T6, T10, T14 (rs8099917) on Bio-Rad Dx.

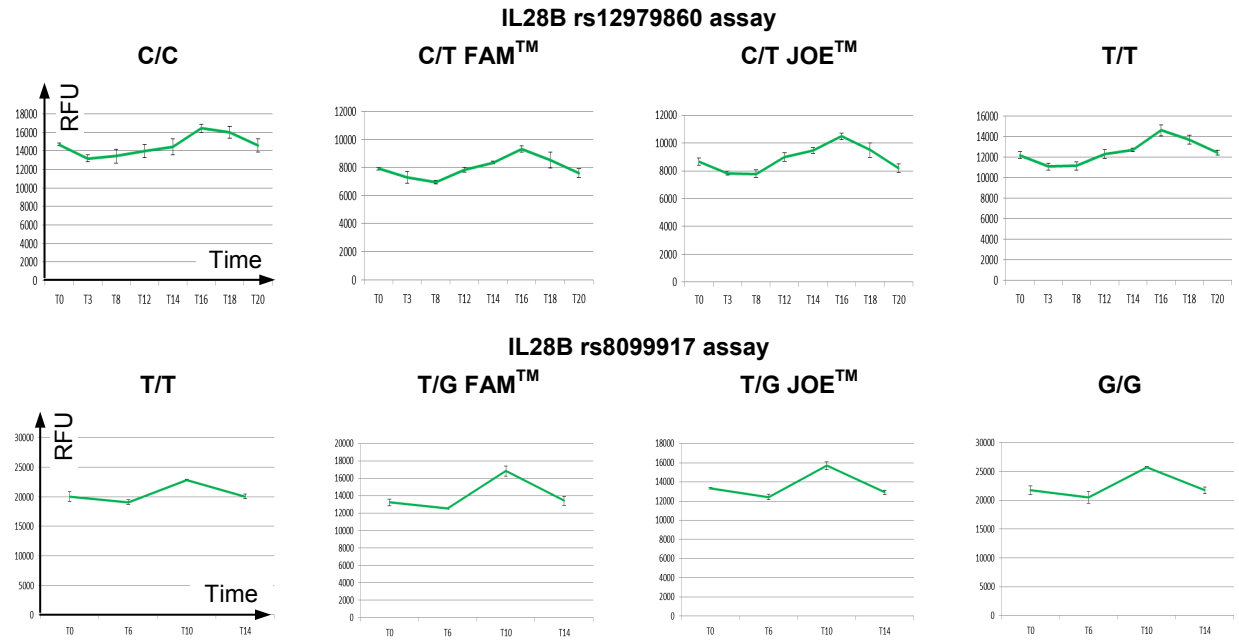


Fig. 45: RFU values of fluorescence signals of IL28B rs12979860 and IL28B rs8099917 assays at T0, T3, T8, T12, T14, T16, T18, T20 (rs12979860) and at T0, T6, T10, T14 (rs8099917) on Bio-Rad Dx.

6.3.2 CALR MUTATION ASSAY

ANALYTICAL SPECIFICITY

To ensure a high specificity of the CALR MUTATION assay, highly specific amplification primers were selected and the protocol established with stringent reaction conditions. To test the specificity, primer sequences were aligned in public genome databases (NCBI Blast). The sequences showed no nonspecific sequence pairing. In addition, the assay showed no nonspecific bands and primer-dimers by gel electrophoresis. This was also confirmed by Sanger sequencing of PCR products. The tested amplicons exclusively contained the target sequence and the sequencing produced well-defined electropherograms with sharp peaks separated by even spaces and little background interference at the peak baseline (Fig. 51), which indicates absence of competing amplification products and specific amplification of the target sequence. Mutation-specific detection was confirmed by the post-PCR treatment of samples, the optimized visualization on a 4% agarose gel and the presence of specific controls:

- post-PCR treatment allowed identification of mutations with a sequence length very similar to the wild-type sequence (± 4-5 bp);
- 4% agarose gels allow proper band separation;
- comparison with positive controls facilitates discrimination of various sample sequences (wild-type/insertions/deletions).

ANALYTICAL SENSITIVITY

We determined the limit of detection (LOD) of the CALR MUTATION assay by testing serial dilutions of a plasmid containing a mutated sequence on the background of a plasmid with the wild-type sequence. We tested the ratios 100%, 70%, 50%, 25%, 10% and 5% for the Type-1/Type-2 mutations and a rare mutation (sequence length -4 bp than the wild-type sequence). The LOD was 5% of mutated allele for the Type-2 and the rare mutation and 10% for the Type-1 mutation (Fig. 46). Therefore, we assigned the CALR MUTATION assay a final LOD of 10% of mutated allele.

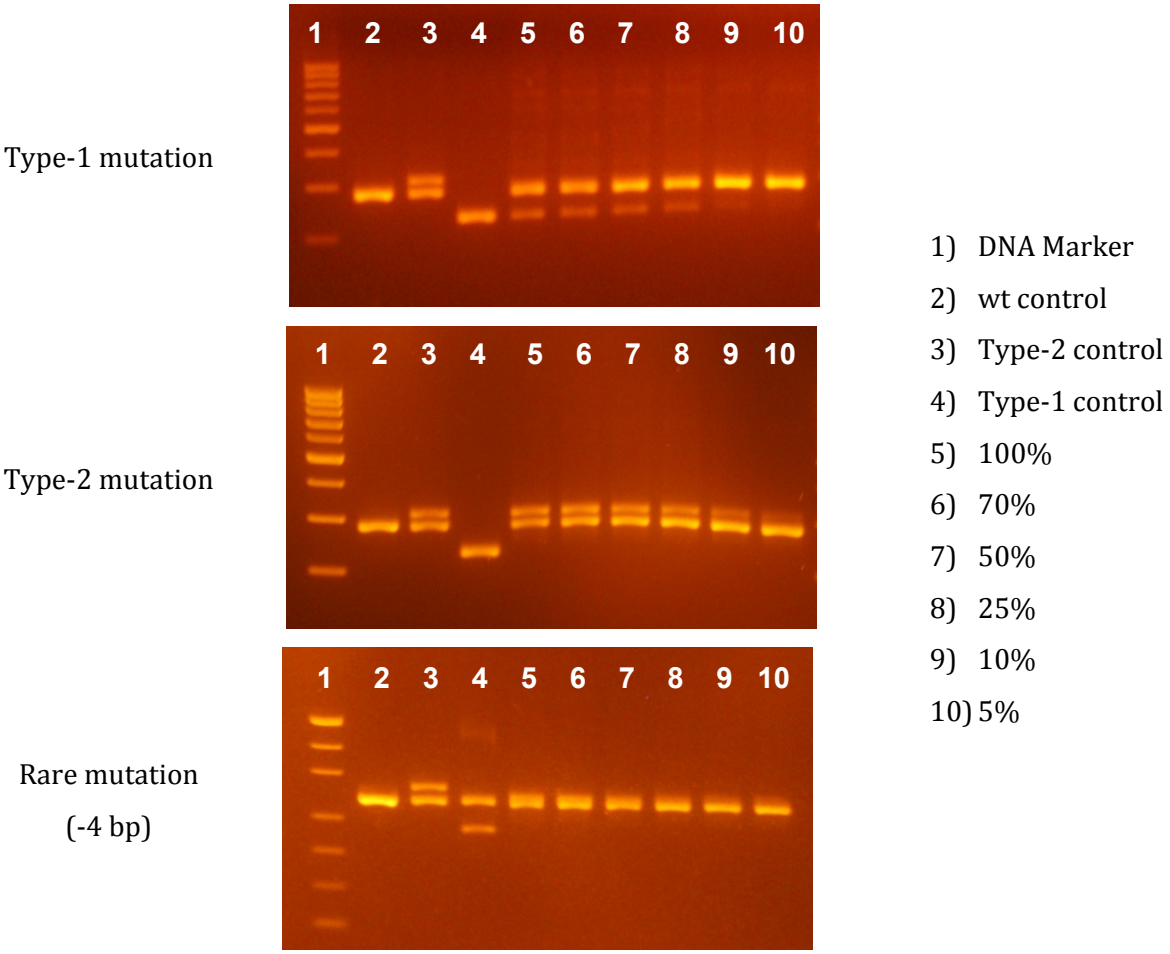


Fig. 46: Gels with positive controls (2, 3, 4) and serial dilutions (5, 6, 7, 8, 9, 10) of plasmids containing a mutated sequence (Type-1/Type-2 mutations and a rare mutation). The two specific bands are visible till the 5% (Type-2 and rare mutation) and the 10% (Type-1 mutation) dilution of mutated allele in a wild-type background. Therefore, the LOD of the assay was established at 10% of mutated allele.

DETECTABLE RANGE OF DNA

To confirm the correct detection of samples with different DNA concentrations using the CALR MUTATION assay, we tested the 50% and 25% dilutions of the Type-1 and Type-2 mutation controls and the wild-type (wt) plasmid control at 3 different concentrations (5 ng/µL, 10 ng/µL and 20 ng/µL of total DNA). Similar to the dilutions used to determine the LOD, the 50% and 25% dilutions were prepared on background of the wt plasmid control. The CALR MUTATION assay clearly detected samples in the concentration range between 5 ng/µL (corresponding to 25 ng/rx) and 20 ng/µL (100 ng/rx). Specific bands were well visible for both 50% and 25% dilutions (Fig. 47).

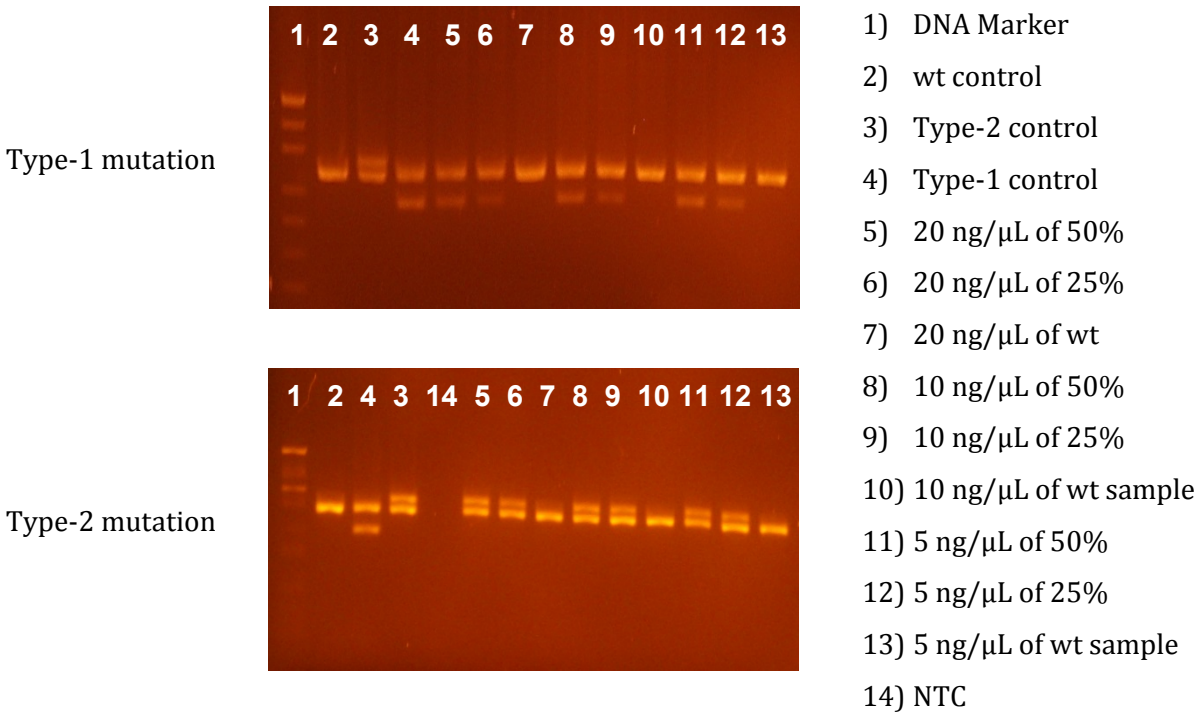
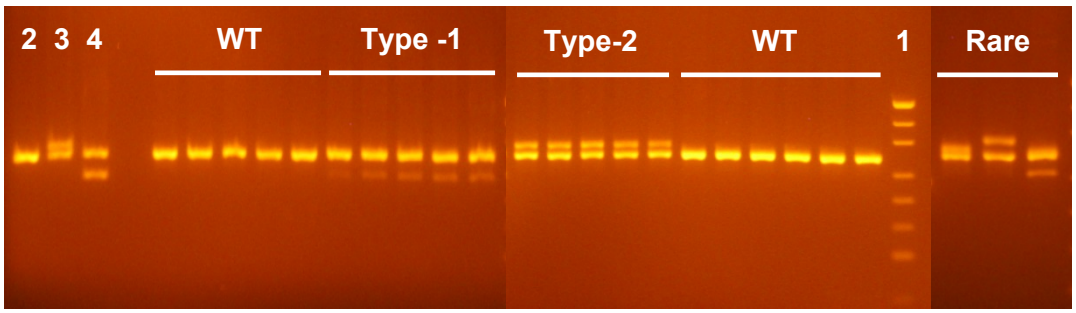


Fig. 47: Gels with positive controls (2, 3, 4), 25% and 50% mutated plasmids (Type-1 and Type-2 mutations) and wt plasmid at different concentrations (5 ng/µL, 10 ng/µL and 20 ng/µL) of total DNA (5 to 13). All target-specific bands are visible in this range of DNA concentration.

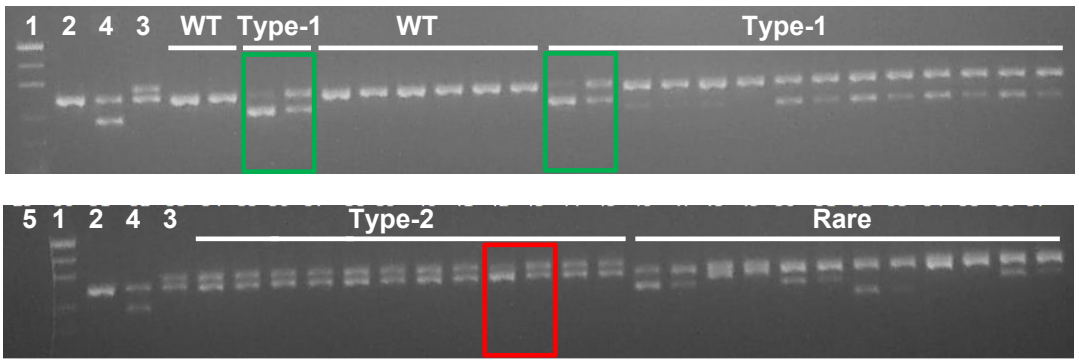
DIAGNOSTIC SENSITIVITY AND SPECIFICITY

We determined diagnostic sensitivity and specificity of CALR MUTATION assay testing a total of 34 clinical samples (21 wild-type samples, 5 samples with Type-1 mutation, 5 samples with Type-2 mutation, 3 samples with rare mutations). The CALR MUTATION assay correctly detected all tested samples (Fig. 48). According to these results, the diagnostic sensitivity and specificity of the assay were 100%. Additionally, we received results from the partner lab of the University of Florence testing the CALR MUTATION assay (external evaluation) on additional 31 clinical samples (15 wild-type samples, 6 samples with Type-1 mutation, 5 samples with Type-2 mutation and 5 samples with rare mutations). All tested samples were correctly detected by the CALR MUTATION assay, confirming the diagnostic sensitivity and specificity (Fig. 49). In addition, simultaneous testing of samples with and without the post-PCR reagent allowed identification of all samples that carried *CALR* mutations, which cause only a slight shift in *CALR* sequence length (red rectangle in Fig. 49), as well as of all 100% mutated samples (all rectangles in Fig. 49).



