UNIVERSITA' DEGLI STUDI DI TRIESTE

XXVIII CICLO DEL DOTTORATO DI RICERCA
IN SCIENZE DELLA RIPRODUZIONE E DELLO SVILUPPO

indirizzo Genetico Molecolare

DEVELOPMENT OF FUNCTIONAL ASSAYS TO DETERMINE THE PATHOGENIC VARIANTS OBTAINED BY NEXT GENERATION SEQUENCING IN IHNERITED THROMBOCYTOPENIAS

Sviluppo di saggi funzionali per verificare la patogenicità delle varianti ottenute mediante tecniche di sequenziamento di nuova generazione in geni implicati nelle piastrinopenie ereditarie

Settore scientifico-disciplinare: MED/03 Genetica Medica

Dottoranda

Michela Faleschini

Coordinatore

Prof.ssa Giuliana Decorti

Supervisore di tesi

Prof.ssa Anna Savoia

ANNO ACCADEMICO 2014/2015

ABSTRACT

Inherited thrombocytopenias (IT) are a heterogeneous group of diseases characterized by platelet count lower than $150x10^9$/L. They are clinically and genetically heterogeneous diseases, with mutations in at least 30 causative. However, these genes account for approximately 50% of the IT patients, suggesting that novel forms are still to be characterized.

For this reason, in collaboration with Medical Genetic Unit of Policlinico Sant'Orsola Malpighi in Bologna and the department of General Medicine 3 , IRCCS San Matteo in Pavia, our laboratory is applying the Next Generation Sequencing (NGS) technologies to identify mutations and genes responsible for the disease in thombocytopenic families. However, understanding the effects of the thousands variants identified remains a major problem in diseases like ITs, which are mainly autosomal dominant diseases caused by private, often missense, mutations. Whereas the deleterious effect of nonsense or frameshift variants is clear, that of the amino acid substitutions, which are classified as variants of unknown significance (VUS) is not always obvious. Therefore, it is fundamental to develop assays to tackle the pathogenicity question with functional studies.

My PhD work fits into this project and is focused on the development of functional assays to test the pathogenicity of the variants of the IT genes encoding for cytosckeleton components, such as *ACTN1* and *TUBB1*.

Thanks to Whole Exome Sequencing (WES) in a family suffering from an IT of unknown origin, we identified a single-base substitution located within exon 19 of the *ACTN1* gene (c.2305G>A), which was predicted to result in a missense mutation, p.Glu769Lys. In order to identify *ACTN1* mutations in our cohort of patients, we searched for *ACTN1* mutations in 127 probands affected and identified nine different heterozygous missense variants in ten families. Three were known amino acid substitutions (p.Gly225Lys, p.Arg738Trp and p.Arg752Gln) that were previously reported (Kunishima, al. 2013). The other six were not present in either the dbSNP or the 1000 Genomes Project datasets and therefore regarded as novel variants of *ACTN1*.

For each of these novel variants, we evaluated their potential effect on protein function using three different pathogenicity prediction tools: SIFT, PolyPhen2 and Mutation Taster. We also performed segregation analysis which was confirmed in all but one (p.Asp666Val) cases.

To confirm the genetic analysis, we performed an immunofluorescence analysis of human fibroblasts overexpressing wild-type and mutated forms of *ACTN1*. Confocal microscopy analysis revealed a well organized cytoskeleton in which α-actinin1 colocalizes with actin along the filaments in cells transfected with the wild type construct. On the contrary, all but one (p.Asp666Val) mutant constructs presented an abnormal distribution of actin, which was no longer able to form filaments causing an apparent disruption in the cytoskeletal structure. Determination of the pathogenicity of variants identified through screening of mutation has allowed clinicians to define ACTN1-Related Thrombocytopenia as a mild form of thrombocytopenia without platelet dysfunction. These data have been reported in a paper published on Blood (Bottega et al, 2015).
In the last year, NGS analysis detected *ACTN1* variants in another 6 families. For the 3 novel amino acid substitutions identified we performed immunofluorescence analysis. The deleterious effects on cytoskeleton organization were confirmed only for 2 mutations. NGS analysis identified also potential pathogenetic variants in *TUBB1*, another cytoskeleton component whose variants could be tested functionally using the same strategy as for those affecting *ACTN1*. No family members are available for segregation study. However, we generated constructs to perform functional studies and confirm the molecular diagnosis in these families.

NGS is a revolutionary technique that over the past few years has been providing significant breakthroughs in the knowledge of the molecular bases of Mendelian diseases, including ITs. Its application in the diagnostic process of IT would allow us to screen directly at least all the known IT genes. However, understanding the effects of the thousands variants identified remains a major problem. For this reason, it is fundamental to develop further functional studies for other genes.

RIASSUNTO

Le piastrinopenie ereditarie (IT) sono un gruppo di malattie rare caratterizzate da una bassa conta piastrinica (inferiore a 150.000/mL) e da un elevato grado di eterogeneità sia dal punto di vista clinico che genetico. Sebbene ad oggi siano state identificate mutazioni in almeno 30 geni diversi, il 50% dei pazienti rimane ancora privo di una diagnosi molecolare, in parte per la complessità dell'inquadramento diagnostico e in parte perché molti pazienti sono probabilmente affetti da forme di IT non ancora caratterizzate. Per questo motivo il nostro laboratorio, in collaborazione con l'Unità di Genetica Medica del Policlinico Sant'Orsola-Malpighi di Bologna e La Medicina Generale 3, IRCCS San Matteo di Pavia, è impegnato in un progetto volto ad analizzare i pazienti affetti da piastrinopenia ereditaria mediante tecniche di "next-generation sequencing" (NGS) allo scopo di identificare le mutazioni e i geni responsabili della malattia. Tuttavia, la comprensione degli effetti delle migliaia di varianti identificate con queste tecnologie resta un grave problema in disordini come le piastrinopenie, che sono principalmente malattie autosomiche dominanti causate da mutazioni "private", spesso missense. Mentre per le mutazioni "deleterie" come quelle nonsenso, *frameshift* o determinate da grandi delezioni del gene la relazione variante/patogenicità è spesso immediata a causa delle importanti conseguenze strutturali sulla composizione proteica, determinare il ruolo patogenetico di una variante *missense* è un processo complesso che richiede studi funzionali mirati.

Il mio progetto di dottorato si inserisce in questo contesto e ha come scopo lo sviluppo di saggi funzionali per verificare la patogenicità delle varianti di *ACTN1* e *TUBB1*, due geni del citoscheletro che sono stati recentemente identificati come responsabili di IT.

In seguito all'analisi di WES su un paziente piastrinopenico risultato negativo allo screening per tutti i geni allora noti, nel 2013 è stata identificata la sostituzione nucleotidica c.2305G>A all'interno dell'esone 19 del gene *ACTN1*. Questa scoperta, unita alla volontà di fornire una diagnosi molecolare ai molti pazienti della nostra coorte che ne erano ancora privi, ci ha spinto a condurre uno screening di mutazioni su questo gene in altri 127 pazienti in cui le forme note di IT erano già state escluse. Grazie a questo studio sono state identificate altre 9 varianti missense in 10 famiglie, di queste 3 (p.Glu225Lys, p.Arg738Trp e p.Arg752Gln) erano già state descritte (Kunishima et

al.,2013), mentre 6 erano nuove varianti non presenti nelle banche dati degli SNP (p.Asp22Asn, p.Arg46Trp, p.Gly215Arg, p.Asp666Val, p.Thr737Asn e p.Gly764Ser). Per ciascuna delle varianti identificate sono stati considerati gli effetti sulla struttura e la funzionalità della proteina mediante l'utilizzo di programmi bioinformatici di predizione di patogenicità quali SIFT, PolyPhen2 e Mutation Taster. E' stata, inoltre, valutata la segregazione delle varianti nelle rispettive famiglie che è stata confermata in tutti i casi, ad eccezione della sostituzione p.Asp666Val. Per confermare le analisi genetiche effettuate su queste varianti, abbiamo sviluppato un saggio di immunofluorescenza su una linea di fibroblasti immortalizzati transfettata con il plasmide codificante per la forma wild-type (wt) di *ACTN1* o con i costrutti codificanti per le proteine mutate generati mediante mutagenesi sito-specifica del plasmide wt. L'analisi al microscopio confocale ha rivelato l'assenza di fasci di actina e una completa disorganizzazione del citoscheletro in tutte le cellule trasfettate con tutte le nuove varianti descritte ad eccezione di una (Asp666Val), confermando così la loro patogenicità. La determinazione della patogenicità delle varianti identificate attraverso lo screening di mutazioni ha permesso ai clinici di definire ACTN1-RT come una forma lieve di trombocitopenia, senza disfunzione piastrinica. Questi dati sono stati oggetto di una pubblicazione sulla rivista scientifica "Blood" (Bottega et al, 2015). Le analisi di next generation sequencing hanno portato in quest'ultimo anno all'identificazione di altre 6 varianti nel gene *ACTN1* di cui 3 nuove. Anche in questo caso abbiamo effettuato analisi di immunofluorescenza confermando gli effetti deleteri sull'organizzazione del citoscheletro solo per 2 mutazioni. Lo stesso approccio ha permesso anche l'identificazione di potenziali varianti patogenetiche in TUBB1, un altro componente citoscheletrico le cui sostituzioni nucleotidiche potrebbero essere verificate da punto di vista funzionale utilizzando la stessa strategia messa a punto per α-actinina1. Sfortunatamente, ad oggi, non sono disponibili parenti per lo studio di segregazione. Tuttavia, abbiamo generato i costrutti necessari per eseguire gli studi funzionali e confermare la diagnosi molecolare di queste famiglie.

Le tecnologie di sequenziamento di nuova generazione sono uno strumento rivoluzionario che ha permesso, nel corso degli ultimi anni, scoperte significative, le quali hanno migliorato la comprensione delle basi molecolari delle malattie mendeliane, comprese le piastrinopenie. La loro applicazione nel processo diagnostico delle IT potrebbe consentirci di "screenare" in un'unica analisi tutti i geni causativi noti. Tuttavia, la comprensione degli effetti funzionali delle migliaia di varianti identificate resta un grave problema e, per questo motivo, è di fondamentale importanza sviluppare dei saggi che permettano di confermare o meno l'effetto patogenetico delle sostituzioni nucleotidiche anche in altre classi di geni.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Biogenesis of Platelets

Platelets are small, anucleate, discoid cells roughly 1-3 µm in diameter that play an essential role in hemostasis, wound healing, angiogenesis, inflammation, and innate immunity.

Platelets are generated from the cytoplasm of precursor cells, megakaryocytes (MKs), through a complex process named megakaryocitopoiesis (**Figure 1**).