1) DNA Marker
2) wt control
3) Type-2 control
4) Type-1 control
WT) wt samples
Type-1) Type-1 mutated samples
Type-2) Type-2 mutated samples
Rare) Samples with rare mutations (in order: -4bp, +4bp, -23bp)

Fig. 48: Gel with positive controls (2, 3, 4) and 11 wt samples (WT), 5 samples with Type-1 mutation (Type-1), 5 samples with Type-2 mutation (Type-2) and 3 samples with rare mutations (Rare). Target-specific bands in all samples are detected.



1) DNA Marker

2) wt control

3) Type-2 control

4) Type-1 control

5) NTC

WT) wt samples (with and without post-PCR treatment)

Type-1) Type-1 mutated samples (with and without post-PCR treatment)

Type-2) Type-2 mutated samples (with and without post-PCR treatment)

Rare) Samples with rare mutations (in order: -44bp, -12bp, -32bp, -64bp, -4bp -32bp with and without post-PCR treatment)

Fig. 49: Gels with positive controls (2, 3, 4), 4 wt samples (WT), 8 samples with Type-1 mutation (Type-1), 6 samples with Type-2 mutation (Type-2) and 6 samples with rare mutations (Rare). Target-specific bands in all samples are detected. Non-corresponding results of a sample with and without post-PCR treatment (rectangles) indicate 100% mutated sample. 100% mutated sample with Type-2 mutation (red rectangle) would be difficult to distinguish from wt control (2) without this treatment.

STABILITY STUDY

We started a stability study to determine the shelf-life of the reagents of the CALR MUTATION assay. During this study aliquots of freshly prepared reagents were stored at the recommended storing conditions and were/will be tested at different points in time and the results compared to the data obtained at the starting point (T0). The stability study was divided into two parts. On the one side, the shelf-life of the amplification mix, the controls (wt/Type-1/Type-2 mutation) and the post-PCR reagent were determined. On the other side, the shelf-life of the sequencing primers was assessed. To test the amplification mix, the controls and the post-PCR reagent, we prepared three batches of the amplification mix, each with reagents from different reagent stock solutions. At the same time, we prepared the three positive controls (wt/Type-1/Type-2) and the post-PCR reagent. To test the sequencing primers, we prepared three batches of amplification mixes containing the primers. We divided mixes and controls into aliquots in order to avoid that any reagent undergoes more than two freeze-thaw cycles, which might lead to degradation of the components. The aliquots were stored at -20°C/-30°C. Each batch of amplification mix was tested with the three positive controls. Only for testing the amplification mix, we treated an aliquot of each control with the post-PCR reagent followed by denaturation at 95°C for 5 minutes. At the end, we loaded in the gel both treated and untreated controls. So far, we have tested each batch of amplification mix at T0 (preparation date) and T3 (after 3 months, approximately 90 days). Further data points will be collected at T6, T10, T12 and T14 (months). However, before conclusion of this study we could assign a shelf-life of 6 months on the basis of the similarity of the reagent composition with other IVDs developed at AB ANALITICA, according to the European Standard EN 13640:2002[1]. The assay results at each data point were evaluated for the presence/absence of specific and nonspecific bands/products, band intensity and separation. At data point T3 (3 months) we neither observed any nonspecific bands nor variation in intensity and separation of specific bands for all three batches of the reagents (Fig. 50 Assay reagents). The amplification mixes with the sequencing primers also did not show any sign of degradation over time (Fig. 50 Primers for sequencing).

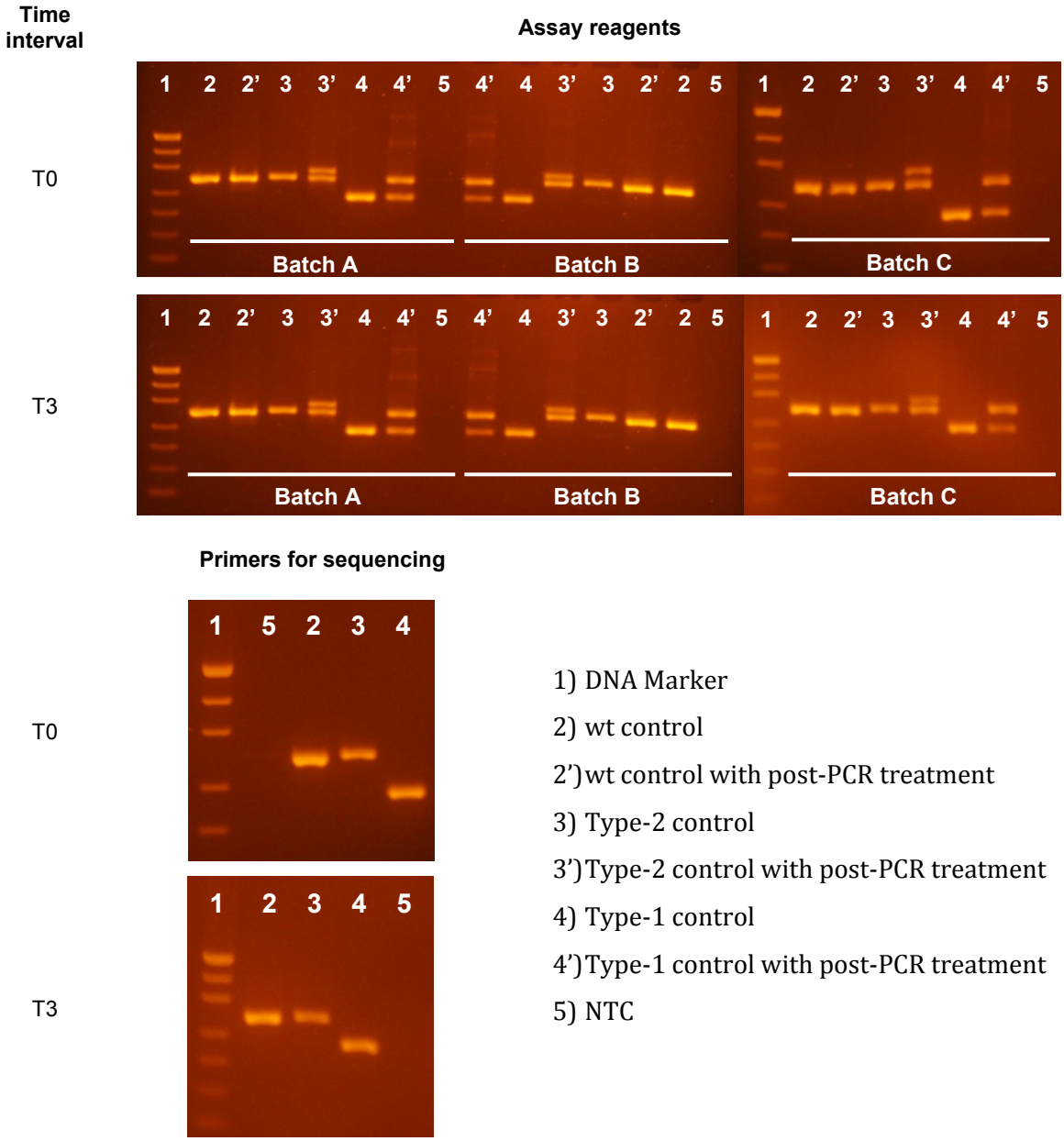


Fig. 50: Gels of T0 and T3 points in the stability study of CALR MUTATION assay. This study is performed with the three positive controls to assess the shelf-life of the sequencing primers. The controls were loaded after/without the post-PCR treatment to assess the shelf-life of the assay reagents.

SANGER SEQUENCING

We designed primers for use in Sanger sequencing of PCR products. Sequencing is required for precise identification of the mutation in a sample. The sequencing primers of the CALR MUTATION assay are specific for the target sequence and showed well defined peaks and little background in the electropherogram. This allows rapid identification of the mutation position in the sequence. Since a mutated DNA sequence overlaps with the wild-type sequence, the location of the mutation site in the sequence may be easily identified (Fig. 51).

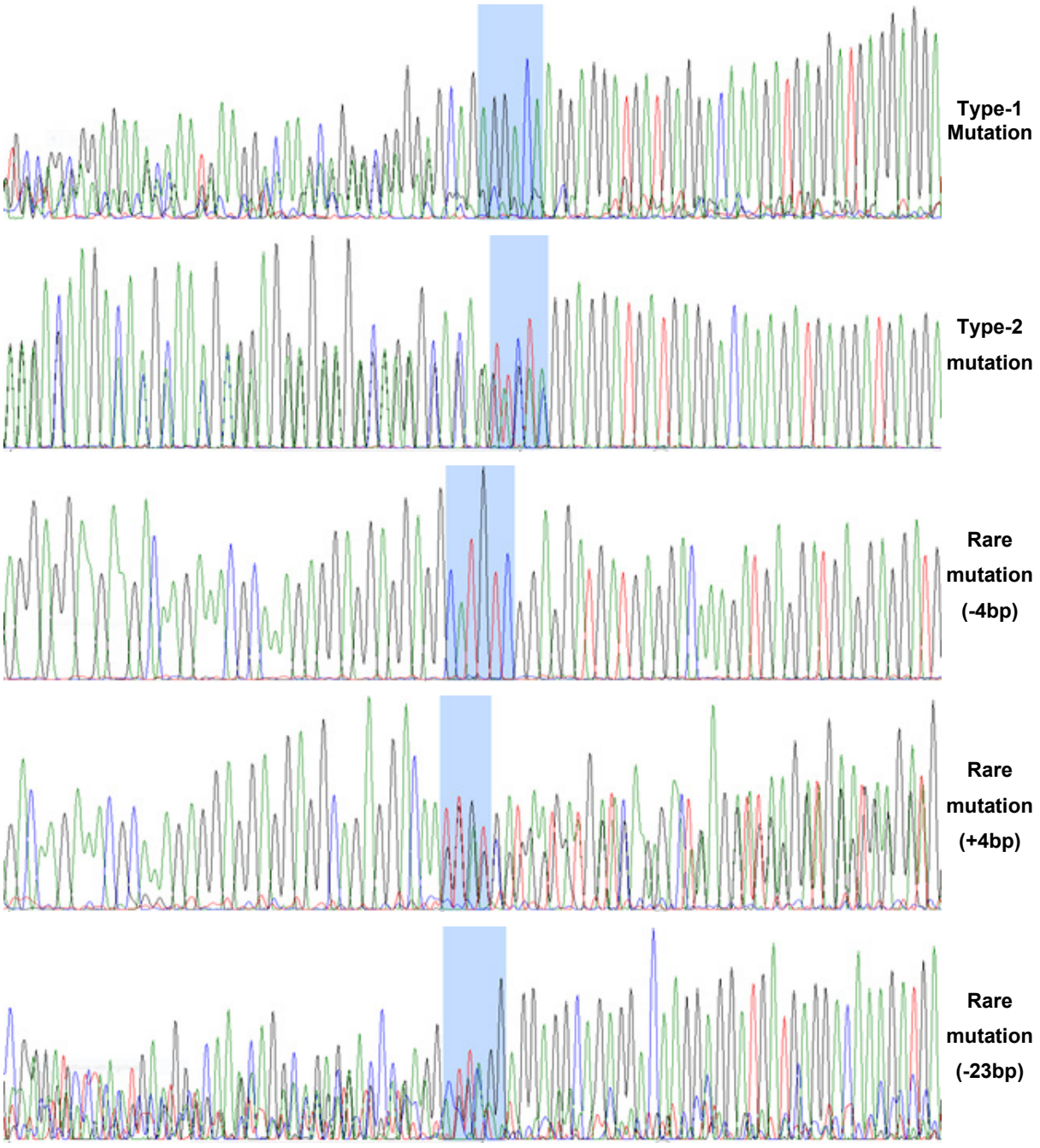


Fig. 51: Electropherograms of samples with Type-1 mutation, Type-2 mutation and three rare mutations. Mutation start is highlighted.

6.3.3 MPL W515L/K ASSAY

ANALYTICAL SPECIFICITY

To ensure a high specificity of the MPL W515L/K assay, highly specific amplification primers were selected and the protocol established using stringent reaction conditions. To test specificity, the primer and probe sequences were aligned in public genome databases (NCBI Blast). The sequences showed no non-specific sequence pairing. To check for possible crosstalk reactions between *MPL* W515K and *MPL* W515L (the respective sequences differing by only two nucleotides) we simultaneously tested each target specific amplification mix with positive controls for the two mutations (5%, 2.5% and 1% of mutated allele). We did not observe any crosstalk between probes for *MPL* W515K and *MPL* W515L (Fig. 52). Analytical specificity of the assay was confirmed by Sanger sequencing of eight PCR products. The tested amplicons exclusively contained the target sequence. The sequencing produced a well-defined electropherogram with sharp peaks separated by even spaces and little background interference at the peak baseline (Fig. 56). This indicates the absence of competing amplification products and a specific amplification of the target sequence.

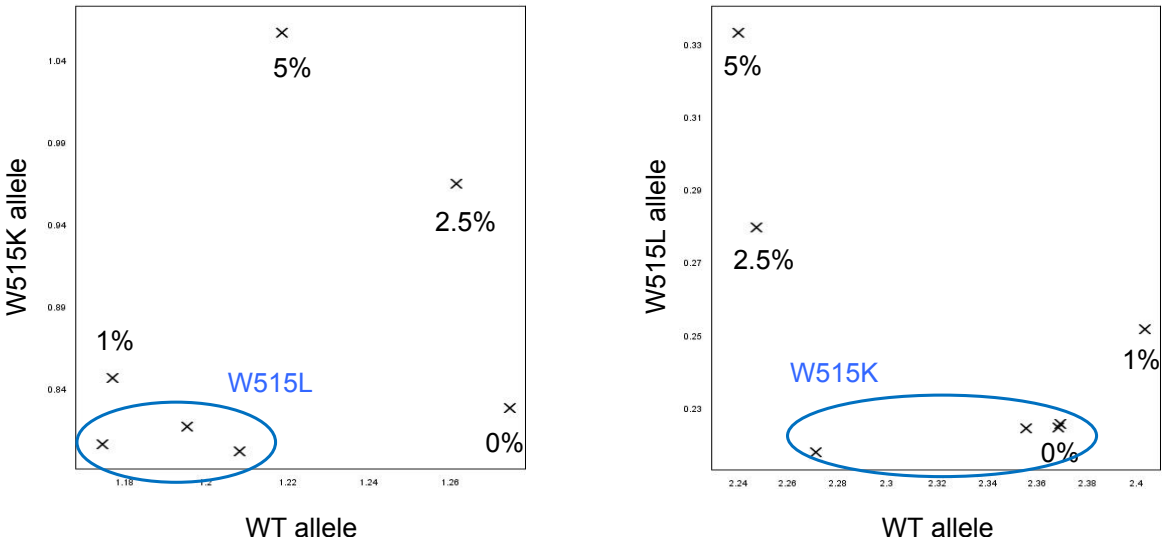


Fig. 52: Allelic discrimination plots of 5%, 2.5%, 1% mutated controls for *MPL* W515K and *MPL* W515L mutations using the mix for amplification of the W515K target (on the left) and the mix for amplification of the W515L target (on the right). Non-specific target reaction controls are detected as wild-type (below 0%, blue circles).

ANALYTICAL SENSITIVITY

We determined the cut-off value for the MPL W515L/K assay comparing the ΔRn ratios of positive controls and clinical samples with wild-type *MPL* (Tab. 19). For both targets, the ΔRn ratios of the tested wild-type samples were below the corresponding 1.5% positive control. Therefore, we assigned a cut-off value of 1.5% (of mutated allele).

	ΔRn ratio W515K/WT	Results W515K	ΔRn ratio W515L/WT	Results W515L
100%	7.587		2.367	
1.5%	0.786		0.134	
1%	0.684		0.103	
0%	0.684		0.092	
Sample 1	0.681	<0%	0.090	<0%
Sample 2	0.676	<0%	0.089	<0%
Sample 3	0.686	<1.5%	0.091	<0%
Sample 4	0.679	<0%	0.090	<0%
Sample 5	0.676	<0%	0.093	<1%
Sample 6	0.699	<1.5%	0.094	<1%
Sample 7	0.690	<1.5%	0.091	<0%
Sample 8	0.718	<1.5%	0.093	<1%
Sample 9	0.683	<0%	0.090	<0%
Sample 10	0.694	<1.5%	0.088	<0%
Sample 11	0.668	<0%	0.088	<0%

Tab. 19: Assessment of the assay cut-off value comparing ΔRn ratios ($\Delta Rn_{MUT}/\Delta Rn_{WT}$) of positive controls and clinical samples with wild-type *MPL*. For both targets, the ΔRn ratios of the tested wild-type samples were below the corresponding 1.5% positive control (cut-off).

RANGE OF DNA QUANTITY

To identify the DNA concentration range, in which optimal results can be obtained with the MPL W515L/K assay, we tested different dilutions of two wild-type samples (5 ng/µL, 10 ng/µL, 16 ng/µL, 20 ng/µL). For the whole concentration range tested (5 to 20 ng/µL corresponding to 25 to 100 ng/rx), the correct mutation status (wild-type) was assigned to the samples (Tab. 20).

	ΔRn ratio W515K/WT	Results W515K	ΔRn ratio W515L/WT	Results W515L
100%	10.035		2.345	
1.5%	0.712		0.110	
0%	0.633		0.092	
Sample 1 (20 ng/µL)	0.658	WT	0.085	WT
Sample 1 (16 ng/µL)	0.645	WT	0.085	WT
Sample 1 (10 ng/µL)	0.640	WT	0.087	WT
Sample 1 (5 ng/µL)	0.648	WT	0.094	WT
Sample 2 (20 ng/µL)	0.669	WT	0.087	WT
Sample 2 (16 ng/µL)	0.665	WT	0.087	WT
Sample 2 (10 ng/µL)	0.669	WT	0.088	WT
Sample 2 (5 ng/µL)	0.648	WT	0.092	WT

Tab. 20: Genotyping of two wild-type known samples at different DNA concentrations (5 ng/µL, 10 ng/µL, 16 ng/µL, 20 ng/µL) with the MPL W515L/K assay.

DIAGNOSTIC SENSITIVITY AND SPECIFICITY

As a preliminary analysis of the diagnostic sensitivity and diagnostic specificity of the MPL W515L/K assay, we analyzed 19 clinical samples (11 with wild-type *MPL* and 8 with mutated *MPL*). All wild-type samples were correctly identified. As expected, these samples showed ΔRn ratios below the assay cut-off (Fig. 53). One W515K and four W515L mutated samples were correctly identified by the assay and the calculated allele concentrations were within or close to the correct range (Fig. 54). Three samples carrying mutations other than or in addition to W515K and W515L were not correctly identified by the assay. These atypical mutations (W515K+c.1542G>A, W515R and W515S) had to be identified by Sanger sequencing (Fig. 56). With the W515K+c.1542G>A mutated sample the assay had shown no amplification signal while the W515R and W515S alleles were detected as wild-type (Fig. 55). This result represents a further confirmation of the high analytical sensitivity of the assay, but, at the same time, indicates a limited diagnostic sensitivity.

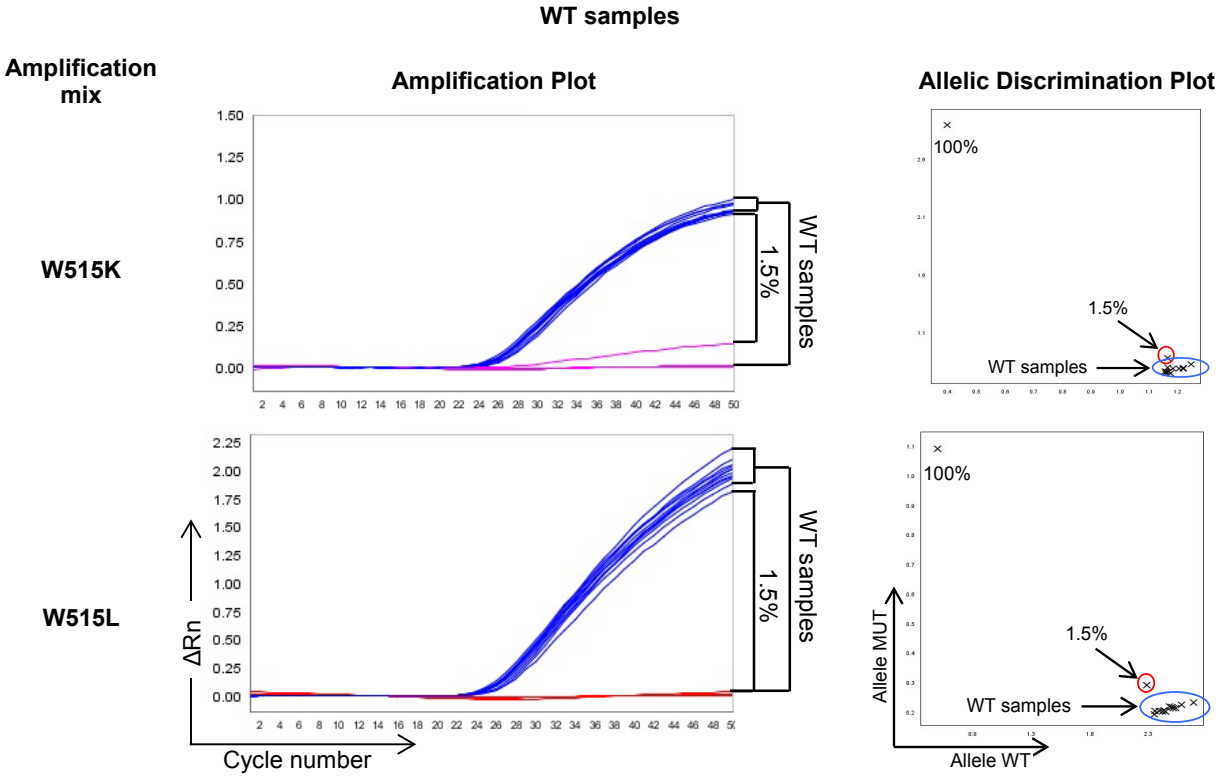


Fig. 53: Amplification and allelic discrimination plots of 11 wild-type samples and the 1.5% positive control (cut-off). ΔRn values of wild-type samples are higher for the wild-type allele and lower for the mutated allele compared to the cut-off (below the cut-off in the scatter plot).

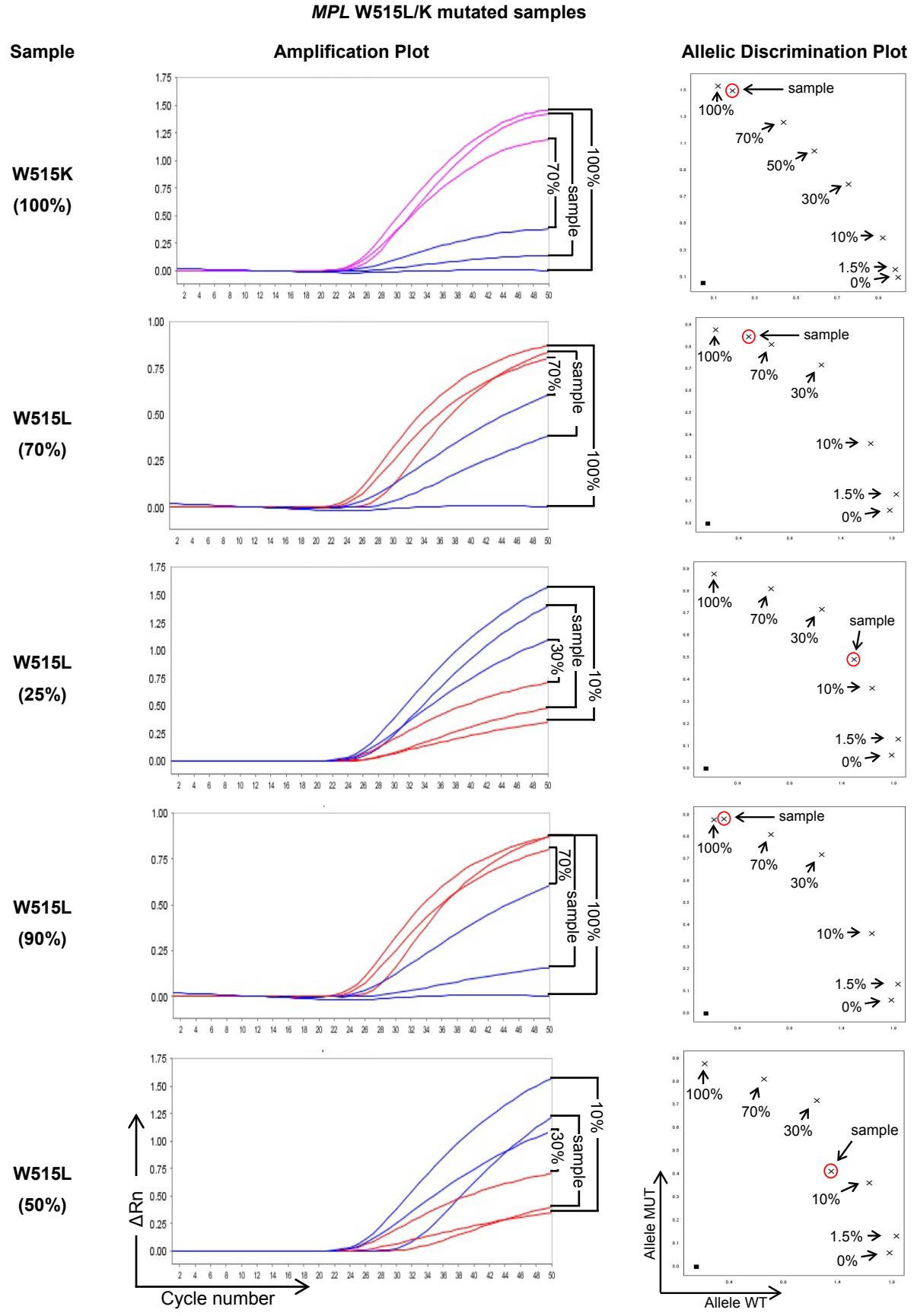


Fig. 54: Amplification and allelic discrimination plots of *MPL* W515L/K mutated samples with known concentration of mutated allele correctly identified by the MPL W515L/K assay. Mutated allele concentration of samples was assigned in/close to the correct concentration range.

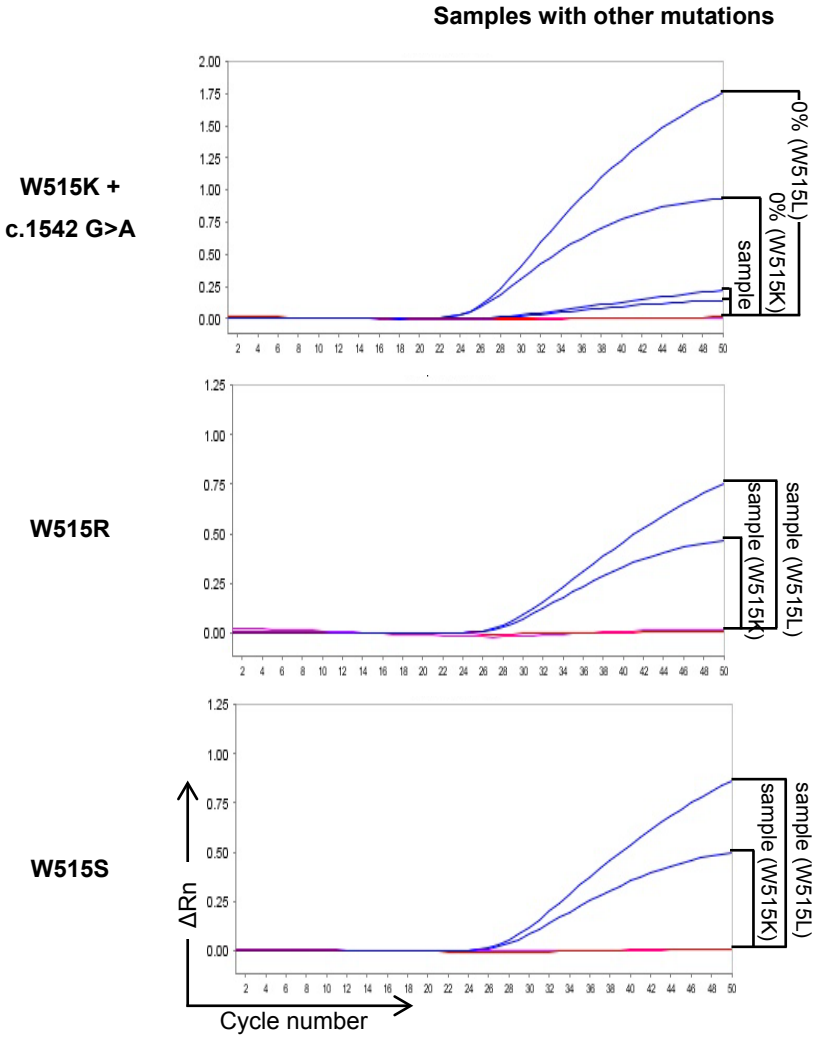


Fig. 55: Amplification of samples with other mutations not correctly identified by the MPL W515L/K assay. The sample with the W515K+c.1542G>A mutations showed no amplification signals, samples with the W515R and W515S mutations were identified as wild-type. The amplification mix is indicated in brackets.