In early phase, the megakaryocitopoiesis occurs within yolk sac and fetal liver and it consists of hematopoietic stem cells (HSCs) differentiation and maturation into MKs [1]. The differentiation of MKs in bone marrow is mediated by thrombopoietin (THPO), which binds to the c-MPL receptor on the surface of the cell. This binding promotes endomitosis, a process implying DNA replication (up to 64 fold) without cell division. Proportionately with diploidy, MK cytoplasmic volume increases, becomes full of platelet's specific organelles, granules, cytoskeletal proteins and develop a highly tortuous invaginated membrane system (IMS) that acts as a membrane reservoir for platelets production.

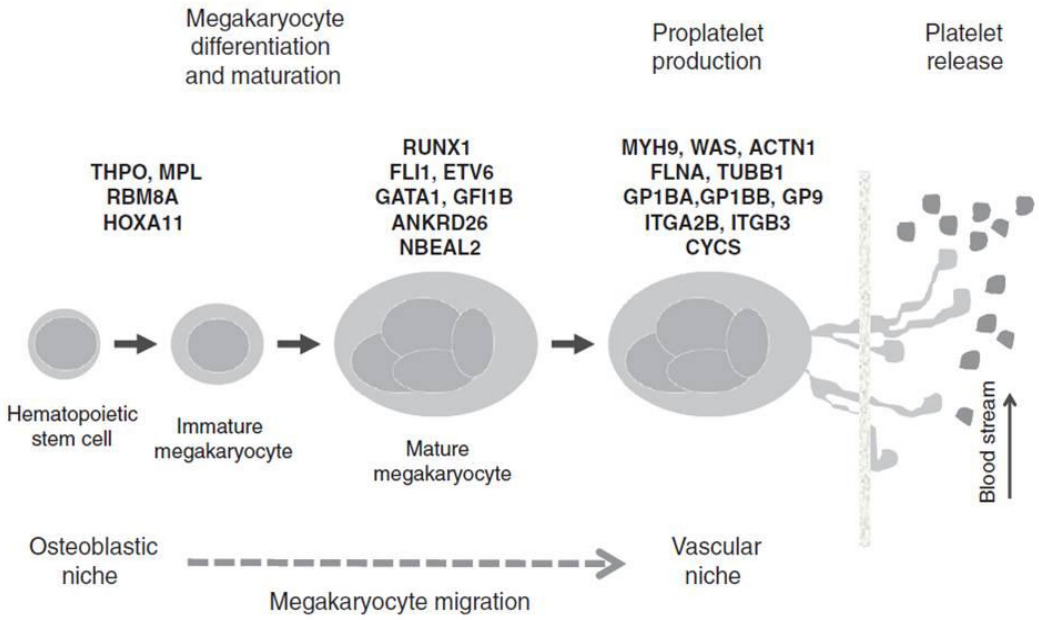

Figure 1. Schematic representation of megakaryopoiesis and platelets release. The cartoon represents the different phases of platelets biogenesis, including differentiation and maturation of megakaryocytes, proplatelets production and platelets release into blood stream. [2]

During this phase, MKs migrate from the osteoblastic to the vascular niche of the bone marrow, where they form pseudopodia. Pseudopodia elongate and branch generating long processes (proplatelets) extending into the sinusoidal blood vessels. Proplatelets formation is mainly driven by apoptotic pathways and remodeling of the microtubule cytoskeleton [3]. Microtubules of α/β-tubulin dimers organize into thick bundles where the pseudopodia originate and elongate forming increasingly narrower bundles as the proplatelets extend. At the distal end, the microtubule bundle loops and generates a teardrop-shaped structure that is released into the blood stream as a preplatelet, an intermediate form then converted into single platelets [4, 5]. Abnormalities in each step of platelets biogenesis due to mutations affecting the different players involved can result in clinical disorders known as inherited thrombocytopenias.

1.2 Inherited Thrombocytopenias

Inherited thrombocytopenias (IT) are a heterogeneous group of diseases characterized by platelet count lower than $150x10^9$/L. ITs are considered exceedingly rare, though their prevalence is unknown because no population-based study has been performed. The most important feature of ITs is increased bleeding tendency, which consists of mucocutaneous haemorrhages, nose bleeds, menorrhagia, and gastrointestinal bleeding. Patients with platelet counts lower than $20\text{-}30x10^9$ /L suffer from spontaneous often life-threating haemorrhages since birth; in most cases the degree of thrombocytopenia is mild and bleeding is occasional and related to trauma or surgery. However, some patients show a bleeding tendency disproportionate to the degree of thrombocytopenia because of defects associated with platelet function [6]. Furthermore, approximately half of the ITs are syndromic disorders, associated with additional defects that could affect nearly all organs and system.

The large variability in clinical phenotype derives from a wide genetic heterogeneity, with mutations in at least 20 genes responsible for the ITs identified so far. However, these genes account for approximately 50% of the IT patients, suggesting that novel forms are still to be characterized.

1.2.1 Classification of inherited thrombocytopenias

Based on the stage of platelets biogenesis in which the genetic defect occurs, we can classify ITs in three major groups [2, 7] (**Table 1**):

- **Defects in MKs production**. When the early phases of megakaryopoiesis are defective, proliferation of the committed cells is impaired resulting in absence or reduction of MKs in the bone marrow (BM). The best-known amegakaryocytic thrombocytopenia is CAMT, an autosomal recessive disease caused by mutations in *MPL*, the gene encoding for the THPO receptor [8].

  There are at least another two ITs that could be included in this group, radioulnar synostosis with amegakaryocytic thrombocytopenia (RUSAT) and thrombocytopenia-absent radius syndrome (TAR), both associated with skeletal defects of the radius [9-11].

  Because of the process involved, these forms of IT are usually severe from birth. Moreover, CAMT and RUSAT even if less frequently evolve into BM aplasia within the first years of

life and results in death whenever patients do not undergo successful haematopoietic stem cell transplantation. On the contrary, the platelet count in patients with TAR ameliorates during the first years of life and patients usually do not develop additional cytopenias.

- **Defects in MKs maturation.** Defects of megakaryopoiesis may also appear later, during the maturation of MKs. Interestingly, most of the ITs with defects in this phase are associated with alterations of genes encoding for transcription factors. These constitute a complex network of numerous elements including FLI1, a regulator of divergence of bipotential megakaryocytic/erythroid progenitors into platelets. Loss of one copy of *FLI1* is responsible for Paris-Trousseau thrombocytopenia (TCPT) and Jacobsen Syndrome (JBS) [12]. Another transcription factor governing the maturation of MKs and erythroid cells is GATA1. Variations in this gene cause an X-linked thrombocytopenia with either dyserythropoietic anemia (XLTDA) or β-thalassemia (XLTT) [13]. The phenotype of the GATA1-related disease is similar to individuals affected by GFI1B-related thrombocytopenia, suggesting that GFI1B and GATA1 are transcription factors that cooperate in the same cascade of events controlling the expression of genes essential for platelet and red cell formation. In these patients maturation of MKs is severely compromised and platelets are large with reduced α-granule content [14]. When α-granules lack completely, individuals are affected with gray platelet syndrome (GPS), which is due to mutations in *NBEAL2* gene [15, 16].

  Another group of ITs with defects in the MK maturation is characterized by an increased risk of developing hematological malignancies :

  1) the familial platelet disorder with propensity to acute myelogenous leukemia (FPD-AML), which is caused by mutations of *RUNX1* [17, 18];

  2) ANKRD26-related thrombocytopenia (ANKRD26-RT), due to mutation in 5'UTR of *ANKRD26* [19] and

  3) thrombocytopenia caused by mutations in *ETV6* (ETV6-RT), a tumor suppressor whose somatic alteration have been associated with leukemia or myelodysplastic syndromes [20, 21]

Disease (abbreviation in this paper, OMIM entry)	Inheritance	Gene (chromosome localization)	Other features
Defective megakaryocytic differentiation			
Congenital amegakaryocytic thrombocytopenia (CAMT, 604498)	AR	*MPL* (1p34)	Evolves into bone marrow aplasia in infancy. Normal-sized platelets
Thrombocytopenia with absent radii (TAR, 274000)	AR	*RBM8A* (1q21·1)	Platelet count tends to rise and often normalizes in adulthood. Reduced megakaryocytes. Normal-sized platelets. Bilateral radial aplasia +/− other malformations
Congenital thrombocytopenia with radio-ulnar synostosis (CTRUS, 605432)	AD	*HOXA11* (7p15-14)	Radio-ulnar synostosis +/− other defects. Possible evolution into aplastic anaemia. Normal-sized platelets
Defective megakaryocyte maturation			
Familial platelet disorder and predisposition to acute myeloid leukaemia (FPD/AML, 601399)	AD	*CBFA2* (21q22)	Possible development of leukaemia or MDS. Normal-sized platelets
Paris-Trousseau thrombocytopenia (TCPT, 188025/600588), Jacobsen syndrome (JBS, 147791)	AD	Large deletion (11q23-ter)	Cardiac and facial defects, developmental delay +/− other defects. Large platelets
GATA1-related diseases (*GATA1*-RDs, Dyserythropoietic anaemia with thrombocytopenia, 300367 – X-linked thrombocytopenia with thalassaemia, 314050)	XL	*GATA1* (Xp11)	Haemolytic anaemia, possible unbalanced globin chain synthesis, possible congenital erythropoietic porphyria. Large platelets
GFI1B-related thrombocytopenia (*GFI1B*-RT, nd)	AD	*GFI1B* (9q34·13)	Red blood cells anisocytosis. Large platelets
ANKRD26-related thrombocytopenia (THC2, 313900)	AD	*ANKRD26* (10p2)	Possible development of leukaemia or MDS. Normal-sized platelets
Gray platelet syndrome (GPS, 139090)	AR	*NBEAL2* (3p21·1)	Thrombocytopenia worsens with age. Evolutive myelofibrosis and splenomegaly. Giant platelets
Defective proplatelet formation and/or platelet release			
MYH9-related disease (*MYH9*-RD, nd)	AD	*MYH9* (22q12-13)	Cataracts, nephropathy and/or deafness. Liver enzymes may be elevated. Giant platelets
ACTN1-related thrombocytopenia (*ACTN1*-RT, nd)	AD	*ACTN1* (14q24)	Large platelets
FLNA-related thrombocytopenia (*FLNA*-RT, nd)	XL	*FLNA* (Xq28)	Periventricular nodular heterotopia (MIM 300049). Large platelets
Wiskott–Aldrich syndrome (WAS, 301000)	XL	*WAS* (Xp11)	Severe immunodeficiency leading to death in infancy. Small platelets
X-linked thrombocytopenia (XLT, 313900)			Mild immunodeficiency. Small platelets
Bernard-Soulier syndrome (BSS, 231200)			
Biallelic	AR	*GP1BA* (17p13), *GP1BB* (22q11), *GP9* (3q21)	Giant platelets
Monoallelic	AD	*GP1BA* (17p13)	Large platelets
ITGA2B/ITGB3-related thrombocytopenia (*ITGA2B/ITGB3-RT*, nd)	AD	*ITGA2B* (17q21·31), *ITGB3* (17q21·32)	Large platelets
TUBB1-related thrombocytopenia (TUBB1-RT, nd)	AD	*TUBB1* (6p21·3)	Giant platelets
CYCS-related thrombocytopenia (*CYCS-RT*, 612004)	AD	*CYCS* (7p15·3)	Normal-sized platelets

OMIM, Online Mendelian Inheritance in Man; AD, autosomal dominant; AR, autosomal recessive; XL, linked to chromosome X; MDS, myelodysplastic syndrome; nd, not defined.

Table 1. Main features of inherited thrombocytopenias classified according to the main mechanism of defective platelet production [7].

- **Defects in proplatelet formation and platelet release.** The largest group of ITs is associated with defects in proplatelet formation and platelet release, whereas MK differentiation and maturation are preserved. Mutations affecting the GPIb-IX-V complex (Bernard-Soulier syndrome) or filamin 1 (FLNA-related thrombocytopenia), as well as the

disruption of cytoskeleton structure induced by mutations in *ACTN1* (ACTN1-related thrombocytopenia) and *TUBB1* (TUBB1-related thrombocytopenia) or the constitutive activation of the GPIIb-IIIa complex induced by monoallelic mutations in *ITGA2* or *ITGB3* (ITGA2/ ITGB3-related thrombocytopenia) prevent proplatelets formation [22-28]. In addition to impaired formation of proplatelets, thrombocytopenia might be related to premature, ectopic release of platelets, as it occurs in MYH9-related thrombocytopenia, CYCS-related thrombocytopenia, Wiskott–Aldrich syndrome and X-linked thrombocytopenia [29-31].

In particular, my research has focused on two genes, *ACTN1* and *TUBB1*, encoding for cytoskeleton components, which have recently been identified as responsible for two different forms of IT.

1.2.1.1 *ACTN1* and α-Actinin

ACTN1 is the gene encoding α-actinin 1, one of the two non-muscle isoforms of α-actinin mainly expressed in megakaryocytes and platelets. Alpha-actinin is a cytoskeletal actin-binding protein and a member of the spectrin superfamily, which comprises spectrin, dystrophin and their homologues and isoforms. It is organized in an actin binding domain (ABD) constituted of two calponin homology domains (CHD) at the N-terminus, four spectrin repeats and a calmodulin-like domain (CaM) at the C-terminus. Antiparallel molecules dimerize resulting in rod-like structures that are able to cross-link the actin filaments into bundles thanks to the presence of the ABD at each boundary [32] (**Figure 2**).

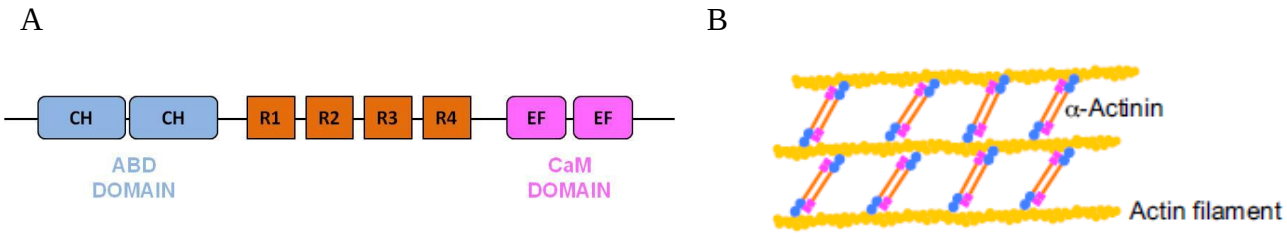

Figure 2. α-Actinin domain structure and interaction with Actinin. a) α-actinin-1 consists of an N-terminal actin-binding domain (ABD), composed of two calponin homology domains (CH), four spectrin repeats (R1–R4), and a C-terminal calmodulin-like domain (CaM). Two molecules form an antiparallel dimer; b) α-actinin-1 cross-links actin filaments into actin-filament bundles (Adapted from Kunishima et al., 2013).

1.2.1.2 *TUBB1* and ß-tubulin

TUBB1 is the gene encoding for ß1-tubulin, the predominant isoform of ß-tubulin present in MKs and platelets. α-ß tubulin heterodimers are the principal structural subunit of microtubules, which play a central role in platelet biogenesis [33] (**Figure 3**).

As mentioned before, megakariocytes generate platelets by remodeling their cytoplasm into long proplatelet extension, a microtubules-driven process. Just prior to proplatelet formation, microtubules consolidate in a mass just beneath the cortical plasma membrane [34].

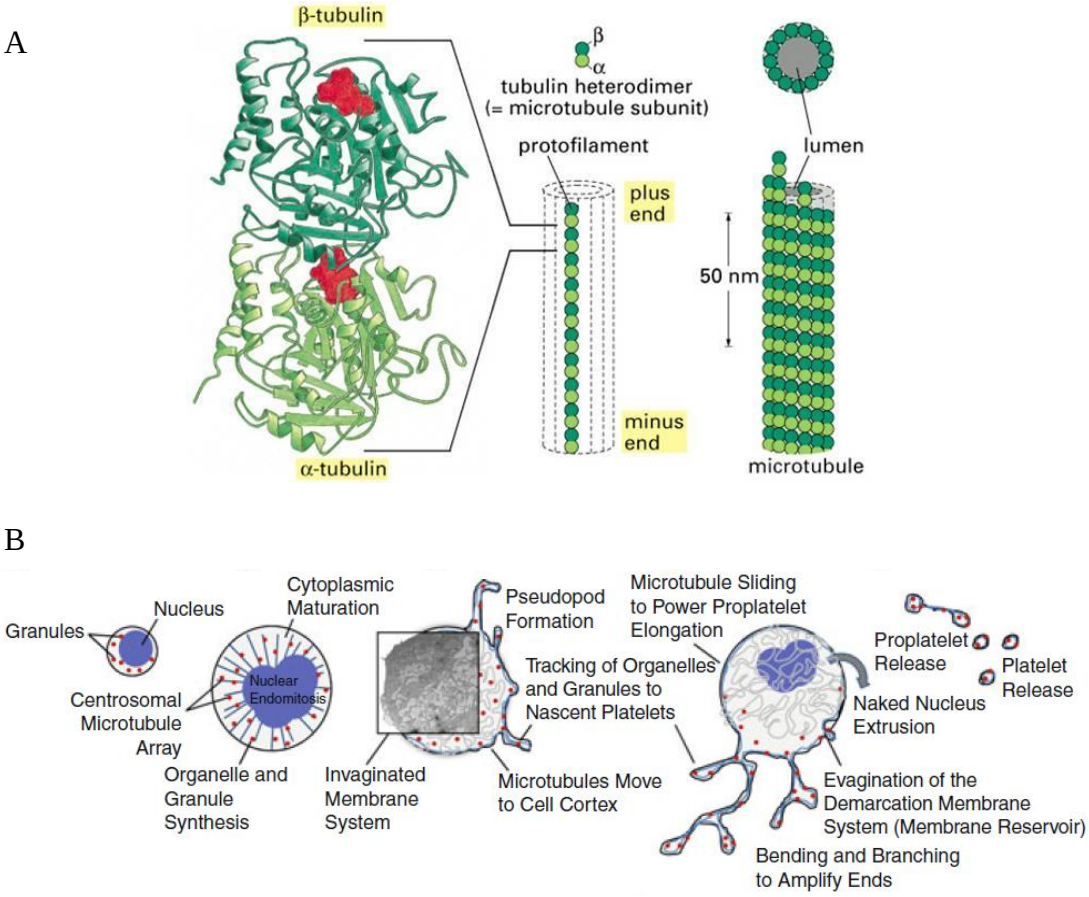

Figure 3. Structure and functions of microtubules in platelet biogenesis. a) Schematic representation of α-ß tubulin heterodimer, protofilament and microtubule (Alberts, Molecular Biology of the Cell, 5th Ed.); b) Cytoskeletal mechanisms of proplatelet production and platelet release: immature megakaryocytes will undergo repeated cycles of nuclear endomitosis for the purpose of supporting organelle synthesis, and dramatic cytoplasmic maturation and expansion. Prior to the onset of proplatelet formation, centrosomes disassemble and microtubules translocate to the cell cortex. Thick bundles of microtubules fill the shafts and cortex of broad pseudopodia that are subsequently extended by the megakaryocyte. Sliding of overlapping microtubules drives proplatelet elongation as organelles are tracked into proplatelet ends. Proplatelet formation continues to expand throughout the cell while bending and branching amplify existing proplatelet ends. Sliding movements by microtubules in the shaft elongate released proplatelets further, and separate the ends from the shaft, mediating platelet release (Thon and Italiano, 2012).

Proplatelet formation begins when these microtubules align into bundles and fill the cortex of the first blunt processes extended by megakaryocytes. Proplatelet shafts continue to become filled with thick bundles of hundreds of microtubules that undergo a thinning phase (to ~20 microtubules) and loop around within the proplatelet to reenter the shaft forming buds at the proplatelet tip [35]. Proplatelet extension is due to the continuous polymerization of tubulin bundles at their free plus ends, and dynein-powered sliding of overlapping microtubules [1]. Because of the key role of microtubules in platelet formation, mice deficient in ß1-tubulin develop moderate thrombocytopenia as a result of reduced proplatelet formation and their spherocytic platelets carry a structurally defective marginal band and reduced microtubule content [33]. More recently Kunishima and colleagues reported *TUBB1* mutations affecting microtubule assembly in the context of inherited thrombocytopenia [25, 36]

1.2.2 ITs diagnosis

The recognition of genetic origin of ITs is often hampered by several difficulties. First of all, the clinical recognition of ITs is often delayed because an overlapping between ITs and acquired forms of thrombocytopenia. Thus, to distinguish these forms, it is important to carefully analyze the family medical history of probands. Moreover, ITs could be often misdiagnosed because of ambiguous platelet count that results both from the combination of ethnic-, age- and sex variables and from errors in count measurement by cell counters, especially for patients with very large platelets [37]. Once an IT is suspected, to avoid the problem of variable expressivity of clinical features, the diagnostic process usually take advantage of the algorithm proposed by the Italian Platelet Study Group in 2003 [38] (**Figure 4**).