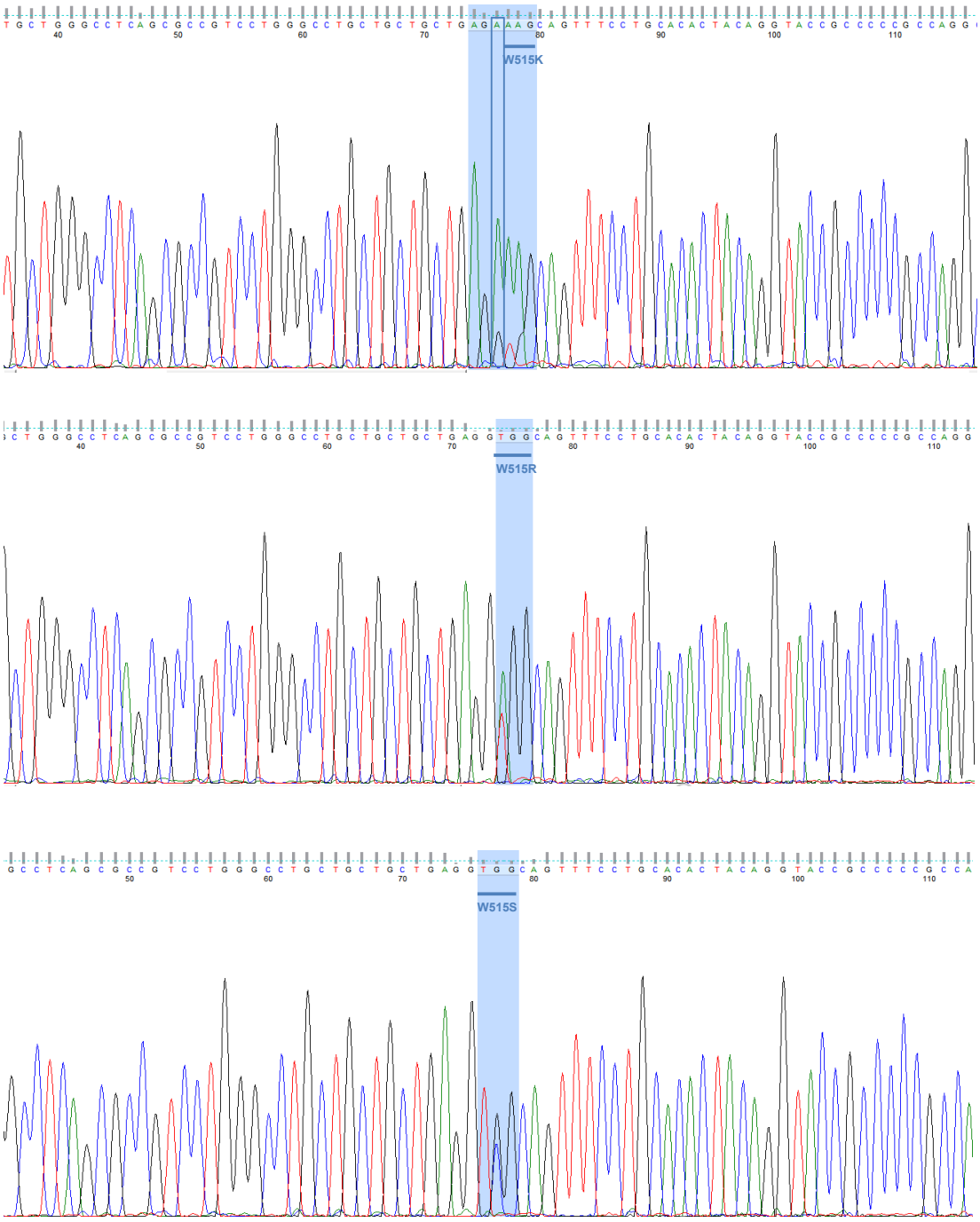


Fig. 56: Electropherograms of the three mutated samples not correctly identified by the MPL W515L/K assay. Two of these sequences contain rare *MPL* W515 mutations (W515R and W515S) and the third has the *MPL* W515K mutation together with a nucleotide substitution in codon 514 (W515K+c.1542G>A).

6.4 CE MARKING

Both IL28B assays and the CALR MUTATION assay fulfilled the “essential requirements" of safety, health, design and manufacture (Annex I of the Directive 98/79/EC[239], transposed by GD 332/2000), which must be met by *in vitro* diagnostic medical devices when they are manufactured and licensed.

For devices other than those covered by Annex II of the Directive, the manufacturer shall, in order to affix the CE marking, draw up the EC declaration of conformity (Annex III of the Directive 98/79/EC[239]) before placing the devices on the market.

As required, we prepared the technical documentation for the conformity of the product with the requirements of the Directive, including:

- general descriptions of the products
- *Kit for detection and genotyping of polymorphism rs12979860 in the human gene coding for Interleukin 28B by real-time PCR*
- *Kit for detection and genotyping of polymorphism rs8099917 in the human gene coding for Interleukin 28B by real-time PCR*
- *Kit for detection of mutations in exon 9 of the CALR gene*;
- documentation of the quality system (requirements and regulations followed by the manufacturer, documented as industrial strategies and procedures are part of the product dossier);
- design information, including the determination of the characteristics of the basic materials, characteristics and limitation of the device performance and manufacturing methods;
- descriptions and explanations necessary to understand the above mentioned characteristics;
- risk analysis;
- performance evaluation data (the reference methods, materials and measurement units used);
- design calculations;
- test reports;
- labels and instructions for use;
- stability studies.

6.4.1 REALQUALITY RS-IL28B rs12979860

DICHIARAZIONE DI CONFORMITÀ • DECLARATION OF CONFORMITY

AB ANALITICA®
ADVANCED BIOMEDICINE
www.abanalitica.it

DICHIARAZIONE DI CONFORMITA' DEL FABBRICANTE
Declaration of conformity of the manufacturer

Fabbricante
Manufacturer

AB ANALITICA
via Svizzera, 16
35127 Padova
(Italy)

CE

Dichiariamo sotto la nostra esclusiva responsabilità che il prodotto al quale questa dichiarazione si riferisce, è conforme alla direttiva 98/79/CE del 27 ottobre 1998 concernente i dispositivi medico diagnostici in vitro.

We hereby declare in our own responsibility, that the product specified below to whom this certificate refers is manufactured according to the regulations of the Directive 98/79/EC of 27 October 1998 on in vitro diagnostics medical device.

Nome commerciale *Brand name*	**REALQUALITY RS-IL28B RS12979860**
Codice *Reference*	**RQ-87-48** **RQ-87-96**
Classificazione CE EC Classification	**Altro tipo di IVD** ***Other type of IVD***
Valutazione della conformità *Conformity assessment route*	**Allegato III** ***Annex III***

Data di notifica al Ministero della Salute: Padova, 28/01/2013
Date of Notification to Italian Ministry of Health: Padua, 28th, January 2013

Data di emissione: 12/11/2014
Date of issue: 12th, November 2014

AB ANALITICA Srl
Via Svizzera, 16 - 35127 PADOVA
Tel 049.761698 - Fax 049.8709510
P. IVA 02375470289

Dott. Vincenzo Putrone
Amministratore Unico
Managing Director and Legal Representative

RS-IL28B_RS12979860_ce_decl_20141112

AB ANALITICA Via Svizzera 16 – 35127 PADOVA (ITALY) - Tel. +39 049 761698; Fax +39 049 8709510
e-mail: info@abanalitica.it; web: www.abanalitica.it

Fig. 57: Declaration of conformity of the manufacturer for the *in vitro* diagnostic device REALQUALITY RS-IL28B rs12979860, product codes RQ-87-48 and RQ-87-96. The declaration of conformity includes information on the manufacturer, data for identification and classification of the device (commercial name) and data concerning CE conformity (Annex III).

Fig. 58: Presentation of the kit REALQUALITY RS-IL28B rs12979860, product code RQ-87-48, comprising BOX RG, BOX PC and instructions for use.
- BOX RG contains: IL28B rs12979860 Real time mix (amplification mix)
- BOX PC contains: PC IL28B C/C (positive control for C/C genotype), PC IL28B C/T (positive control for C/T genotype), PC IL28B T/T (positive control for T/T genotype)

The label on each box includes the following information: intended use, batch number, expiration date, included reagents (with corresponding batch number, expiration date, quantity), storage temperature, number of tests, R/S statements or symbol according to the EN980:2003 (if required), CE mark, IVD symbol, manufacturer information.

6.4.2 REALQUALITY RQ-IL28B rs8099917

DICHIARAZIONE DI CONFORMITÀ • DECLARATION OF CONFORMITY

AB ANALITICA®
ADVANCED BIOMEDICINE
www.abanalitica.it

DICHIARAZIONE DI CONFORMITA' DEL FABBRICANTE
Declaration of conformity of the manufacturer

Fabbricante
Manufacturer

AB ANALITICA
via Svizzera, 16
35127 Padova
(Italy)

CE

Dichiariamo sotto la nostra esclusiva responsabilità che il prodotto al quale questa dichiarazione si riferisce, è conforme alla direttiva 98/79/CE del 27 ottobre 1998 concernente i dispositivi medico diagnostici in vitro.

We hereby declare in our own responsibility, that the product specified below to whom this certificate refers is manufactured according to the regulations of the Directive 98/79/EC of 27 October 1998 on in vitro diagnostics medical device.

Nome commerciale *Brand name*	**REALQUALITY RQ-IL28B rs8099917**
Codice *Reference*	**RQ-91-48** **RQ-91-96**
Classificazione CE EC Classification	**Altro tipo di IVD** ***Other type of IVD***
Valutazione della conformità *Conformity assessment route*	**Allegato III** ***Annex III***

Data di notifica al Ministero della Salute: Padova, 16/01/2014
Date of Notification to Italian Ministry of Health: Padua, 16th, January 2014

Data di emissione: 12/11/2014
Date of issue: 12th, November 2014

AB ANALITICA Srl
Via Svizzera, 16 - 35127 PADOVA
Tel 049.761698 - Fax 049.8709510
P. IVA 02375470289

Dott. Vincenzo Putrone
Amministratore Unico
Managing Director and Legal Representative

RQ-IL28B_rs8099917_ce_decl_20141112

AB ANALITICA Via Svizzera 16 – 35127 PADOVA (ITALY) - Tel. +39 049 761698; Fax +39 049 8709510
e-mail: info@abanalitica.it; web: www.abanalitica.it

Fig. 59: Declaration of conformity of the manufacturer for the *in vitro* diagnostic device REALQUALITY RQ-IL28B rs8099917, product codes RQ-91-48 and RQ-91-96. The declaration of conformity includes information on the manufacturer, data for identification and classification of the device (commercial name) and data concerning CE conformity (Annex III).

Fig. 60: Presentation of the kit REALQUALITY RQ-IL28B rs8099917, product code RQ-91-48, comprising BOX RG, BOX PC and instructions for use.
- BOX RG contains: IL28B rs8099917 Real time mix (amplification mix)
- BOX PC contains: PC IL28B T/T (positive control for T/T genotype), PC IL28B T/G (positive control for T/G genotype), PC IL28B G/G (positive control for G/G genotype)

The label on each box includes the following information: intended use, batch number, expiration date, included reagents (with corresponding batch number, expiration date, quantity), storage temperature, number of tests, R/S statements or symbol according to the EN980:2003 (if required), CE mark, IVD symbol, manufacturer information.

6.4.3 GENEQUALITY CALR MUTATION

DICHIARAZIONE DI CONFORMITA' • DECLARATION OF CONFORMITY

AB ANALITICA®
ADVANCED BIOMEDICINE
www.abanalitica.it

DICHIARAZIONE DI CONFORMITA' DEL FABBRICANTE
Declaration of conformity of the manufacturer

Fabbricante
Manufacturer

AB ANALITICA
via Svizzera, 16
35127 Padova
(Italy)

CE

Dichiariamo sotto la nostra esclusiva responsabilità che il prodotto al quale questa dichiarazione si riferisce, è conforme alla direttiva 98/79/CE del 27 ottobre 1998 concernente i dispositivi medico diagnostici in vitro.

We hereby declare in our own responsibility, that the product specified below to whom this certificate refers is manufactured according to the regulations of the Directive 98/79/EC of 27 October 1998 on in vitro diagnostics medical device.

Nome commerciale *Brand name*	**GENEQUALITY CALR MUTATION**
Codice *Reference*	**04-40A-50** **04-40R-50**
Classificazione CE EC Classification	**Altro tipo di IVD** ***Other type of IVD***
Valutazione della conformità *Conformity assessment route*	**Allegato III** ***Annex III***

Data di notifica al Ministero della Salute: Padova, 22/07/2014
Date of Notification to Italian Ministry of Health: Padua, 22th, July 2014

Data di emissione: 12/11/2014
Date of issue: 12th, November 2014

Dott. Vincenzo Putrone
Amministratore Unico
Managing Director and Legal Representative

GQ-CALR_MUTATION_ce_decl_20141112

AB ANALITICA Via Svizzera 16 – 35127 PADOVA (ITALY) - Tel. +39 049 761698; Fax +39 049 8709510
e-mail: info@abanalitica.it; web: www.abanalitica.it

Fig. 61: Declaration of conformity of the manufacturer for the *in vitro* diagnostic device GENEQUALITY CALR MUTATION, product codes 04-40A-50 and 04-40R-50. The declaration of conformity includes information on the manufacturer, data for identification and classification of the device (commercial name) and data concerning CE conformity (Annex III).

Fig. 62: Presentation of the kit GENEQUALITY CALR MUTATION, product code 04-40A-50, comprising: BOX RG, BOX PC, BOX F, BOX A, instructions for use.

- BOX RG contains: CALR Master Mix (amplification mix), Primer CALR F 10 µM (forward primer for sequencing), Primer CALR R 10 µM (reverse primer for sequencing)
- BOX PC contains: PC CALR WT (wild-type positive control), PC CALR DEL (Type-1 mutation positive control), PC CALR INS (Type-2 mutation positive control), Post-PCR Reagent (reagent for post-PCR treatment)
- BOX F contains: Bromophenol Blue (agarose gel loading buffer), Ethidium bromide (DNA-intercalating agent), MW Marker (Molecular Weight Marker)
- BOX A contains: Agarose, 50X TAE (buffer solution containing Tris base, acetic acid and EDTA, pH=8).