The first step consists of discrimination of the syndromic forms from the nonsyndromic ones. The latter are then classified according to the platelet size. Finally, to select potentially candidate genes, different investigations tests, such as blood film and bone marrow evaluation, ristocetin response analysis and immunofluorescence test, are performed. Despite of the application of diagnostic algorithm, in about the 50% of the IT patients mutations are not identified and families remain without a molecular diagnosis [6].

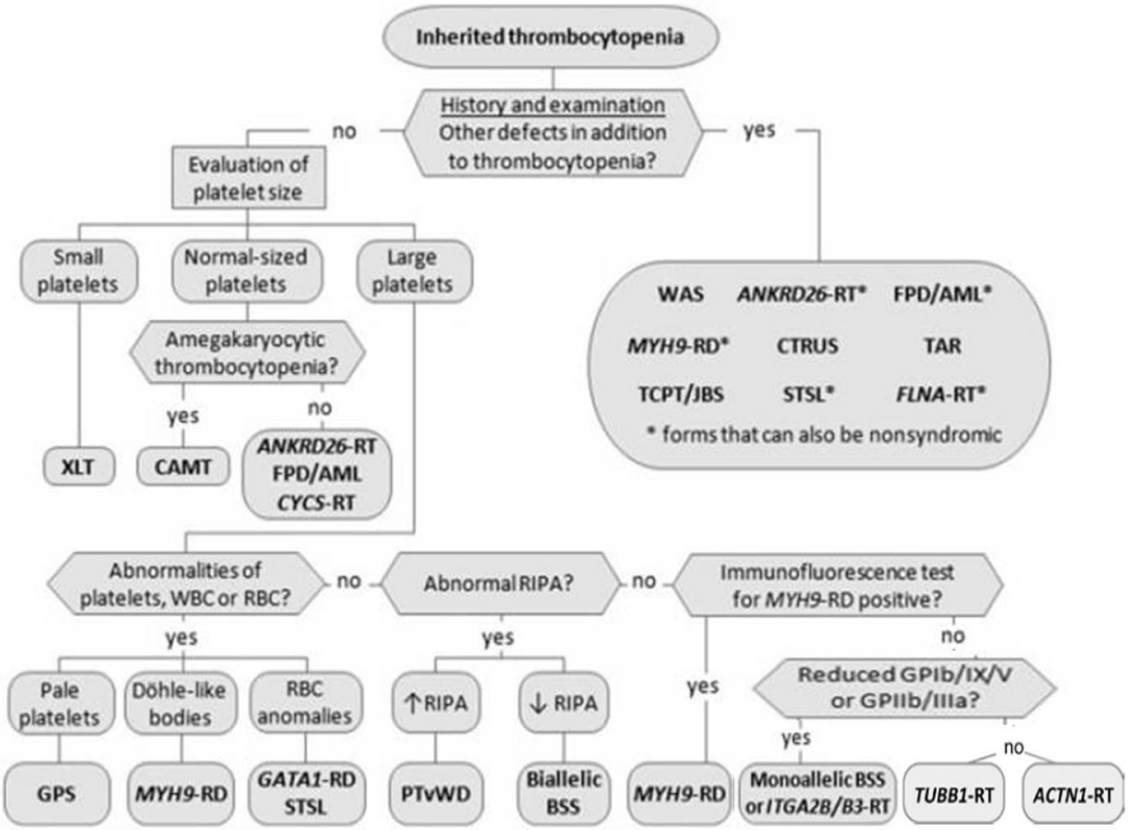

Figure 4. Diagnostic algorithm for ITs. Medical history and physical examination are sufficient in most cases for suspecting syndromic forms. For the nonsyndromic forms, further investigations of the platelets size and laboratory tests are necessary (Adapted from Balduini et al, 2013).

1.3 Next Generation Sequencing

DNA sequencing has come a long way since the days of two-dimensional chromatography in the 1970s. With the advent of capillary electrophoresis (CE)- based sequencing in 1977, scientists gained the ability to sequence the full genome of any species in a reliable, reproducible manner [39]. This technology, known as "Sanger Sequencing", had dominated the industry for almost 30 years and led to a number of accomplishments, including the completion of the human genome sequence. In recent years, sequence data have become more and more relevant for the assessment of disease at the molecular level and the ever increasing need for large-scale analysis of large quantities of samples has highlighted the limitations of automated Sanger sequencing, despite many technical improvements occurs in this "era". For this reason over the past ten years, there has been a fundamental shift away from the application of automated Sanger sequencing for genome analysis and the development of Next Generation Sequencing (NGS) technologies able of producing large volumes of sequence data in a short time and at low cost [40]. Important applications include whole genomes and transcriptomes sequencing, genome resequencing, DNA-protein interactions through chromatin immunoprecipitation sequencing, discovering noncoding RNAs, metagenomics, and other applications that will appear over the next few years.

Application of NGS strategies has provided new insights into the field of ITs with the identification of novel IT genes, such as *NBEAL2*, *GFI1B*, *RBM8A* and *PRKACG* responsible for gray platelet syndrome, thrombocytopenia and absent radii and PRKACG-related thrombocytopenia, respectively [14, 15, 41, 42]. In 2015, mutations were identified also in *FYB* [43], *SLFN14* [44] and *ETV6* [20, 21], driving up to 30 the IT genes currently known and further reducing the number of families that remain without a molecular diagnosis.

Application of this technology will have a strong impact not only in cloning novel genes but also in molecular genetic testing using target panels of all the genes responsible for ITs.

2. AIM OF THE STUDY

Diagnosis of IT is a complex difficult process due to high grade of heterogeneity from both the clinical and genetic point of view. Moreover, almost half of the IT patients still remain without molecular diagnosis because they suffer from forms that are not identified yet. For this reason, in collaboration with Medical Genetic Unit of Policlinico Sant'Orsola Malpighi in Bologna and the Department of General Medicine 3, IRCCS San Matteo in Pavia, our laboratory is involved in a project aimed to studying IT patients using NGS technologies. The purpose of this project is to:

1) simplify and shorten the time for the molecular diagnosis using a target sequencing panel of known causative genes;

2) perform whole exome sequencing in patients without mutations identified by the target sequencing panel in order to identify new candidate genes.

My PhD work fits into this project since it is focused on development of functional assays to test pathogenicity of the variants identify by NGS. Indeed, whereas there is no doubt on the deleterious effects of nonsense or frameshift, the pathogenic role of missense alterations is not always obvious. In this condition, search of the disease-causing gene is further prevented when the model of inheritance requires only mutated allele. Therefore, in diseases like ITs, which are mainly autosomal dominant diseases caused by private, often missense, the development of targeted functional studies is absolutely necessary to determine the pathogenicity of missense variants.

3. MATERIALS AND METHODS

3.1 Patients

Our cohort of patients included 239 consecutive probands with IT examined at the Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS) Policlinico San Matteo Foundation, in Pavia, Italy, until 2013. For mutational screening of *ACTN1* we enrolled 128 individuals without a certain diagnosis after the IT diagnostic work-up. In particular, according to their clinical features, patients were previously analyzed for mutations in *MYH9, CYCS, GP1BB, GP1BA, GP9* and *ANKRD26* genes without finding any causative mutation. Patients enrolled from 2014 and some cases of our cohort still unsolved were analyzed during 2015 by means of next generation sequencing technologies (N=100).

3.2 Whole Exome Sequencing

Whole exome DNA from patient's whole blood was captured using the solid-phase NimbleGen SeqCap EZ Exome 44Mb array (Nimblegen Inc., Madison, WI, USA) and sequenced as 91/100 bp paired-end reads on Illumina HiSeq2000 platform (Illumina Inc., San Diego, CA, USA).
Generated reads were checked with FastQC (http://www.bioinformatics.babraham.ac.uk/ publications.html) and aligned with BWA (Li and Durbin, 2010) to the reference genome hg19. Aligned reads were treated for realignment and base quality score recalibration with GATK (De Pristo et al., 2011), and for duplicate removal with PicardTools (http://picartools.sourceforge.net). Alignment statistics were collected by SAMtools (http://samtools.sourceforge.net./) and GATK. Coverage statistics over the targeted regions were calculated with GATK. Variant calling and filtering by quality were performed by GATK. Variants passing quality filters were annotated using ANNOVAR (Wang and Li, 2010) against NCBI RefGene (http://www.ncbi.nlm.nih.gov).

3.3 Target Sequencing

Target sequence analysis were performed using Ion Torrent Personal Genome Machine (IPGM) platform. Sequencing primers were designed on the coding and intronic flanking regions of the IT genes using the Ion Ampliseq Designer software (https://www. ampliseq.com/browse.action). Following the manufacturer's recommendations (Life Technologies), two multiplex PCRs were

carried out for each sample using Ion AmpliSeq library kit 2.0. Emulsion-PCR and enrichment reactions were performed on the template using Ion One Touch 2 system and
the enriched-template quality was analyzed using Qubit 2.0 Fluorometer (Invitrogen Corporation). Sequencing reactions were performed using Ion PGMTM Sequencing 200 Kit v2. Sequencing data were analyzed using Ion Torrent Suite software (v 4.0). Using the plug-in Coverage Analysis (TSCA v 4.0), we evaluated the quality of data sequencing. Data were aligned with hg19 human genomic sequence using the plug-in Variant Caller (TSVC v.4.0). Functional annotations of all the sequence variants were performed using the w-annovar software (http://wannovar.usc.edu/). The exons containing variants were confirmed by Sanger sequencing using standard conditions in an ABI 3100 automated sequencer (Applied Biosystems, Foster City, CA).

3.4 Sanger Sequencing

ACTN1 was screened for mutations using genomic DNA extracted from peripheral blood. Mutational analysis was performed by polymerase chain reaction (PCR) amplification using primers covering exons 1-8 and 17-21 of the gene and the relative exon/intron boundaries. PCR was carried out in 35 μL of total reaction volume with 25 ng of genomic DNA, 10μM of each primer, and Kapa 2G Fast Hot Start ReadyMix 2X (KapaBiosystems, Cape Town, South Africa). PCR products were bidirectionally sequenced using the ABI PRISM BigDye v3.1 Terminator Cycle Sequencing Ready Reaction Kit and ABI PRISM 3130xl sequencer (Applied Biosystems, Foster City, CA).

3.5 Bioinformatic Analysis

Multiple species alignment of α-actinin 1 was generated using the Clustal Omega program (https://www.ebi.ac.uk/Tools/msa/clustalw2/). Analysis of protein domains was performed using the database Prosite (http://prosite.expasy.org). The effect of the missense variations was evaluated using three pathogenicity prediction programs: PoliPhen-2 (http://genetics.bwh.harvard.edu/pph2/), Mutation Taster (http://www.mutationtaster.org/) and SIFT (http://sift.jcvi.org).