The label on each box includes the following information: intended use, batch number, expiration date, included reagents (with corresponding batch number, expiration date, quantity), storage temperature, number of tests, R/S statements or symbol according to the EN980:2003 (if required), CE mark, IVD symbol, manufacturer information.

7 CONCLUSIONS

The general objective of this doctorate project was to develop commercial tests for personalized medicine in two different contexts: pharmacogenomics and oncogenomics. Pharmacogenomic targets in this study were two single nucleotide polymorphisms near the locus of the *IL28B* gene that are important predictors for the outcome in chronically infected HCV patients treated with the standard therapy (PEG-IFN/RBV). The oncogenomic targets addressed in this study were mutations in the oncogenes *MPL* and *CALR* that are relevant for the diagnosis and prognosis of BCR-ABL1-negative myeloproliferative neoplasms.

During the three years of the doctorate, the studies on IL28B SNPs and CALR mutations were completed with the development of commercial CE-IVD marked kits. Unfortunately, the research on the MPL W515L/K assay could not be completed, because new molecular targets were added as test requirements (all W515 mutations instead of just W515L/K) in order to improve the assay diagnostic sensitivity.

All four assays employed standard methods that have a number of advantages including:

- highly diffused technologies present in the vast majority of clinical labs;
- simple technology that does not require expertise for results interpretations;
- cost reduction by using inexpensive reagents;
- cost reduction by allowing processing of high sample numbers in a single run (up to 96 for *IL28B* polymorphisms and up to 50 for *CALR* mutations)

All these aspects make these tests highly affordable and easy-to-use for the majority of laboratories.

Therefore, the IVD tests developed during this doctorate represent a valuable contribution to the diffusion of personalized medicine in health care.

7.1 REALQUALITY RS-IL28B rs12979860 AND REALQUALITY RQ-IL28B rs8099917

The Hepatitis C virus (HCV) represents a significant health problem worldwide with approximately 170 million infected persons. Over 70% of individuals with an acute HCV infection develop the chronic form (CHC) and are at significant risk of progressive liver disease and subsequent liver cirrhosis and hepatocellular carcinoma (HCC). Until 2011, the standard treatment for chronic HCV infection was weekly administration of pegylated interferon α (PEG-IFN) in combination with daily doses of ribavirin (RBV). However, less than 50% of patients infected with HCV genotype 1 (HCV GT1) undergoing this treatment achieve a sustained virological response (SVR)[29,248]. The likelihood of attaining SVR is influenced by both viral (HCV genotype, baseline viral load) and host factors (age, sex, degree of fibrosis). In 2009, genome-wide association studies (GWAS) uncovered the pivotal role of host genetic factors, i.e. two single nucleotide polymorphisms (SNPs) located near the interleukin-28B (*IL28B*) gene (encoding for IFN-λ3) in the response to PEG-IFN/RBV therapy, especially in patients infected with HCV GT1[11,12,151–153]. Patients carrying the favorable *IL28B* genotypes (CC for rs12979860 and TT for rs8099917) are associated with two-fold higher SVR rates compared to patients with the risk alleles for both SNPs (T and G, respectively). As a result, prediction of the treatment outcome, especially non-responsiveness to PEG-IFN/RBV therapy, has been greatly improved by genotyping for *IL28B* SNPs, enabling personalization of CHC therapy. Newly developed treatments involving direct-acting antivirals (DAAs), like non-structural 3/4A protease inhibitors (PIs), have shown promising results in terms of higher SVR rates in HCV GT1 patients (>70%)[124,127,249] but at the same time have raised concerns on the generation of resistant viral variants and significant side effects[250]. Furthermore, PIs are expensive and not yet available in many countries[251]. Although *IL28B* testing offers just minimal additional information to help clinician's decisions due to the availability of DAAs, pre-treatment evaluation is still recommended to establish the therapeutic schedule[252]. There is also evidence that these SNPs may affect viral kinetics even in the context of IFN-free regimens[169].

As a response to the need for pharmacogenetic tests, which allow genotyping of these informative SNPs in clinical practice, we decided to develop two assays for allelic discrimination based on real-time PCR with TaqMan® probes. This is a standard and cost-effective method that has been in use by most analysis laboratories for several years. The final assays were very reliable in the identification of variants of the *IL28B* SNPs rs12979860 and rs8099917 with a diagnostic sensitivity and specificity of 99.50% and a 99.48%, respectively (calculated on approximately 200 clinical samples). Diagnostic sensitivity and specificity values indicate the accuracy of the test to give reliable results when compared to a reference method (observed/expected). Both assays demonstrated full capability to correctly assign the genotype

of samples with DNA concentrations between 2 ng/rx (analytical sensitivity) and 250 ng/rx (highest DNA quantity to amplify). Both devices were assigned a shelf-life of 12 months, but we planned to extend the time span during which assay performances are monitored to 18 months (already done with REALQUALITY RS-IL28B rs12979860).

REALQUALITY RS-IL28B rs12979860 code RQ-87 and REALQUALITY RQ-IL28B rs8099917 code RQ-91 were notified with the Ministry of Health and received the CE marking. This mandatory conformity marking is required for sale of these products as IVD devices within the European Economic Area (EEA) and guarantees that EU legislation in matter of safety, health and environmental protection is being respected.

The assays must be performed using genomic DNA purified from whole blood. The assays resulted compatible with the following extraction systems: BioRobot EZ1 with *EZ1 DSP Blood* (Qiagen), *QIAamp DNA Blood Mini Kit* (Qiagen), *NucleoSpin® Blood* (Mackerey-Nagel), *chemagic Prepito®-D* (PerkinElmer), *High Pure PCR Template Preparation Kit* (Roche), *MagNA Pure* (Roche), *Maxwell 16 Blood DNA Purification Kit* (Promega), *QuickGene-Mini80 and QuickGene-810* (Fujifilm). The assays were validated on Applied Biosystems StepOne™ and StepOnePlus™ Real-Time PCR Systems, Applied Biosystems 7500 Fast and 7500 Fast Dx Real-Time PCR Systems, Applied Biosystems 7300 Real-Time PCR System, Bio-Rad Dx Real-Time System and Bio-Rad CFX96™ Real-Time PCR Detection System. Compatibility of the assays with additional real-time PCR systems (Applied Biosystems 7500 and 7900 Real-Time PCR Systems and LightCycler® 480 System by Roche) which work with a reaction volume of 25 µL and FAM™ and JOE™ calibrated dyes, was established. As a result, REALQUALITY RS-IL28B rs12979860 and REALQUALITY RQ-IL28B rs8099917 allow genotyping of the two most informative polymorphisms for HCV therapy with most real-time PCR systems. Samples can be simultaneously tested for both polymorphisms seen the identical thermal profile for both assays.

The strengths of these devices compared to other commercial kits include:

- the ready-to-use amplification mix reduces the manual steps thus minimizing the risk of contamination;
- no extra reagents are needed;
- the controls for all three genotypes are ready-to-use and do not need to be diluted;
- the dUTP/UNG system in the amplification mix prevents carry-over contaminations from previous amplifications;
- the master mix contains ROX™ as passive reference for reduction of non-PCR-related fluctuations in fluorescence signal.

7.2 GENEQUALITY CALR MUTATION

BCR-ABL1-negative myeloproliferative neoplasms (MPNs) are characterized by clonal molecular markers. *JAK2* mutations are almost invariably found in polycythemia vera (PV), whereas approximately 90% of patients with essential thrombocytemia (ET) and primary myelofibrosis (PMF) carry one of the mutations listed below in order of relative frequency: the valine-to-phenylalanine substitution in the Janus kinase 2 gene (*JAK2* V617F), insertions and/or deletions in calreticulin encoding gene (*CALR*), the amino acid substitutions tryptophan to leucine or lysine (W515L/K) in the myeloproliferative leukemia virus oncogene (*MPL*). The recent discovery of *CALR* mutations[210,211] was not only important in the field of diagnostics[253], as it is now a major diagnostic criterion in the new proposed WHO classification for BCR-ABL1-negative MPN, but also for its prognostic value[212–214]. In ET and PMF, *CALR* mutations correlated with male gender, younger age, lower hemoglobin levels, lower leukocyte count, higher platelet count and a higher chance of thrombosis-free survival[212,214]. The prognosis in *CALR*-mutated PMF is further influenced by the coexistence (unfavorable) of *ASXL1* mutations[200,219] or other prognostically detrimental mutations[220].

In order to address the new diagnostic needs in the field of BCR-ABL1-negative MPNs, we decided to develop an assay for the detection of mutations in exon 9 of the *CALR* gene. We based this multiplex assay on conventional PCR, a very simple and common method that gives results that are easy to interpret. Thanks to the low costs of this technology (due to inexpensive reagents) the assay can be used for the screening of suspected patients that are negative for the *JAK2* V617F mutation.

Diagnostic specificity and sensitivity of the assay were measured on a total of 36 wild-type and 29 mutated clinical samples. Since the assay method was based on visualization of amplification products on agarose gel, the length of the target sequence was established using a reference method (Sanger sequencing) for all tested samples. All mutated samples contained insertions, deletions or complex ins-del that generated sequences with at least 4 bp difference in length compared to the wild-type sequence. The results showed that the assay had an accuracy of 100%. We were able to overcome the well-known limit in precision associated with analysis by agarose gel electrophoresis by using a post-PCR treatment that allowed detection of DNA sequences of very similar size. The assay detection limit was 10% of mutated allele on a wild-type allele background. The assay was also able to correctly discriminate samples with DNA concentrations from 25 to 100 ng/rx. A stability study to determine the shelf-life of the reagents is still on-going. Preliminary data indicate that the performances of the assay remained stable for at least 3 months. We could, however, safely assign the device a shelf-life of 6 months based on the similarity of the reagent composition to other IVD developed at AB ANALITICA, according to the European Standard EN 13640:2002[1].

The GENEQUALITY CALR MUTATION device, code 04-40R-50 and code 04-40A-50, was finally notified with the Ministry of Health and placed in the market as the first CE-IVD device for the detection of mutations in exon 9 of the *CALR* gene. Code 04-40A-50 contains all reagents necessary for both PCR amplification and gel electrophoresis whereas code 04-40R-50 does not provide the reagents for visualization of PCR products. The reagents for PCR amplification include a ready-to-use master mix, which minimizes the number of manual steps, thus lowering the risk of contaminations, and ready-to-use positive controls for the wild-type sequence and the two most common mutations (Type-1 and Type-2). Both device formats additionally contain sequencing primers. Bidirectional sequencing is optional, but can be very useful when the precise identification of a particular mutation detected by the assay is required. In this regard, development of interpretation software for automated identification of the mutation type based on the DNA sequence would be a valuable addition to the assay.

7.3 MPL W515L/K ASSAY

To complete the panel of the genetic tests facilitating diagnosis of BCR-ABL1-negative MPNs, we decided to develop an assay for the detection and (semi-)quantification of the two main mutations in the *MPL* gene (*MPL* W515L and *MPL* W515K). Based on our experience with the *JAK2* V617F semi-quantitative assay (REALQUALITY RS-JAK-2 V617F), we decided to use a qPCR based on TaqMan® MGB probes. Preliminary results were obtained on the Applied Biosystems StepOnePlus™ Real-Time PCR System. On this instrument we established the cut-off value (threshold between "positive", i.e. above, and "negative", i.e. below, results) of 1.5%. Although the assay was able of correctly attributing the wild-type status to 11/11 wild-type clinical samples (comparison with a reference method), the distance of ΔRn ratios (ΔRn_{MUT}/ ΔRn_{WT}) between the cut-off and the wild-type controls was very small with both amplification mixes. This represents a substantial risk of "false positive" results due to inter-assay variation.

When testing the assay with mutated samples, we found that the presence of other mutations in or close to codon 515 prevented the correct identification of mutated samples. Two samples carrying less frequent mutations in codon 515 of *MPL* (W515R and W515S) produced no signal for the assayed mutations but only a wild-type signal in the expected ΔRn range. A sample with the *MPL* W515K mutation accompanied by a point mutation in codon 514 (c.1542 G>A) generated no fluorescence signal at all. This atypical mutation in codon 514 is a silent mutation or a synonymous SNP (AGG>AGA, both code for Arginine) and, therefore, likely to be present in both alleles of the respective gene. These results may be explained by the high specificity of the probes used in the assays. It seems that hybridization of the probes (specific for the wild-type, *MPL* W515K and *MPL* W515L sequences) to the target sequence did not take place in the presence of the mismatches. Furthermore, this high specificity may be partly due to the stringent amplification conditions used in this assay.

Although detection of *MPL* W515 mutations other than *MPL* W515L and *MPL* W515K was not part of the initial objective of this project, we decided to design a new multiplex assay for the simultaneous semi-quantification of all *MPL* W515 mutations (L, K, R, A, S). For this purpose, the next step of this project will be to accurately design new probes with the introduction of "strategic" mismatches in order to reduce the specificity for the target sequence and favor crosstalk hybrids between similar sequences.

8 PERSPECTIVES

Since the number of genes carrying subclonal mutations used in the stratification of patients with primary myelofibrosis or advanced- polycythemia vera/essential thrombocythemia has rapidly increased in the last months, the need for a complete molecular panel for the detection of all informative mutations is emerging. Next-generation sequencing may be the most suitable technique for this purpose. Therefore, we will be focusing on a further development of a NGS-based system for the detection of mutations in 8 genes (*JAK2, CALR, MPL, ASXL1, EZH2, IDH1, IDH2* and *SRSF2*) that are associated with a high risk according to the DIPSS-plus prognostic model. Patients carrying mutations in these genes should be immediately allocated to therapeutic intervention. Compared to conventional methods used by many laboratories, NGS displays several advantages including high sensitivity and specificity, and allows high-throughput screening of patients, thus potentially lowering the costs for diagnosis.

9 REFERENCES

1. European Committee for Standardization (CEN). Stability testing of in vitro diagnostic reagents - EN13640:2002. (2002).

2. Food and Drug Administration. Paving the Way for Personalized Medicine: FDA's Role in a New Era of Medical Product Development. (2013).

3. National Research Council (US) Committee on A Framework for Developing a New Taxonomy of Disease. Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease. *Natl. Acad. Press* 120 (2011). at <http://www.nap.edu/catalog.php?record_id=13284>

4. Vogel, M. *et al.* Initial presentation of acute hepatitis C virus (HCV) infection among HIV-negative and HIV-positive individuals-experience from 2 large German networks on the study of acute HCV infection. *Clin. Infect. Dis.* **49,** 317–9 (2009).

5. Sutton, A. J. *et al.* A comparison between the force of infection estimates for blood-borne viruses in injecting drug user populations across the European Union: a modelling study. *J. Viral Hepat.* **15,** 809–16 (2008).