3.6 Cloning Procedures

A full-length *ACTN1* sequence was amplified from normal platelet cDNA and constructed into mammalian expression vector pcDNA3.1 (Invitrogen, San Diego, CA, USA) with a 5'Myc tag sequence. *ACTN1* mutants were generated by PCR using specific mutagenesis primers (view table below). The parental methylated DNA was eliminated by digesting with *DpnI* restriction enzyme. The construct obtained was transformed in TOP10 competent cells. The full-length sequences, frame, and orientation were confirmed by sequencing. The same stategy was used to cloning pcDNA3.1-Myc-TUBB1.

3.7 Immunofluorescence Assay

For immunofluorescence assay human fibroblasts (PD220) were seeded on chamberslides and transfected with myc-tagged wild type or mutant ACTN1 plasmids. After 16 hours cells were fixed with 4% parafolmaldehyde for 20 minutes and then permeabilized with 0.1% Triton X-100. For detection of ACTN1 primary antibody against c-myc (9E10, Santa Cruz Biotechnology) was used with anti mouse FITC secondary antibody (F0479, DakoCytomation),while actin filaments were stained with AlexaFluor594 conjugated phalloidin (Invitrogen).

Images were obtained with a Nikon C1si confocal microscope, containing 488nm, argon laser line and 561 nm diode laser. Light was delivered to the sample with an 80/20 reflector. Electronic zoom was kept at minimum values for measurements to reduce potential bleaching. The images were acquired using a 60X Plan Apo objectives (with a corresponding NA of 1.4), collecting series of optical images at 1 μm z resolution step size. The corresponding voxel size was 100x100x1000 nm (X×Y×Z). Images were processed for z-projection (maximum intensity), brightness and contrast regulation using ImageJ 1.45 (NIH,Bethesda, USA).

3.8 Platelet Studies

Platelet studies was performed by clinicians as reported in Bottega, Marconi, Faleschini et al., 2015

4. RESULTS AND DISCUSSION

4.1 Identification of c.2305G>A variant in *ACTN1* gene by Whole Exome Sequencing

In order to identify the causative gene in a family with an IT of unknown origin in which candidate genes were previously excluded, we performed Whole Exome Sequencing (WES) in the proband of Family 1 (F1). The WES experiments allowed us to identify a single-base substitution located within exon 19 of the *ACTN1* gene (c.2305G>A), which was predicted to result in a missense mutation, p.Glu769Lys. The mutation was absent in the public dbSNP and the 1000 Genomes Project dataset (http://www.ncbi.nlm.nih.gov/SNP/ ; http://www.1000genomes.org/category/dbsnp). Moreover, it occurs in a very conserved position, according to PhyloP score. The variant was predicted to be Benign from Polyphen2, while SIFT and Mutation Taster indicated it as deleterious. The variant was confirmed by Sanger sequencing and was observed to correctly segregate with the disease in family F1 (**Figure 1**). According to the putative causative role of this mutation, at the same time Kunishima and colleagues identified *ACTN1* as a novel gene responsible for macrothrombocytopenia [26].

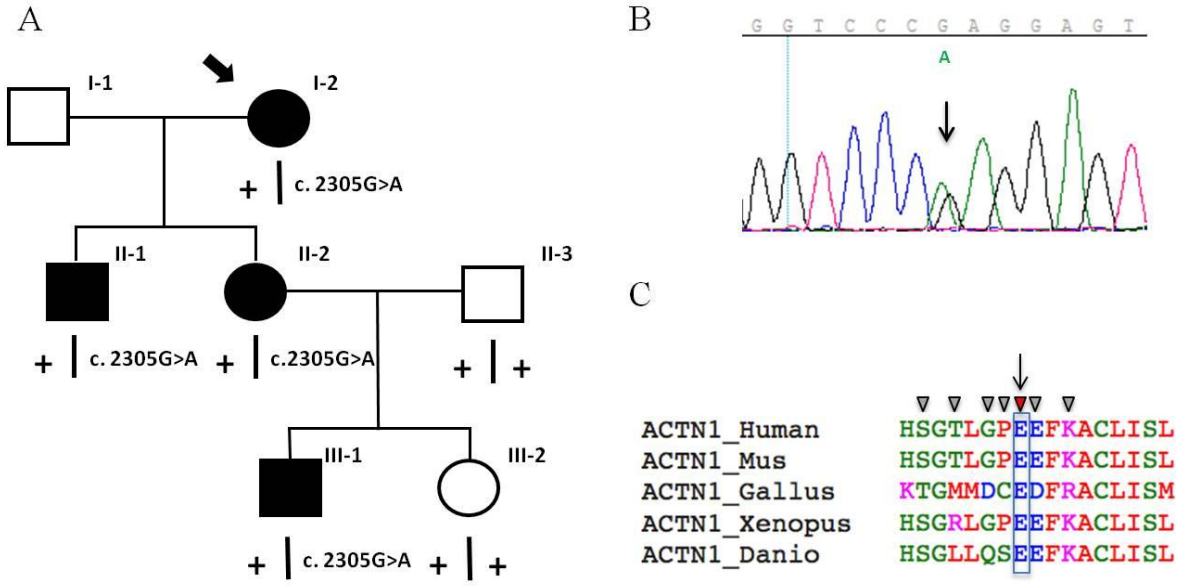

Figure 1. Identification of c.2305G>A variant in *ACTN1* gene. A) Pedigree of Family 1; B) Chromatogram of mutation confirmed in Sanger sequencing: arrow indicates the aminoacid substitution c.2305G>A; C) Aminoacid conservation: arrow indicates the aminoacid substitution Glu769Lys .

4.2 Identification of *ACTN1* mutations in 10 families

According to protein structure and mutations identified by Kunishima and colleagues, we decided to restrict the mutational screening to exons that encode for functional domains. For this reason, we performed direct Sanger sequencing of exons 1-8 and 17-21 of the *ACTN1* gene in 127 probands.
We identified nine different heterozygous missense variants in ten families (**Table1**). Three were known amino acid substitutions (p.Gly225Lys, p.Arg738Trp and p.Arg752Gln) that were previously reported [26]. The other six were not present in either the dbSNP or the 1000 Genomes Project datasets and therefore regarded as novel variants of ACTN1.

Family	Mutation		Exon	Type of mutation	References
	Nucleotide	Protein			
F2	c.64G>A	p.Asp22Asn	1	Missense	**New mutation**
F3	c.136C>T	p.Arg46Trp	2	Missense	**New mutation**
F4	c.136C>T	p.Arg46Trp	2	Missense	**New mutation**
F5	c.673G>A	p.Glu225Arg	7	Missense	Kunishima et al.,2013
F6	c.751G>A	p.Gly251Arg	8	Missense	**New mutation**
F7	c.1997A>T	p.Asp666Val	17	Missense	**New mutation**
F8	c.2210C>A	p.Thr737Asn	18	Missense	**New mutation**
F9	c.2212C<T	p.Arg738Trp	18	Missense	Kunishima et al.,2013
F10	c.2255G<A	p.Arg752Gln	18	Missense	Kunishima et al.,2013
F11	c.2290G>A	p.Gly764Ser	19	Missense	**New mutation**

Table 1. Variants identified in *ACTN1* gene. Nucleotide A of the ATG translation initiaton start site of the *ACTN1* gene cDNA in GenBank sequence NM_001130004.1 is indicated as nucleotide +1.

Interestingly, the c.136C>T (p.Arg46Trp) variant affects the same residue as it does c.137G>A (p.Arg46Gln), the mutation described by Kunishima and Guéguen [26, 45]. Although the two Italian families carrying p.Arg46Trp were not aware of a relationship, both of them come from a little village in the north of Italy, suggesting a possible founder effect for this mutation. All the missense variants segregate with macrothrombocytopenia within the individual pedigrees when family members were available, except for c.1997A>T (p.Asp666Val) (**Figure 2**). In Family 7 (F7),

in fact, c.1997A>T was identified in the proband but not in her mother and maternal aunt despite their clinical features were suggestive of macrothrombocytopenia, suggesting that the variant was inherited from the healthy father whose DNA sample was not available.

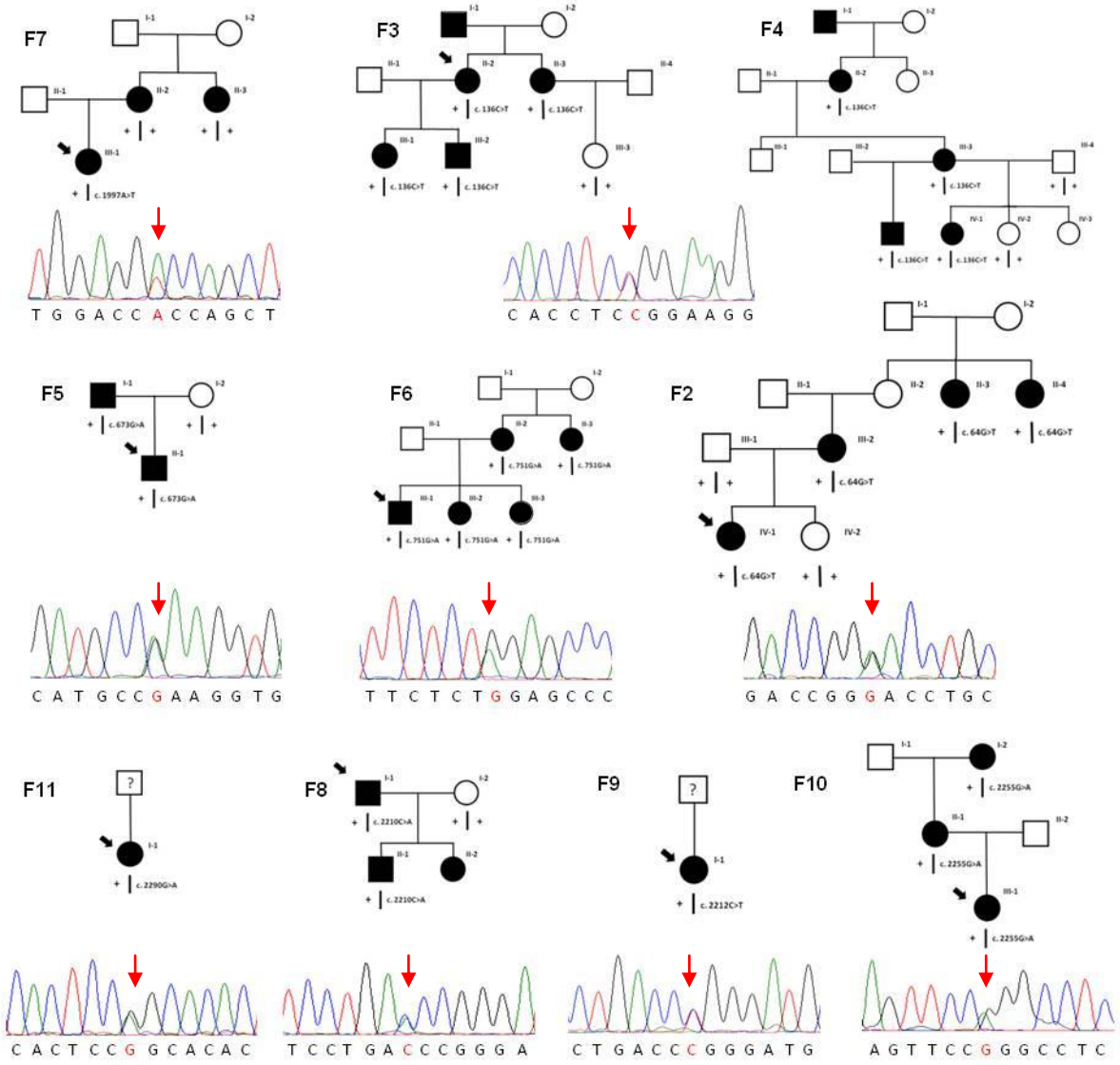

Figure 2. Mutation analysis of *ACTN1* families. Pedigrees of ten families with a heterozygous allelic variant of *ACTN1* identified by Sanger sequencing. The black arrows indicate the probands of each family. Plus symbol indicates a wild type allele in the segregation analysis. The chromatogram of each heterozygous mutation is shown under the corresponding pedigree. Red arrows indicate the nucleotide substitution in each chromatogram.

4.3 Evaluation of the effect of missense mutations on α-actinin1 structure

The multiple-sequence alignment indicated that all the six novel missense variants hit amino acid residues highly conserved from zebrafish. Furthermore, their potential effect on protein function was evaluated using three different pathogenicity prediction tools: SIFT, Polyphen and Mutation Taster (**Table 2**). Regarding p.Asp22Asn, p.Arg46Trp, p.Gly251Arg, and p.Thr737Asn, all programs predicted their deleterious effects on α-actinin1 function. In contrast, the p.Gly764Ser substitution was tolerated using SIFT even if the variant occurs in a highly conserved functional domain (calmodulin-like domain, CaM). Instead, the significance of p.Asp666Val was unclear: although it does not segregate its pathogenicity scores were high.

Mutation		PolyPhen-2	SIFT	MutationTaster
Nucleotide	Protein			
c.64G>A	p.Asp22Asn	possibly damaging	not tolerated	disease causing
c.136C>T	p.Arg46Trp	probably damaging	not tolerated	disease causing
c.751G>A	p.Gly251Arg	possibly damaging	not tolerated	disease causing
c.1997A>T	p.Asp666Val	probably damaging	not tolerated	disease causing
c.2210C>A	p.Thr737Asn	possibly damaging	not tolerated	disease causing
c.2290G>A	p.Gly764Ser	possibly damaging	tolerated	disease causing

Table 2. Bioinformatic analysis for predicting the functional effect of the novel *ACTN1* variants. Effect of variations was evaluated using three pathogenicity prediction programs: PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (http:// http://sift.jcvi.org/), and Mutation Taster (http://www.mutationtaster.org/).

4.4 Functional analysis of the novel missense mutations

Whereas the pathogenetic role of nonsense and frameshift variants or large deletions is clear (usually for RNA or protein degradation) that of other variants, such as missense or small in-frame deletions/duplications is of uncertain significance (VUS). Unraveling the effect of the VUS is a complex process that requires targeted functional studies to explore whether a single amino acid substitution alters or not the protein function, for example modifying its folding and therefore stability, or its functional domains.

For this reason, in order to determine the pathogenetic role of the seven novel missense variants identified in this study, we performed an immunofluorescence analysis after transfection of wild type or mutant *ACTN1* cDNA cloned into mammalian expression vector pcDNA3.1 with a 5' Myc tag sequence [26] into human fibroblasts (**Figure 3A**). When cells were transfected with the wild type construct, we observed a well-organized cytoskeleton in which α-actinin1 colocalizes in large part with actin along the filaments.

On the contrary, cells transfected with mutant constructs presented an abnormal distribution of actin, which was no longer able to form filaments causing an apparent disruption in the cytoskeletal structure, as observed in other studies testing the effect of different *ACTN1* missense mutations [26, 45]. This was particularly evident when we compared the actin staining in cells expressing and not expressing the mutant constructs. In addition, when mutant α-actinin1 was expressed, the staining appears widespread in the cytoplasm and the colocalization with actin was not specific. These phenotypic features are shared by all but one (p.Asp666Val) mutation. Consistent with segregation analysis, the correct organization of the cytoskeleton in cells expressing p.Asp666Val excluded this variant as a disease-causing mutation. It is worth noting that p.Asp666Val is outside the actin binding (ABD) and calmodulin-like (CaM) domains (**Figure 3B**), suggesting that only alterations of these functional regions are compatible with the disease.

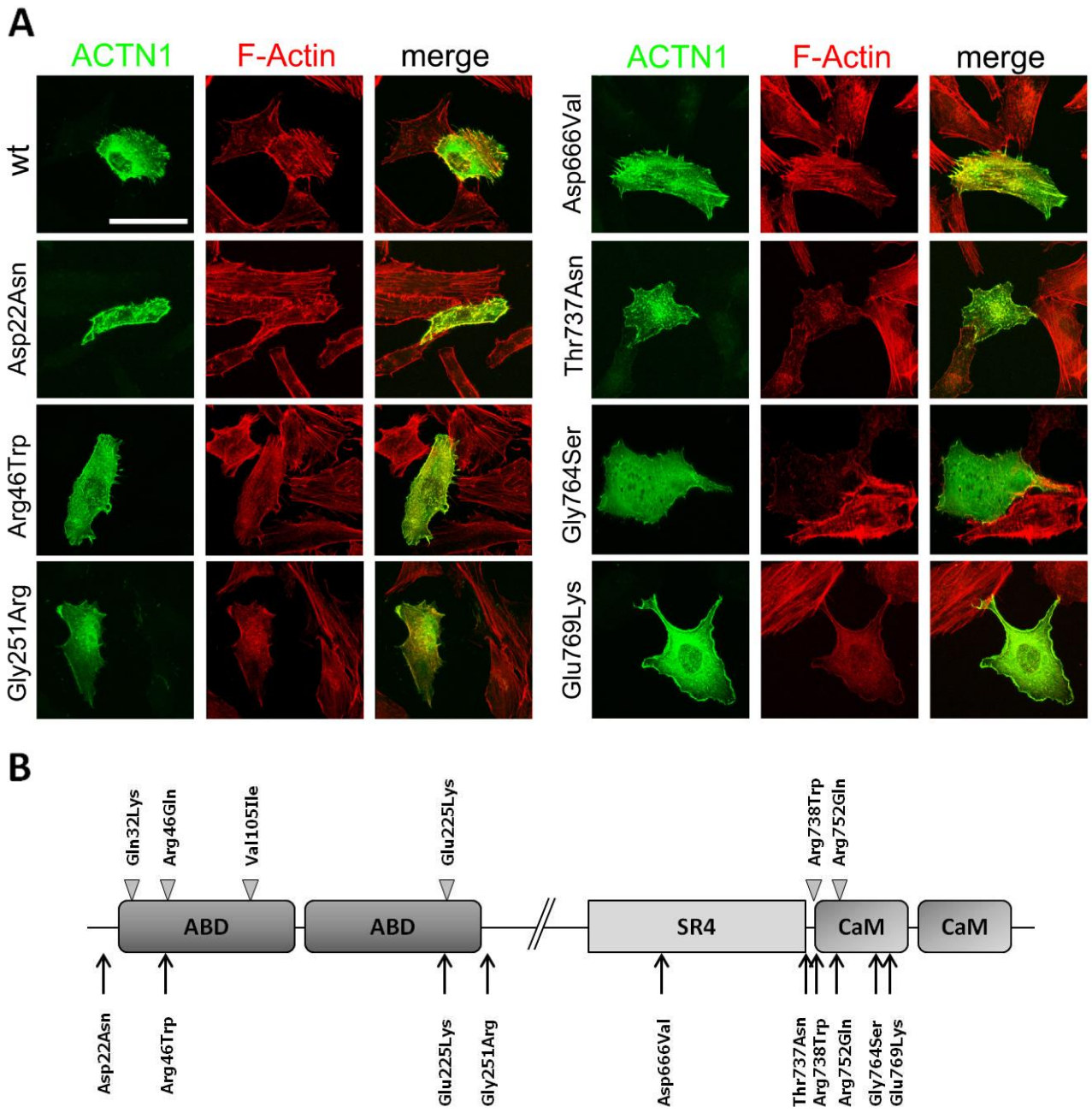

Figure 3. Functional studies of novel *ACTN1* variants. (A) Immunofluorescence analysis in PD220 fibroblast cell line transiently transfected according to standard procedures. Both wild type (top panel) or mutant (lower panels) *ACTN1* cDNAs were cloned into the pcDNA3.1-Myc tagged expression vector. The subcellular localization of exogenous α-actinin1 (green) was examined using c-myc antibodies (9E10; Santa Cruz Biotechnology, Dallas, Texas, U.S.A.) while the actin filaments were stained with AlexaFluor594 (red) conjugated phalloidin (Invitrogen, Marseille, France). Images were obtained with a Nikon C1si confocal microscope using a 60X Plan Apo objectives. Images were processed for z-projection (maximum intensity), brightness and contrast regulation using ImageJ 1.45 (NIH, Bethesda, USA). The cells shown are representative of three independent experiments. Scale bar =50 µm. (B) Domain structure of α-actinin and localization of *ACTN1* mutations identified in Japanese families (arrowheads) and in this paper (arrows). The p.Arg46Gln mutation was also identified in a French family.

4.5 Clinical features of ACTN1-Related Thrombocytopenia (ACTN1-RT)

Determination of the pathogenicity of variants identified through screening of mutation has allowed clinicians to define the clinical and laboratory features of this new form of thrombocytopenia. As reported in **Table 3**, in most cases thrombocytopenia was mild and both mean platelet volume and diameter were significantly higher in patients than in healthy subjects. The mean reticulated platelet count was significantly lower whereas the serum level of TPO was only slightly higher than in controls (these data are consistent with normal bone marrow megakaryocyte concentration and reduced platelet production). Regarding functional activity, *in vitro* platelet aggregation was within the normal range in the 17 investigated patients. The GPIbα, GPIX, GPIIb and GPIIIa glycoproteins were expressed on platelet surface at normal level, as indicated by flow cytometry of in the 24 investigated patients. Briefly, the study of affected individuals lead us to classify ACTN1-RT as a mild form of macrothrombocytopenia without platelet dysfunction, confirming that thrombocytopenia derives from defects of the phase of proplatelet formation and platelet release.