6. Ohto, H. *et al.* Transmission of hepatitis C virus from mothers to infants. The Vertical Transmission of Hepatitis C Virus Collaborative Study Group. *N. Engl. J. Med.* **330,** 744–50 (1994).

7. Haley, R. W. & Fischer, R. P. Commercial tattooing as a potentially important source of hepatitis C infection. Clinical epidemiology of 626 consecutive patients unaware of their hepatitis C serologic status. *Medicine (Baltimore).* **80,** 134–51 (2001).

8. Chu, C. M., Yeh, C. T. & Liaw, Y. F. Fulminant hepatic failure in acute hepatitis C: increased risk in chronic carriers of hepatitis B virus. *Gut* **45,** 613–7 (1999).

9. Alter, M. J. *et al.* The prevalence of hepatitis C virus infection in the United States, 1988 through 1994. *N. Engl. J. Med.* **341,** 556–62 (1999).

10. Lauer, G. M. & Walker, B. D. Hepatitis C virus infection. *N. Engl. J. Med.* **345,** 41–52 (2001).

11. Thomas, D. L. *et al.* Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* **461,** 798–801 (2009).

12. Rauch, A. *et al.* Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology* **138,** 1338–45, 1345.e1–7 (2010).

13. Mauss, S., Berg, T., Rockstroh, J., Sarrazin, C. & Wedemeyer, H. *Hepatology 2013: A Clinical Textbook*. (Flying Publisher, 2013).

14. Mathurin, P. *et al.* Slow progression rate of fibrosis in hepatitis C virus patients with persistently normal alanine transaminase activity. *Hepatology* **27,** 868–72 (1998).

15. Zignego, A. L. & Craxì, A. Extrahepatic manifestations of hepatitis C virus infection. *Clin. Liver Dis.* **12,** 611–36, ix (2008).

16. Tong, M. J., El-Farra, N. S., Reikes, A. R. & Co, R. L. Clinical outcomes after transfusion-associated hepatitis C. *N. Engl. J. Med.* **332,** 1463–6 (1995).

17. Poynard, T., Bedossa, P. & Opolon, P. Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet* **349,** 825–32 (1997).

18. Planas, R. *et al.* Natural history of decompensated hepatitis C virus-related cirrhosis. A study of 200 patients. *J. Hepatol.* **40,** 823–30 (2004).

19. Di Bisceglie, A. M. Hepatitis C and hepatocellular carcinoma. *Hepatology* **26,** 34S–38S (1997).

20. Fattovich, G. *et al.* Morbidity and mortality in compensated cirrhosis type C: a retrospective follow-up study of 384 patients. *Gastroenterology* **112,** 463–72 (1997).

21. Svirtlih, N. *et al.* Older age at the time of liver biopsy is the important risk factor for advanced fibrosis in patients with chronic hepatitis C. *Hepatogastroenterology.* **54,** 2324–7 (2007).

22. Vogt, M. *et al.* Prevalence and clinical outcome of hepatitis C infection in children who underwent cardiac surgery before the implementation of blood-donor screening. *N. Engl. J. Med.* **341,** 866–70 (1999).

23. Sterling, R. K. *et al.* A comparison of the spectrum of chronic hepatitis C virus between Caucasians and African Americans. *Clin. Gastroenterol. Hepatol.* **2,** 469–73 (2004).

24. Gitto, S., Micco, L., Conti, F., Andreone, P. & Bernardi, M. Alcohol and viral hepatitis: a mini-review. *Dig. Liver Dis.* **41,** 67–70 (2009).

25. Jonsson, J. R., Purdie, D. M., Clouston, A. D. & Powell, E. E. Recognition of genetic factors influencing the progression of hepatitis C : potential for personalized therapy. *Mol. Diagn. Ther.* **12,** 209–18 (2008).

26. Simmonds, P. Genetic diversity and evolution of hepatitis C virus--15 years on. *J. Gen. Virol.* **85,** 3173–88 (2004).

27. Simmonds, P. *et al.* Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology* **42,** 962–73 (2005).

28. Manns, M. P. *et al.* Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* **358,** 958–65 (2001).

29. Fried, M. W. *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N. Engl. J. Med.* **347,** 975–82 (2002).

30. Bradley, D. *et al.* Hepatitis C virus: buoyant density of the factor VIII-derived isolate in sucrose. *J. Med. Virol.* **34,** 206–8 (1991).

31. Thomssen, R. *et al.* Association of hepatitis C virus in human sera with beta-lipoprotein. *Med. Microbiol. Immunol.* **181,** 293–300 (1992).

32. Thomssen, R., Bonk, S. & Thiele, A. Density heterogeneities of hepatitis C virus in human sera due to the binding of beta-lipoproteins and immunoglobulins. *Med. Microbiol. Immunol.* **182,** 329–34 (1993).

33. Agnello, V., Abel, G., Elfahal, M., Knight, G. B. & Zhang, Q. X. Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 12766–71 (1999).

34. André, P. *et al.* Characterization of low- and very-low-density hepatitis C virus RNA-containing particles. *J. Virol.* **76,** 6919–28 (2002).

35. Nahmias, Y., Casali, M., Barbe, L., Berthiaume, F. & Yarmush, M. L. Liver endothelial cells promote LDL-R expression and the uptake of HCV-like particles in primary rat and human hepatocytes. *Hepatology* **43,** 257–65 (2006).

36. Voisset, C. *et al.* High-density lipoproteins reduce the neutralizing effect of hepatitis C virus (HCV)-infected patient antibodies by promoting HCV entry. *J. Gen. Virol.* **87,** 2577–81 (2006).

37. Heller, T. *et al.* An in vitro model of hepatitis C virion production. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 2579–83 (2005).

38. Wakita, T. *et al.* Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat. Med.* **11,** 791–6 (2005).

39. Yu, X. *et al.* Cryo-electron microscopy and three-dimensional reconstructions of hepatitis C virus particles. *Virology* **367,** 126–34 (2007).

40. Takahashi, K. *et al.* p26 protein and 33-nm particle associated with nucleocapsid of hepatitis C virus recovered from the circulation of infected hosts. *Virology* **191,** 431–4 (1992).

41. Song, Y. *et al.* The hepatitis C virus RNA 3'-untranslated region strongly enhances translation directed by the internal ribosome entry site. *J. Virol.* **80,** 11579–88 (2006).

42. McLauchlan, J., Lemberg, M. K., Hope, G. & Martoglio, B. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *EMBO J.* **21,** 3980–8 (2002).

43. Matsumoto, M., Hwang, S. B., Jeng, K. S., Zhu, N. & Lai, M. M. Homotypic interaction and multimerization of hepatitis C virus core protein. *Virology* **218,** 43–51 (1996).

44. Ait-Goughoulte, M. *et al.* Core protein cleavage by signal peptide peptidase is required for hepatitis C virus-like particle assembly. *J. Gen. Virol.* **87,** 855–60 (2006).

45. Santolini, E., Migliaccio, G. & La Monica, N. Biosynthesis and biochemical properties of the hepatitis C virus core protein. *J. Virol.* **68,** 3631–41 (1994).

46. Duvet, S. *et al.* Glycosylation of the hepatitis C virus envelope protein E1 occurs posttranslationally in a mannosylphosphoryldolichol-deficient CHO mutant cell line. *Glycobiology* **12,** 95–101 (2002).

47. Cocquerel, L., Meunier, J. C., Pillez, A., Wychowski, C. & Dubuisson, J. A retention signal necessary and sufficient for endoplasmic reticulum localization maps to the transmembrane domain of hepatitis C virus glycoprotein E2. *J. Virol.* **72,** 2183–91 (1998).

48. Duvet, S. *et al.* Hepatitis C virus glycoprotein complex localization in the endoplasmic reticulum involves a determinant for retention and not retrieval. *J. Biol. Chem.* **273,** 32088–95 (1998).

49. Cocquerel, L. *et al.* The transmembrane domain of hepatitis C virus glycoprotein E1 is a signal for static retention in the endoplasmic reticulum. *J. Virol.* **73,** 2641–9 (1999).

50. Cocquerel, L. *et al.* Coexpression of hepatitis C virus envelope proteins E1 and E2 in cis improves the stability of membrane insertion of E2. *J. Gen. Virol.* **82,** 1629–35 (2001).

51. Op De Beeck, A. *et al.* The transmembrane domains of hepatitis C virus envelope glycoproteins E1 and E2 play a major role in heterodimerization. *J. Biol. Chem.* **275,** 31428–37 (2000).

52. Weiner, A. J. *et al.* Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. *Virology* **180,** 842–8 (1991).

53. Kato, N. Molecular virology of hepatitis C virus. *Acta Med. Okayama* **55,** 133–59 (2001).

54. Helle, F. *et al.* Role of N-linked glycans in the functions of hepatitis C virus envelope proteins incorporated into infectious virions. *J. Virol.* **84,** 11905–15 (2010).

55. Haqshenas, G., Mackenzie, J. M., Dong, X. & Gowans, E. J. Hepatitis C virus p7 protein is localized in the endoplasmic reticulum when it is encoded by a replication-competent genome. *J. Gen. Virol.* **88,** 134–42 (2007).

56. Pavlović, D. *et al.* The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 6104–8 (2003).

57. Griffin, S. D. C. *et al.* The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS Lett.* **535,** 34–8 (2003).

58. Sakai, A. *et al.* The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 11646–51 (2003).

59. Stapleford, K. A. & Lindenbach, B. D. Hepatitis C virus NS2 coordinates virus particle assembly through physical interactions with the E1-E2 glycoprotein and NS3-NS4A enzyme complexes. *J. Virol.* **85,** 1706–17 (2011).

60. Erdtmann, L. *et al.* The hepatitis C virus NS2 protein is an inhibitor of CIDE-B-induced apoptosis. *J. Biol. Chem.* **278,** 18256–64 (2003).

61. Yang, X.-J. *et al.* HCV NS2 protein inhibits cell proliferation and induces cell cycle arrest in the S-phase in mammalian cells through down-regulation of cyclin A expression. *Virus Res.* **121,** 134–43 (2006).

62. Kaukinen, P., Sillanpää, M., Nousiainen, L., Melén, K. & Julkunen, I. Hepatitis C virus NS2 protease inhibits host cell antiviral response by inhibiting IKKε and TBK1 functions. *J. Med. Virol.* **85,** 71–82 (2013).

63. Jin, L. & Peterson, D. L. Expression, isolation, and characterization of the hepatitis C virus ATPase/RNA helicase. *Arch. Biochem. Biophys.* **323,** 47–53 (1995).

64. Kim, D. W., Gwack, Y., Han, J. H. & Choe, J. C-terminal domain of the hepatitis C virus NS3 protein contains an RNA helicase activity. *Biochem. Biophys. Res. Commun.* **215,** 160–6 (1995).

65. Malcolm, B. A. *et al.* SCH 503034, a mechanism-based inhibitor of hepatitis C virus NS3 protease, suppresses polyprotein maturation and enhances the antiviral activity of alpha interferon in replicon cells. *Antimicrob. Agents Chemother.* **50,** 1013–20 (2006).

66. Perni, R. B. *et al.* Preclinical profile of VX-950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3-4A serine protease. *Antimicrob. Agents Chemother.* **50,** 899–909 (2006).

67. Wölk, B. *et al.* Subcellular localization, stability, and trans-cleavage competence of the hepatitis C virus NS3-NS4A complex expressed in tetracycline-regulated cell lines. *J. Virol.* **74,** 2293–304 (2000).

68. Blight, K. J. Charged residues in hepatitis C virus NS4B are critical for multiple NS4B functions in RNA replication. *J. Virol.* **85,** 8158–71 (2011).

69. Einav, S., Elazar, M., Danieli, T. & Glenn, J. S. A nucleotide binding motif in hepatitis C virus (HCV) NS4B mediates HCV RNA replication. *J. Virol.* **78,** 11288–95 (2004).

70. Einav, S. *et al.* Discovery of a hepatitis C target and its pharmacological inhibitors by microfluidic affinity analysis. *Nat. Biotechnol.* **26,** 1019–27 (2008).

71. Egger, D. *et al.* Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J. Virol.* **76,** 5974–84 (2002).

72. Chatterji, U. *et al.* The isomerase active site of cyclophilin A is critical for hepatitis C virus replication. *J. Biol. Chem.* **284,** 16998–7005 (2009).

73. Enomoto, N. *et al.* Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b. Sensitivity to interferon is conferred by amino acid substitutions in the NS5A region. *J. Clin. Invest.* **96,** 224–30 (1995).

74. Enomoto, N. *et al.* Mutations in the nonstructural protein 5A gene and response to interferon in patients with chronic hepatitis C virus 1b infection. *N. Engl. J. Med.* **334,** 77–81 (1996).

75. Behrens, S. E., Tomei, L. & De Francesco, R. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* **15,** 12–22 (1996).

76. Roussel, J. *et al.* Characterization of the expression of the hepatitis C virus F protein. *J. Gen. Virol.* **84,** 1751–9 (2003).

77. Walewski, J. L., Keller, T. R., Stump, D. D. & Branch, A. D. Evidence for a new hepatitis C virus antigen encoded in an overlapping reading frame. *RNA* **7,** 710–21 (2001).

78. Komurian-Pradel, F. *et al.* Antigenic relevance of F protein in chronic hepatitis C virus infection. *Hepatology* **40,** 900–9 (2004).

79. Germi, R. *et al.* Cellular glycosaminoglycans and low density lipoprotein receptor are involved in hepatitis C virus adsorption. *J. Med. Virol.* **68,** 206–15 (2002).

80. Barth, H. *et al.* Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J. Biol. Chem.* **278,** 41003–12 (2003).

81. Heo, T.-H. *et al.* Incomplete humoral immunity against hepatitis C virus is linked with distinct recognition of putative multiple receptors by E2 envelope glycoprotein. *J. Immunol.* **173,** 446–55 (2004).

82. Monazahian, M. *et al.* Low density lipoprotein receptor as a candidate receptor for hepatitis C virus. *J. Med. Virol.* **57,** 223–9 (1999).

83. Wünschmann, S., Medh, J. D., Klinzmann, D., Schmidt, W. N. & Stapleton, J. T. Characterization of hepatitis C virus (HCV) and HCV E2 interactions with CD81 and the low-density lipoprotein receptor. *J. Virol.* **74,** 10055–62 (2000).

84. Molina, S. *et al.* The low-density lipoprotein receptor plays a role in the infection of primary human hepatocytes by hepatitis C virus. *J. Hepatol.* **46,** 411–9 (2007).

85. Scarselli, E. *et al.* The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO J.* **21,** 5017–25 (2002).

86. Kapadia, S. B., Barth, H., Baumert, T., McKeating, J. A. & Chisari, F. V. Initiation of hepatitis C virus infection is dependent on cholesterol and cooperativity between CD81 and scavenger receptor B type I. *J. Virol.* **81,** 374–83 (2007).