The data obtained in this part of my PhD work have been published at the beginning of 2015 on the scientific journal "Blood" (Bottega, Marconi, Faleschini et al.,2015). One of immunofluorescence included in this work was also selected for the cover of the journal.

Family (N. of patients)	Mean age at diagnosis Years (range)	WHO bleeding score[a] (N. of patients)	Mean platelet count using cell counter[b] x10^9/L (range)	Mean platelet volume fL (range)	Mean platelet diameter µm (range)	*In vitro* platelet aggregation[c] (N. of patients)	Flow cytometry of platelet glycoproteins (N. of patients)
F1 (4)	38 (3-66)	0 (1), 1 (2), 2 (1)	86 (46-120)	14-3 (12.6-15)	3.5 (2.9-3.9)	normal (4)	normal (4)
F2 (4)	43 (22-55)	0 (1), 1 (3)	107 (89-134)	11,1 (10,1-12)	2.8 (2.7-3)	nd	normal (4)
F3 (4)	46 (26-64)	0 (1), 1 (2), 2 (1)	103 (81-118)	12,5 (10.6-14.7)	3.3 (3-3.7)	normal (4)	normal (4)
F4 (4)	42 (14-72)	0 (1), 1 (1), 2 (2)	95 (66-124)	14.8 (14-15.6)	3.8 (3.5-4.1)	normal (4)	normal (3)
F5 (2)	30 (12-49)	0 (2)	103 (97-110)	11.8 (11.3-12.4)	3 (2.8-3.1)	normal (2)	normal (1)
F6 (6)	23 (7-44)	0 (1), 1 (3), 2 (3)	103 (78-**154**)	12.3 (10.4-14)	3.3 (2.6-4.3)	normal (2)	normal (5)
F8 (2)	58 (34-82)	2 (2)	58 (55-62)	10.5 (10.4-10.6)	3.2 (2.9-3.5)	normal (1)	nd
F9 (1)	nd	0 (1)	110	12.1	2.5	nd	nd
F10 (3)	25 (3-44)	0 (2), 1 (1)	112 (65-**166**)	12.5 (11.3-14.3)	2.8 (2.6-3)	nd	normal (2)
F11 (1)	48	0 (1)	117	14.4	3.3	normal (1)	normal (1)

Table 3. Clinical and laboratory features of families with *ACTN1* mutations. [a]WHO (World Health Organization) bleeding score: grade 0, no bleeding; grade 1, only cutaneous bleeding; grade 2, mild blood loss; grade 3, gross blood loss, requiring transfusion; grade 4, debilitating blood loss, retinal or cerebral associated with fatality. [b]*In vitro* platelet aggregation after collagen (4 µg/mL), ADP (5 mM) and ristocetin (1.5 mg/mL); [c]Individuals were classified as thrombocytopenic based on the reference range of platelet count (150-400 x 10^9/L) and the recently proposed age- and gender specific reference intervals (Savoia et al., 2001). According to these criteria, two three-year-old girls from families 5 and 8 were enrolled despite their platelet count (in bold) was in the normal range; nd, not determined

4.6 Identification of *ACTN1* and *TUBB1* variants through NGS approach

As mentioned before, ITs are characterized to high grade of clinical and genetic heterogeneity. Despite of application of diagnostic algorithm described before, 50% of IT patients remains without molecular diagnosis, suggesting that they are likely to suffer from IT not characterized yet.

For this reason, in collaboration with Medical Genetic Unit of Policlinico Sant'Orsola Malpighi in Bologna and the department of General Medicine 3 , IRCCS San Matteo in Pavia, since last year we have been analyzing these patients using next generation sequencing approaches based on :

- **target sequencing** (Ion Torrent) to identify rapidly variants in known IT genes;
- **whole exome sequencing** in patients negative after target sequencing to identify new genes responsible for the disease.

At present, we have analyzed 120 patients. In 13 individuals, we have identified potential pathogenetic variants of *ACTN1* and *TUBB1*, another cytoskeleton component, whose variants could be tested functionally using the same strategy as for those affecting *ACTN1* (**Table 4**).

Family	Gene	Exon	Mutation		Type of mutation	Patogenicity Prediction
			Nucleotide	Protein		
IT1	ACTN1	2	c.136C>T	p.Arg46Trp	Missense	Bottega et al., 2015
IT2	ACTN1	2	c.136C>T	p.Arg46Trp	Missense	Bottega et al., 2015
IT3	ACTN1	4	c.384G>C	p.Trp128Cys	Missense	Probably Damaging (score=1)
IT4	ACTN1	7	c.673C>T	p.Glu225Lys	Missense	Bottega et al., 2015 Kunishima et al., 2013
IT5	ACTN1	8	c.698C>T	p. Pro233Leu	Missense	Probably Damaging (score=1)
IT6	ACTN1	10	c.982G>A	p.Val328Met	Missense	Possibly Damaging (score=0.85)
IT7	TUBB1	2	c.165C>A	p.Tyr55*	Nonsense	
IT8	TUBB1	4	c.326G>A	p.Gly109Glu	Missense	Probably Damaging (score=1)
IT9						
IT10						
IT11						
IT12	TUBB1	4	c.572A>C	p.Gln191Pro	Missense	Probably Damaging (score=0.99)
IT13	TUBB1	4	c.856G>A	p.Val286Met	Missense	Probably Damaging (score=1)

Table 4. Identification of *ACTN1* and *TUBB1* variants through NGS approach. Nucleotide A of the ATG translation initiaton start site of the *ACTN1* gene cDNA in GenBank sequence NM_001130004.1 is indicated as nucleotide +1. Effect of variations was evaluated using the three pathogenicity prediction programs previously mentioned, but only PolyPhen-2 score is reported in this table.

4.6.1 *ACTN1* variants

The NGS analyses revealed 5 *ACTN1* variants in 6 families. Whereas two of these are known mutations (p.Arg46Trp and p.Glu225Lys), the others are novel variations (p.Trp128Cys, p.Pro233Leu and p.Val328Met). In all these cases, the variants segregate with macrothrombocytopenia within pedigrees when family members were available (**Figure 4**). Notably, we found a homozygous variant in the proband of Family IT6 with a severe thrombocytopenia. Her father, heterozygous for the variant, had a platelet count of 100-120 x10^9/L, while the mother was reported as healthy, even though there are no laboratory tests that confirm her condition.

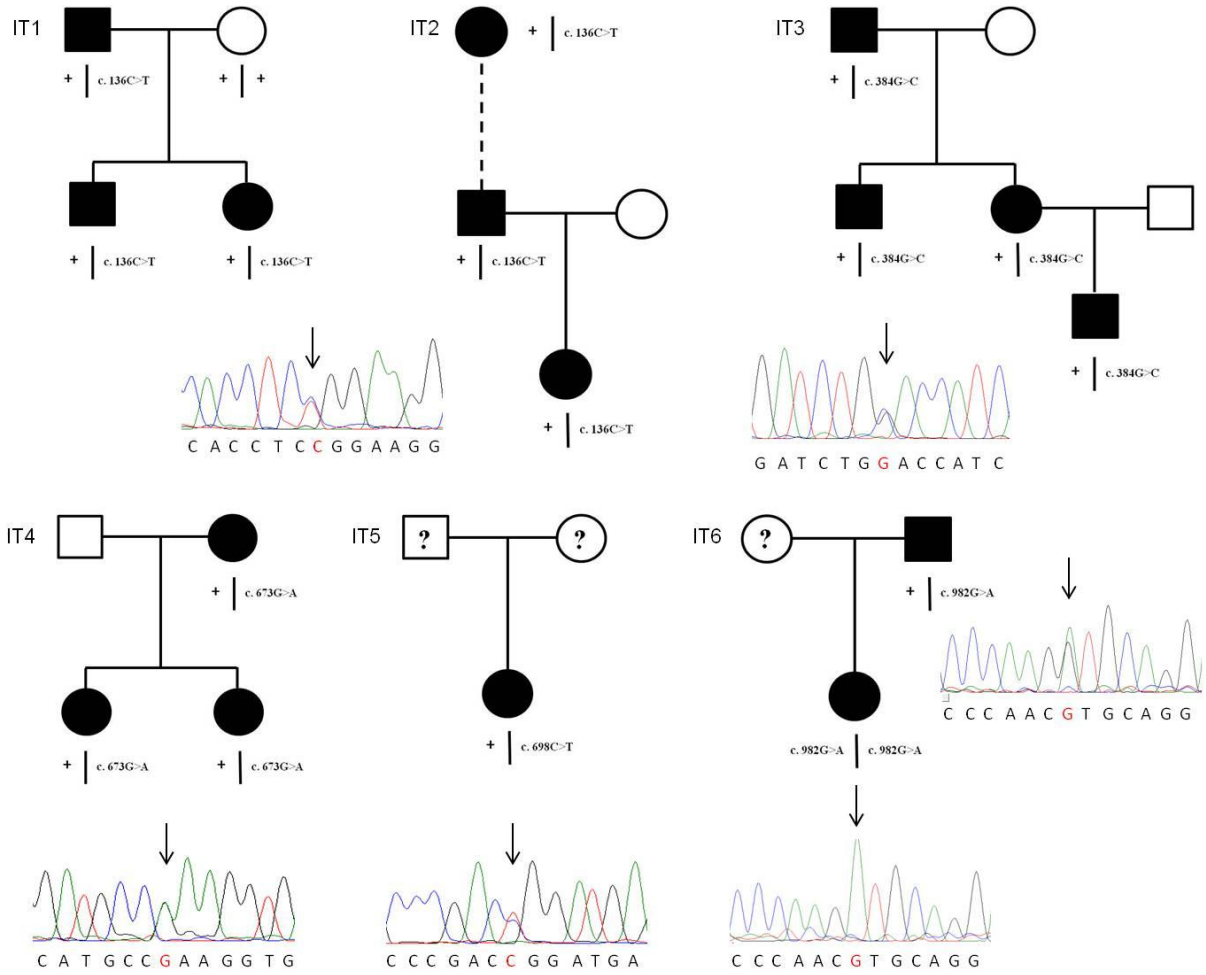

Figure 4. Pedigrees of families with a variant of *ACTN1* identified by NGS approach. Plus symbol indicates a wild type allele in the segregation analysis. The chromatogram of each mutation is shown under the corresponding pedigree. Arrows indicate the nucleotide substitution in each chromatogram.

To confirm the pathogenic role predicted by bioinformatical analysis (**Table 4**), we performed immunofluorescence assay on the three novel variants identified (**Figure 5**).