87. Zeisel, M. B. *et al.* Scavenger receptor class B type I is a key host factor for hepatitis C virus infection required for an entry step closely linked to CD81. *Hepatology* **46,** 1722–31 (2007).

88. Evans, M. J. *et al.* Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature* **446,** 801–5 (2007).

89. Zheng, A. *et al.* Claudin-6 and claudin-9 function as additional coreceptors for hepatitis C virus. *J. Virol.* **81,** 12465–71 (2007).

90. Meertens, L. *et al.* The tight junction proteins claudin-1, -6, and -9 are entry cofactors for hepatitis C virus. *J. Virol.* **82,** 3555–60 (2008).

91. Liu, S. *et al.* Tight junction proteins claudin-1 and occludin control hepatitis C virus entry and are downregulated during infection to prevent superinfection. *J. Virol.* **83,** 2011–4 (2009).

92. Ploss, A. *et al.* Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature* **457,** 882–6 (2009).

93. Lupberger, J. *et al.* EGFR and EphA2 are host factors for hepatitis C virus entry and possible targets for antiviral therapy. *Nat. Med.* **17,** 589–95 (2011).

94. Sainz, B. *et al.* Identification of the Niemann-Pick C1-like 1 cholesterol absorption receptor as a new hepatitis C virus entry factor. *Nat. Med.* **18,** 281–5 (2012).

95. Bartosch, B. *et al.* Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *J. Biol. Chem.* **278,** 41624–30 (2003).

96. Hsu, M. *et al.* Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 7271–6 (2003).

97. Blanchard, E. *et al.* Hepatitis C virus entry depends on clathrin-mediated endocytosis. *J. Virol.* **80,** 6964–72 (2006).

98. Codran, A. *et al.* Entry of hepatitis C virus pseudotypes into primary human hepatocytes by clathrin-dependent endocytosis. *J. Gen. Virol.* **87,** 2583–93 (2006).

99. Meertens, L., Bertaux, C. & Dragic, T. Hepatitis C virus entry requires a critical postinternalization step and delivery to early endosomes via clathrin-coated vesicles. *J. Virol.* **80,** 11571–8 (2006).

100. Lavillette, D. *et al.* Characterization of fusion determinants points to the involvement of three discrete regions of both E1 and E2 glycoproteins in the membrane fusion process of hepatitis C virus. *J. Virol.* **81,** 8752–65 (2007).

101. Valli, M. B. *et al.* Transmission in vitro of hepatitis C virus from persistently infected human B-cells to hepatoma cells by cell-to-cell contact. *J. Med. Virol.* **78,** 192–201 (2006).

102. Witteveldt, J. *et al.* CD81 is dispensable for hepatitis C virus cell-to-cell transmission in hepatoma cells. *J. Gen. Virol.* **90,** 48–58 (2009).

103. Fukushi, S. *et al.* Complete 5' noncoding region is necessary for the efficient internal initiation of hepatitis C virus RNA. *Biochem. Biophys. Res. Commun.* **199,** 425–32 (1994).

104. Honda, M., Beard, M. R., Ping, L. H. & Lemon, S. M. A phylogenetically conserved stem-loop structure at the 5' border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *J. Virol.* **73,** 1165–74 (1999).

105. Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. & Nomoto, A. Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.* **66,** 1476–83 (1992).

106. Wang, C., Sarnow, P. & Siddiqui, A. Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. *J. Virol.* **67,** 3338–44 (1993).

107. Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. & Shimotohno, K. Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 5547–51 (1991).

108. Santolini, E., Pacini, L., Fipaldini, C., Migliaccio, G. & Monica, N. The NS2 protein of hepatitis C virus is a transmembrane polypeptide. *J. Virol.* **69,** 7461–71 (1995).

109. Failla, C., Tomei, L. & De Francesco, R. Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *J. Virol.* **68,** 3753–60 (1994).

110. Tanji, Y., Hijikata, M., Satoh, S., Kaneko, T. & Shimotohno, K. Hepatitis C virus-encoded nonstructural protein NS4A has versatile functions in viral protein processing. *J. Virol.* **69,** 1575–81 (1995).

111. Bartenschlager, R., Lohmann, V., Wilkinson, T. & Koch, J. O. Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. *J. Virol.* **69,** 7519–28 (1995).

112. Lin, C., Thomson, J. A. & Rice, C. M. A central region in the hepatitis C virus NS4A protein allows formation of an active NS3-NS4A serine proteinase complex in vivo and in vitro. *J. Virol.* **69,** 4373–80 (1995).

113. Tomei, L., Failla, C., Vitale, R. L., Bianchi, E. & De Francesco, R. A central hydrophobic domain of the hepatitis C virus NS4A protein is necessary and sufficient for the activation of the NS3 protease. *J. Gen. Virol.* **77 (Pt 5),** 1065–70 (1996).

114. Gastaminza, P. *et al.* Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *J. Virol.* **82,** 2120–9 (2008).

115. Moradpour, D., Englert, C., Wakita, T. & Wands, J. R. Characterization of cell lines allowing tightly regulated expression of hepatitis C virus core protein. *Virology* **222,** 51–63 (1996).

116. Barba, G. *et al.* Hepatitis C virus core protein shows a cytoplasmic localization and associates to cellular lipid storage droplets. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 1200–5 (1997).

117. Miyanari, Y. *et al.* The lipid droplet is an important organelle for hepatitis C virus production. *Nat. Cell Biol.* **9,** 1089–97 (2007).

118. Shavinskaya, A., Boulant, S., Penin, F., McLauchlan, J. & Bartenschlager, R. The lipid droplet binding domain of hepatitis C virus core protein is a major determinant for efficient virus assembly. *J. Biol. Chem.* **282,** 37158–69 (2007).

119. Appel, N. *et al.* Essential role of domain III of nonstructural protein 5A for hepatitis C virus infectious particle assembly. *PLoS Pathog.* **4,** e1000035 (2008).

120. Hoofnagle, J. H. *et al.* Treatment of chronic non-A,non-B hepatitis with recombinant human alpha interferon. A preliminary report. *N. Engl. J. Med.* **315,** 1575–8 (1986).

121. McHutchison, J. G. *et al.* Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N. Engl. J. Med.* **339,** 1485–92 (1998).

122. Poynard, T. *et al.* Randomised trial of interferon alpha2b plus ribavirin for 48 weeks or for 24 weeks versus interferon alpha2b plus placebo for 48 weeks for treatment of chronic infection with hepatitis C virus. International Hepatitis Interventional Therapy Group (IHIT). *Lancet* **352,** 1426–32 (1998).

123. McHutchison, J. G. *et al.* Peginterferon alfa-2b or alfa-2a with ribavirin for treatment of hepatitis C infection. *N. Engl. J. Med.* **361,** 580–93 (2009).

124. Jacobson, I. M. *et al.* Telaprevir for previously untreated chronic hepatitis C virus infection. *N. Engl. J. Med.* **364,** 2405–16 (2011).

125. Poordad, F. *et al.* Boceprevir for untreated chronic HCV genotype 1 infection. *N. Engl. J. Med.* **364,** 1195–206 (2011).

126. Bacon, B. R. *et al.* Boceprevir for previously treated chronic HCV genotype 1 infection. *N. Engl. J. Med.* **364,** 1207–17 (2011).

127. Zeuzem, S. *et al.* Telaprevir for retreatment of HCV infection. *N. Engl. J. Med.* **364,** 2417–28 (2011).

128. Sarrazin, C. *et al.* [Expert opinion on boceprevir- and telaprevir-based triple therapies of chronic hepatitis C]. *Z. Gastroenterol.* **50,** 57–72 (2012).

129. Darnell, J. E., Kerr, I. M. & Stark, G. R. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264,** 1415–21 (1994).

130. Sarasin-Filipowicz, M. *et al.* Interferon signaling and treatment outcome in chronic hepatitis C. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 7034–9 (2008).

131. Gale, M. & Foy, E. M. Evasion of intracellular host defence by hepatitis C virus. *Nature* **436,** 939–45 (2005).

132. INCIVEK™ (telaprevir) Full prescribing information. 1–23 (2011).

133. Welsch, C., Jesudian, A., Zeuzem, S. & Jacobson, I. New direct-acting antiviral agents for the treatment of hepatitis C virus infection and perspectives. *Gut* **61 Suppl 1,** i36–46 (2012).

134. VICTRELIS® (boceprevir) Full prescribing information. 1–30 (2014).

135. Bartenschlager, R. & Lohmann, V. Replication of hepatitis C virus. *J. Gen. Virol.* **81,** 1631–48 (2000).

136. Kieffer, T. L. *et al.* Telaprevir and pegylated interferon-alpha-2a inhibit wild-type and resistant genotype 1 hepatitis C virus replication in patients. *Hepatology* **46,** 631–9 (2007).

137. Sarrazin, C. *et al.* Dynamic hepatitis C virus genotypic and phenotypic changes in patients treated with the protease inhibitor telaprevir. *Gastroenterology* **132,** 1767–77 (2007).

138. Bartels, D. J. *et al.* Natural prevalence of hepatitis C virus variants with decreased sensitivity to NS3.4A protease inhibitors in treatment-naive subjects. *J. Infect. Dis.* **198,** 800–7 (2008).

139. Kuntzen, T. *et al.* Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naïve patients. *Hepatology* **48,** 1769–78 (2008).

140. Shimakami, T. *et al.* Protease inhibitor-resistant hepatitis C virus mutants with reduced fitness from impaired production of infectious virus. *Gastroenterology* **140,** 667–75 (2011).

141. Susser, S. *et al.* Characterization of resistance to the protease inhibitor boceprevir in hepatitis C virus-infected patients. *Hepatology* **50,** 1709–18 (2009).

142. Howe, J. a. *et al.* Frequencies of Resistance-Associated Amino Acid Variants Detected By 454-Sequencing During Combination Treatment With Boceprevir Plus Pegintron (Peginterferon Alfa-2B)/Ribavirin in Hcv (Gt1)-Infected Patients. *J. Hepatol.* **54,** S176 (2011).

143. Coelmont, L. *et al.* DEB025 (Alisporivir) inhibits hepatitis C virus replication by preventing a cyclophilin A induced cis-trans isomerisation in domain II of NS5A. *PLoS One* **5,** e13687 (2010).

144. Lanford, R. E. *et al.* Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* **327,** 198–201 (2010).

145. Welsch, C. & Zeuzem, S. Will interferon-free regimens prevail? *Gastroenterology* **142,** 1351–5 (2012).

146. Akuta, N. *et al.* Association of amino acid substitution pattern in core protein of hepatitis C virus genotype 1b high viral load and non-virological response to interferon-ribavirin combination therapy. *Intervirology* **48,** 372–80 (2005).

147. Conjeevaram, H. S. *et al.* Peginterferon and ribavirin treatment in African American and Caucasian American patients with hepatitis C genotype 1. *Gastroenterology* **131,** 470–7 (2006).

148. Jeffers, L. J., Cassidy, W., Howell, C. D., Hu, S. & Reddy, K. R. Peginterferon alfa-2a (40 kd) and ribavirin for black American patients with chronic HCV genotype 1. *Hepatology* **39,** 1702–8 (2004).

149. Muir, A. J., Bornstein, J. D. & Killenberg, P. G. Peginterferon alfa-2b and ribavirin for the treatment of chronic hepatitis C in blacks and non-Hispanic whites. *N. Engl. J. Med.* **350,** 2265–71 (2004).

150. McHutchison, J. G. *et al.* Adherence to combination therapy enhances sustained response in genotype-1-infected patients with chronic hepatitis C. *Gastroenterology* **123,** 1061–9 (2002).

151. Ge, D. *et al.* Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* **461,** 399–401 (2009).

152. Suppiah, V. *et al.* IL28B is associated with response to chronic hepatitis C interferon-alpha and ribavirin therapy. *Nat. Genet.* **41,** 1100–4 (2009).

153. Tanaka, Y. *et al.* Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat. Genet.* **41,** 1105–9 (2009).

154. Shiffman, M. L. *et al.* Treatment of chronic hepatitis C virus in African Americans with genotypes 2 and 3. *Am. J. Gastroenterol.* **102,** 761–6 (2007).

155. Thompson, A. J. *et al.* Interleukin-28B polymorphism improves viral kinetics and is the strongest pretreatment predictor of sustained virologic response in genotype 1 hepatitis C virus. *Gastroenterology* **139,** 120–9.e18 (2010).

156. Rallón, N. I. *et al.* Association of a single nucleotide polymorphism near the interleukin-28B gene with response to hepatitis C therapy in HIV/hepatitis C virus-coinfected patients. *AIDS* **24,** F23–9 (2010).

157. Fukuhara, T. *et al.* Variants in IL28B in liver recipients and donors correlate with response to peg-interferon and ribavirin therapy for recurrent hepatitis C. *Gastroenterology* **139,** 1577–85, 1585.e1–3 (2010).

158. Charlton, M. R. *et al.* Interleukin-28B polymorphisms are associated with histological recurrence and treatment response following liver transplantation in patients with hepatitis C virus infection. *Hepatology* **53,** 317–24 (2011).

159. Mangia, A. *et al.* An IL28B polymorphism determines treatment response of hepatitis C virus genotype 2 or 3 patients who do not achieve a rapid virologic response. *Gastroenterology* **139,** 821–7, 827.e1 (2010).

160. Sakamoto, N. *et al.* Association of IL28B variants with response to pegylated-interferon alpha plus ribavirin combination therapy reveals intersubgenotypic differences between genotypes 2a and 2b. *J. Med. Virol.* **83,** 871–8 (2011).

161. Yu, M.-L. *et al.* Role of interleukin-28B polymorphisms in the treatment of hepatitis C virus genotype 2 infection in Asian patients. *Hepatology* **53,** 7–13 (2011).

162. Moghaddam, A. *et al.* IL28B genetic variation and treatment response in patients with hepatitis C virus genotype 3 infection. *Hepatology* **53,** 746–54 (2011).

163. Antaki, N. *et al.* IL28B polymorphisms do not predict response to therapy in chronic hepatitis C with HCV genotype 5. *Gut* **61,** 1640–1 (2012).

164. Seto, W.-K. *et al.* Role of IL28B and inosine triphosphatase polymorphisms in the treatment of chronic hepatitis C virus genotype 6 infection. *J. Viral Hepat.* **20,** 470–7 (2013).

165. European Medicines Agency (EMEA) Committee for Medicinal Products for Human Use (CHMP). Guideline on clinical evaluation of medicinal products for the treatment of chronic hepatitis C . 1–23 (2011).

166. Grebely, J. *et al.* Potential role for interleukin-28B genotype in treatment decision-making in recent hepatitis C virus infection. *Hepatology* **52,** 1216–24 (2010).

167. Grebely, J., Matthews, G. V & Dore, G. J. Treatment of acute HCV infection. *Nat. Rev. Gastroenterol. Hepatol.* **8,** 265–74 (2011).

168. Mangia, A. *et al.* Treatment optimization and prediction of HCV clearance in patients with acute HCV infection. *J. Hepatol.* **59,** 221–8 (2013).

169. Matsuura, K., Watanabe, T. & Tanaka, Y. Role of IL28B for chronic hepatitis C treatment toward personalized medicine. *J. Gastroenterol. Hepatol.* **29,** 241–9 (2014).

170. Bota, S. *et al.* Role of interleukin-28B polymorphism as a predictor of sustained virological response in patients with chronic hepatitis C treated with triple therapy: a systematic review and meta-analysis. *Clin. Drug Investig.* **33,** 325–31 (2013).

171. Lok, A. S. *et al.* Preliminary study of two antiviral agents for hepatitis C genotype 1. *N. Engl. J. Med.* **366,** 216–24 (2012).

172. Chu, T. W. *et al.* Effect of IL28B genotype on early viral kinetics during interferon-free treatment of patients with chronic hepatitis C. *Gastroenterology* **142,** 790–5 (2012).

173. Zeuzem, S. *et al.* Faldaprevir and deleobuvir for HCV genotype 1 infection. *N. Engl. J. Med.* **369,** 630–9 (2013).

174. Marcello, T. *et al.* Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation kinetics. *Gastroenterology* **131,** 1887–98 (2006).

175. Robek, M. D., Boyd, B. S. & Chisari, F. V. Lambda interferon inhibits hepatitis B and C virus replication. *J. Virol.* **79,** 3851–4 (2005).

176. Muir, A. J. *et al.* Phase 1b study of pegylated interferon lambda 1 with or without ribavirin in patients with chronic genotype 1 hepatitis C virus infection. *Hepatology* **52,** 822–32 (2010).

177. Kempuraj, D. *et al.* Interleukin-28 and 29 (IL-28 and IL-29): new cytokines with anti-viral activities. *Int. J. Immunopathol. Pharmacol.* **17,** 103–6 (2004).

178. Zhou, Z. *et al.* Type III interferon (IFN) induces a type I IFN-like response in a restricted subset of cells through signaling pathways involving both the Jak-STAT pathway and the mitogen-activated protein kinases. *J. Virol.* **81,** 7749–58 (2007).

179. Asahina, Y. *et al.* Potential relevance of cytoplasmic viral sensors and related regulators involving innate immunity in antiviral response. *Gastroenterology* **134,** 1396–405 (2008).

180. Feld, J. J. *et al.* Hepatic gene expression during treatment with peginterferon and ribavirin: Identifying molecular pathways for treatment response. *Hepatology* **46,** 1548–63 (2007).

181. Honda, M. *et al.* Hepatic ISG expression is associated with genetic variation in interleukin 28B and the outcome of IFN therapy for chronic hepatitis C. *Gastroenterology* **139,** 499–509 (2010).

182. Urban, T. J. *et al.* IL28B genotype is associated with differential expression of intrahepatic interferon-stimulated genes in patients with chronic hepatitis C. *Hepatology* **52,** 1888–96 (2010).

183. Thomas, E. *et al.* HCV infection induces a unique hepatic innate immune response associated with robust production of type III interferons. *Gastroenterology* **142,** 978–88 (2012).

184. Watanabe, T. *et al.* Hepatitis C virus kinetics by administration of pegylated interferon-α in human and chimeric mice carrying human hepatocytes with variants of the IL28B gene. *Gut* **62,** 1340–6 (2013).

185. Zhang, S., Kodys, K., Li, K. & Szabo, G. Human type 2 myeloid dendritic cells produce interferon-λ and amplify interferon-α in response to hepatitis C virus infection. *Gastroenterology* **144,** 414–425.e7 (2013).

186. Prokunina-Olsson, L. *et al.* A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nat. Genet.* **45,** 164–71 (2013).

187. Bibert, S. *et al.* IL28B expression depends on a novel TT/-G polymorphism which improves HCV clearance prediction. *J. Exp. Med.* **210,** 1109–16 (2013).

188. Tefferi, a, Thiele, J., Vannucchi, a M. & Barbui, T. An overview on CALR and CSF3R mutations and a proposal for revision of WHO diagnostic criteria for myeloproliferative neoplasms. *Leukemia* **28,** 1407–13 (2014).

189. Cazzola, M. & Kralovics, R. From Janus kinase 2 to calreticulin: the clinically relevant genomic landscape of myeloproliferative neoplasms. *Blood* **123,** 3714–9 (2014).

190. Hartwig, J. & Italiano, J. The birth of the platelet. *J. Thromb. Haemost.* **1,** 1580–6 (2003).

191. Patel, S. R., Hartwig, J. H. & Italiano, J. E. The biogenesis of platelets from megakaryocyte proplatelets. *J. Clin. Invest.* **115,** 3348–54 (2005).

192. Schulze, H. & Shivdasani, R. A. Mechanisms of thrombopoiesis. *J. Thromb. Haemost.* **3,** 1717–24 (2005).

193. Italiano, J. E. & Shivdasani, R. A. Megakaryocytes and beyond: the birth of platelets. *J. Thromb. Haemost.* **1,** 1174–82 (2003).

194. Malara, A. *et al.* Megakaryocytes contribute to the bone marrow-matrix environment by expressing fibronectin, type IV collagen, and laminin. *Stem Cells* **32,** 926–37 (2014).

195. Balduini, A. *et al.* In vitro megakaryocyte differentiation and proplatelet formation in Ph-negative classical myeloproliferative neoplasms: distinct patterns in the different clinical phenotypes. *PLoS One* **6,** e21015 (2011).

196. Swerdlow, S. H. *et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. (International Agency for Research on Cancer (IARC), 2008).

197. Klampfl, T. *et al.* Genome integrity of myeloproliferative neoplasms in chronic phase and during disease progression. *Blood* **118,** 167–76 (2011).

198. Kralovics, R. Genetic complexity of myeloproliferative neoplasms. *Leukemia* **22,** 1841–8 (2008).

199. Harutyunyan, A., Klampfl, T., Cazzola, M. & Kralovics, R. p53 lesions in leukemic transformation. *N. Engl. J. Med.* **364,** 488–90 (2011).

200. Vannucchi, a M. *et al.* Mutations and prognosis in primary myelofibrosis. *Leukemia* **27,** 1861–9 (2013).

201. James, C. *et al.* A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* **434,** 1144–8 (2005).

202. Kralovics, R. *et al.* A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N. Engl. J. Med.* **352,** 1779–90 (2005).

203. Baxter, E. J. *et al.* Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* **365,** 1054–61 (2005).

204. Campbell, P. J. & Green, A. R. The myeloproliferative disorders. *N. Engl. J. Med.* **355,** 2452–66 (2006).

205. Scott, L. M. *et al.* JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *N. Engl. J. Med.* **356,** 459–68 (2007).

206. Passamonti, F. *et al.* Molecular and clinical features of the myeloproliferative neoplasm associated with JAK2 exon 12 mutations. *Blood* **117,** 2813–6 (2011).

207. Pikman, Y. *et al.* MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med.* **3,** e270 (2006).

208. Pardanani, A. D. *et al.* MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood* **108,** 3472–6 (2006).

209. Schmidt, A. E. & Oh, S. T. Pathology consultation on myeloproliferative neoplasms. *Am. J. Clin. Pathol.* **138,** 12–9 (2012).

210. Klampfl, T. *et al.* Somatic mutations of calreticulin in myeloproliferative neoplasms. *N. Engl. J. Med.* **369,** 2379–90 (2013).

211. Nangalia, J. *et al.* Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *N. Engl. J. Med.* **369,** 2391–405 (2013).

212. Rotunno, G. *et al.* Impact of calreticulin mutations on clinical and hematological phenotype and outcome in essential thrombocythemia. *Blood* **123,** 1552–5 (2014).

213. Tefferi, a *et al.* CALR vs JAK2 vs MPL-mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. *Leukemia* **28,** 1472–7 (2014).

214. Rumi, E. *et al.* JAK2 or CALR mutation status defines subtypes of essential thrombocythemia with substantially different clinical course and outcomes. *Blood* **123,** 1544–51 (2014).

215. Cervantes, F. *et al.* New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. *Blood* **113,** 2895–901 (2009).

216. Passamonti, F. *et al.* A dynamic prognostic model to predict survival in primary myelofibrosis: a study by the IWG-MRT (International Working Group for Myeloproliferative Neoplasms Research and Treatment). *Blood* **115,** 1703–8 (2010).

217. Gangat, N. *et al.* DIPSS plus: a refined Dynamic International Prognostic Scoring System for primary myelofibrosis that incorporates prognostic information from karyotype, platelet count, and transfusion status. *J. Clin. Oncol.* **29,** 392–7 (2011).

218. Tefferi, A. Primary myelofibrosis: 2014 update on diagnosis, risk-stratification, and management. *Am. J. Hematol.* **89,** 915–25 (2014).

219. Tefferi, A. *et al.* CALR and ASXL1 mutations-based molecular prognostication in primary myelofibrosis: an international study of 570 patients. *Leukemia* **28,** 1494–500 (2014).

220. Guglielmelli, P. *et al.* The number of prognostically detrimental mutations and prognosis in primary myelofibrosis: an international study of 797 patients. *Leukemia* **28,** 1804–10 (2014).

221. Ballen, K. K. *et al.* Outcome of transplantation for myelofibrosis. *Biol. Blood Marrow Transplant.* **16,** 358–67 (2010).

222. Tefferi, A. JAK inhibitors for myeloproliferative neoplasms: clarifying facts from myths. *Blood* **119,** 2721–30 (2012).

223. Tefferi, A. Challenges facing JAK inhibitor therapy for myeloproliferative neoplasms. *N. Engl. J. Med.* **366,** 844–6 (2012).

224. Tefferi, A. *et al.* Phase 3 Study Of Pomalidomide In Myeloproliferative Neoplasm (MPN)-Associated Myelofibrosis With RBC-Transfusion-Dependence. *Blood* **122,** 394 (2013).

225. Verstovsek, S. *et al.* Safety and efficacy of INCB018424, a JAK1 and JAK2 inhibitor, in myelofibrosis. *N. Engl. J. Med.* **363,** 1117–27 (2010).

226. Tefferi, A. & Pardanani, A. Serious adverse events during ruxolitinib treatment discontinuation in patients with myelofibrosis. *Mayo Clin. Proc.* **86,** 1188–91 (2011).

227. Pardanani, A. *et al.* Safety and efficacy of TG101348, a selective JAK2 inhibitor, in myelofibrosis. *J. Clin. Oncol.* **29,** 789–96 (2011).

228. Pardanani, A. *et al.* Results Of a Randomized, Double-Blind, Placebo-Controlled Phase III Study (JAKARTA) Of The JAK2-Selective Inhibitor Fedratinib (SAR302503) In Patients With Myelofibrosis (MF). *Blood* **122,** 393 (2013).

229. Pardanani, A. *et al.* Safety and efficacy of CYT387, a JAK1 and JAK2 inhibitor, in myelofibrosis. *Leukemia* **27,** 1322–7 (2013).

230. Komrokji, R. S. *et al.* Results of a Phase 2 Study of Pacritinib (SB1518), a Novel Oral JAK2 Inhibitor, In Patients with Primary, Post-Polycythemia Vera, and Post-Essential Thrombocythemia Myelofibrosis. *ASH Annu. Meet. Abstr.* **118,** 282 (2011).

231. Guglielmelli, P. *et al.* Safety and efficacy of everolimus, a mTOR inhibitor, as single agent in a phase 1/2 study in patients with myelofibrosis. *Blood* **118,** 2069–76 (2011).

232. Bio-Rad. PCR Primer and Probe Chemistries. at <http://www.bio-rad.com/it-it/applications-technologies/pcr-primer-probe-chemistries>

233. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. 1977. *Biotechnology* **24,** 104–8 (1992).

234. Life Technologies. Sanger Sequencing Method. at <http://www.lifetechnologies.com/it/en/home/life-science/sequencing/sanger-sequencing/sanger_sequencing_method.html>

235. Grada, A. & Weinbrecht, K. Next-generation sequencing: methodology and application. *J. Invest. Dermatol.* **133,** e11 (2013).

236. Berglund, E. C., Kiialainen, A. & Syvänen, A.-C. Next-generation sequencing technologies and applications for human genetic history and forensics. *Investig. Genet.* **2,** 23 (2011).

237. Mardis, E. R. Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* **9,** 387–402 (2008).

238. Gogol-Döring, A. & Chen, W. An overview of the analysis of next generation sequencing data. *Methods Mol. Biol.* **802,** 249–57 (2012).

239. The European Parliament and the Council of the European Union. DIRECTIVE 98/79/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 October 1998 on in vitro diagnostic medical devices. *Off. J. Eur. Communities* (1998).

240. Grosse, S. D. & Khoury, M. J. What is the clinical utility of genetic testing? *Genet. Med.* **8,** 448–50 (2006).

241. Lucigen. pGC Blue Cloning & Amplification Kits.

242. Invitrogen. TOPO TA Cloning ®. (2006).

243. Life Technologies. Three Simple Steps Save Time. at <http://www.b2b.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cloning/PCR-cloning/Three-Simple-Steps-Save-Time.html>

244. Shuman, S. Recombination mediated by vaccinia virus DNA topoisomerase I in Escherichia coli is sequence specific. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 10104–8 (1991).

245. Shuman, S. Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA topoisomerase. *J. Biol. Chem.* **269,** 32678–84 (1994).

246. Applied Biosystems. Allelic Discrimination Getting Started Guide. (2010).

247. Poodt, J., Fijnheer, R., Walsh, I. B. B. & Hermans, M. H. A. A sensitive and reliable semi-quantitative real-time PCR assay to detect JAK2 V617F in blood. *Hematol. Oncol.* **24,** 227–33 (2006).

248. Hadziyannis, S. J. *et al.* Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann. Intern. Med.* **140,** 346–55 (2004).

249. Poordad, F. *et al.* Factors that predict response of patients with hepatitis C virus infection to boceprevir. *Gastroenterology* **143,** 608–18.e1–5 (2012).

250. Maasoumy, B. *et al.* Eligibility and safety of triple therapy for hepatitis C: lessons learned from the first experience in a real world setting. *PLoS One* **8,** e55285 (2013).

251. Gellad, Z. F., Reed, S. D. & Muir, A. J. Economic evaluation of direct-acting antiviral therapy in chronic hepatitis C. *Antivir. Ther.* **17,** 1189–99 (2012).

252. Rosso, C. *et al.* IL28B polymorphism genotyping as predictor of rapid virologic response during interferon plus ribavirin treatment in hepatitis C virus genotype 1 patients. *World J. Gastroenterol.* **20,** 13146–52 (2014).

253. Tefferi, A. & Pardanani, A. Genetics: CALR mutations and a new diagnostic algorithm for MPN. *Nat. Rev. Clin. Oncol.* **11,** 125–6 (2014).