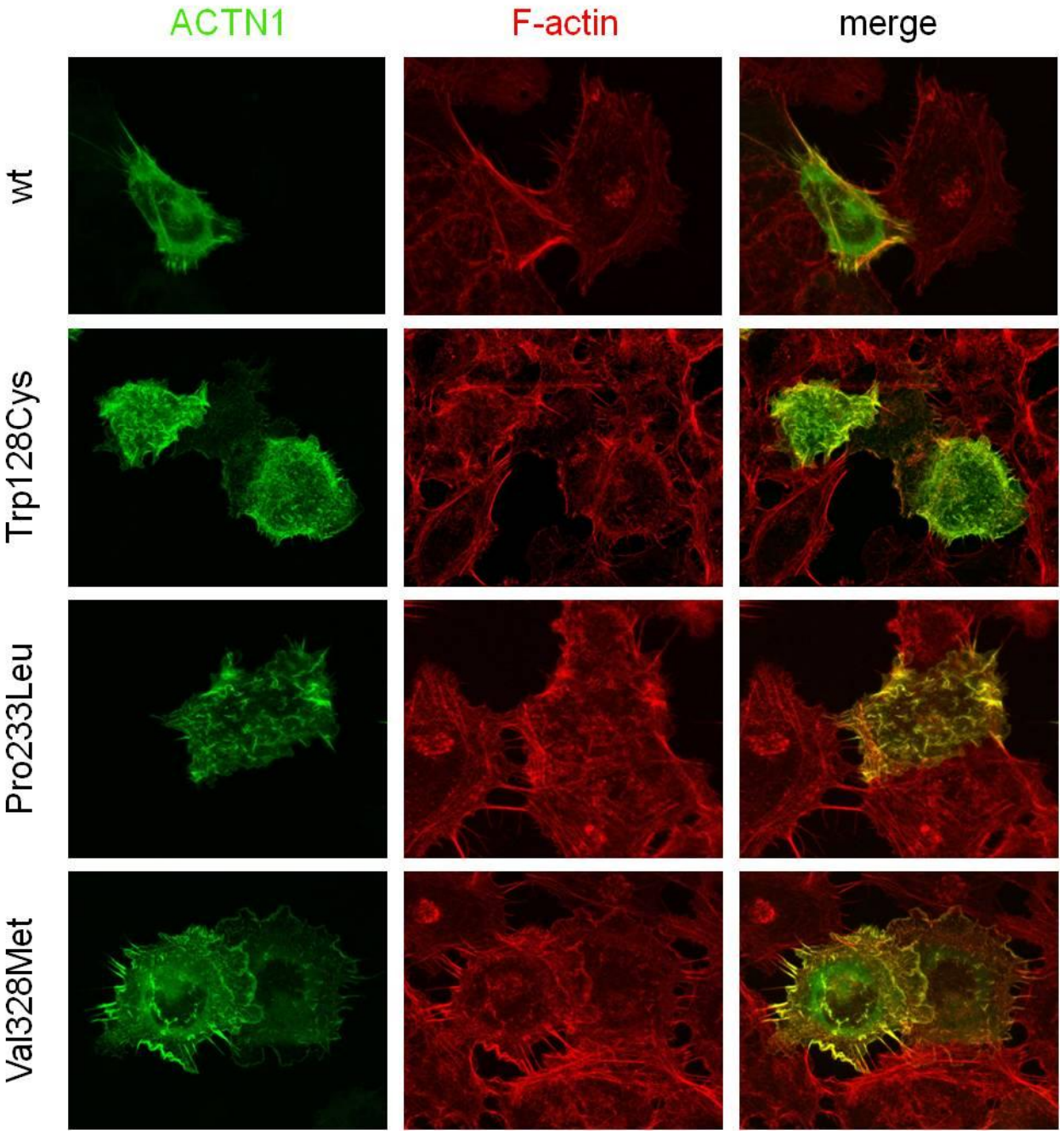

Figure 5. Functional studies of novel *ACTN1* variants. Immunofluorescence analysis in PD220 fibroblast cell line transiently transfected according to standard procedures. Both wild type (top panel) or mutant (lower panels) *ACTN1* cDNAs were cloned into the pcDNA3.1-Myc tagged expression vector. The subcellular localization of exogenous α-actinin1 (green) was examined using c-myc antibodies (9E10; Santa Cruz Biotechnology, Dallas, Texas, U.S.A.) while the actin filaments were stained with AlexaFluor594 (red) conjugated phalloidin (Invitrogen, Marseille, France). Images were obtained with a Nikon C1si confocal microscope using a 60X Plan Apo objectives. Images were processed for z-projection (maximum intensity), brightness and contrast regulation using ImageJ 1.45 (NIH, Bethesda, USA).

While cells transfected with p.Trp128Cys and p.Pro233Leu showed a disorganized cytoskeleton, those expressing p.Val328Met had alterations within cells but not in proximity of the membrane. In these cells, there are no actin filaments in the cytoplasm and α-actinin1 is widespread. However, they colocalize perfectly in the apical regions at the cell-to-cell junctions.
This is particularly interesting if we consider that ,unlike mutations previously reported which reside in actin-binding domain (ABD) or calmodulin-like (CaM) domain, p.Val328Met hits the spacer rod domain, suggesting a novel mechanism for the pathogenesis of ACTN1-related macrothrombocytopenia that does not involve functional domain mutations.

4.6.2 *TUBB1* variants

We detected 4 *TUBB1* variants, including 3 missense (p.Gly109Glu, p.Gln191Pro, p.Val286Met) and one nonsense (p.Tyr55 *), in 7 probands. All these variants were confirmed by Sanger sequencing and had a "deleterious" prediction of pathogenicity (**Table 4**). At present, no family members are available for segregation study. However, we are generating constructs cloning *TUBB1* cDNA into mammalian expression vector pcDNA3.1 with a 5' Myc tag sequence on N-terminus. Once transfected the wild-type and mutant constructs in a immortalized fibroblast cell line, we will evaluate the incorporation of exogenous ß-tubulin at the microtubule level. What we expect to see is a colocalization of endogenous α-tubulin and ß-tubulin exogenous only in cells expressing the wild-type and not the mutated forms of *TUBB1*[25]. This would demonstrate the inability of mutated ß-tubulin to be incorporated into rising microtubules and so the effect of pathogenic variants.

5. CONCLUSIONS

The purpose of my PhD project was to develop functional assays to test the pathogenicity of variants obtained by NGS. These analysis are absolutely necessary to determine the pathogenicity of missense variants because, in contrast to nonsense or frameshift mutations, the pathogenic role of them is not obvious. For this reason, we developed a functional assay on *ACTN1*, a novel gene responsible for macrothrombocytopenia identified through NGS analysis. In particular, we performed immunofluorescence on human fibroblasts overexpressing wild-type and mutated forms of *ACTN1*, in order to discriminate pathogenetic missense variants from those having no effect on actin cytoskeletal structure. Thanks to functional studies, we confirmed ACTN1-RT diagnosis among 15 families of our cohort of patients, describing it as the fourth most frequent (6,3%) form of IT in Italy after Bernard-Soulier syndrome (13%), MYH9-Related Disease (12%) and ANKRD26-Related Thrombocytopenia (10%). The NGS analyses revealed also 4 potential pathogenetic variants of *TUBB1*, another cytoskeleton component, which variants could be tested functionally using the same strategy as for those affecting *ACTN1*.

In conclusion, NGS is a revolutionary technique that over the past few years has been providing significant breakthroughs in the knowledge of the molecular bases of Mendelian diseases, including ITs. Its application in the diagnostic process of IT would allow us to screen directly at least all the known IT genes. However, understanding the effects of the thousands variants identified remains a major problem. For this reason, it is fundamental to develop further functional studies for other genes.

6. BIBLIOGRAPHY

1. Patel SR, Hartwig JH, Italiano JE, Jr.: **The biogenesis of platelets from megakaryocyte proplatelets**. *J Clin Invest* 2005, **115**(12):3348-3354.
2. Savoia A: **Molecular basis of inherited thrombocytopenias**. *Clin Genet* 2016, **89**(2):154-162.
3. Kile BT: **The role of apoptosis in megakaryocytes and platelets**. *Br J Haematol* 2014, **165**(2):217-226.
4. Machlus KR, Italiano JE, Jr.: **The incredible journey: From megakaryocyte development to platelet formation**. *J Cell Biol* 2013, **201**(6):785-796.
5. Thon JN, Italiano JE: **Platelets: production, morphology and ultrastructure**. *Handb Exp Pharmacol* 2012(210):3-22.
6. Balduini CL, Savoia A: **Genetics of familial forms of thrombocytopenia**. *Hum Genet* 2012, **131**(12):1821-1832.
7. Pecci A, Balduini CL: **Lessons in platelet production from inherited thrombocytopenias**. *Br J Haematol* 2014, **165**(2):179-192.
8. Ballmaier M, Germeshausen M: **Congenital amegakaryocytic thrombocytopenia: clinical presentation, diagnosis, and treatment**. *Semin Thromb Hemost* 2011, **37**(6):673-681.
9. Thompson AA, Nguyen LT: **Amegakaryocytic thrombocytopenia and radio-ulnar synostosis are associated with HOXA11 mutation**. *Nat Genet* 2000, **26**(4):397-398.
10. Albers CA, Paul DS, Schulze H, Freson K, Stephens JC, Smethurst PA, Jolley JD, Cvejic A, Kostadima M, Bertone P *et al*: **Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome**. *Nat Genet* 2012, **44**(4):435-439, S431-432.
11. Klopocki E, Schulze H, Strauss G, Ott CE, Hall J, Trotier F, Fleischhauer S, Greenhalgh L, Newbury-Ecob RA, Neumann LM *et al*: **Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome**. *Am J Hum Genet* 2007, **80**(2):232-240.
12. Raslova H, Komura E, Le Couedic JP, Larbret F, Debili N, Feunteun J, Danos O, Albagli O, Vainchenker W, Favier R: **FLI1 monoallelic expression combined with its hemizygous loss underlies Paris-Trousseau/Jacobsen thrombopenia**. *J Clin Invest* 2004, **114**(1):77-84.
13. Millikan PD, Balamohan SM, Raskind WH, Kacena MA: **Inherited thrombocytopenia due to GATA-1 mutations**. *Semin Thromb Hemost* 2011, **37**(6):682-689.
14. Monteferrario D, Bolar NA, Marneth AE, Hebeda KM, Bergevoet SM, Veenstra H, Laros-van Gorkom BA, MacKenzie MA, Khandanpour C, Botezatu L *et al*: **A dominant-negative GFI1B mutation in the gray platelet syndrome**. *N Engl J Med* 2014, **370**(3):245-253.
15. Albers CA, Cvejic A, Favier R, Bouwmans EE, Alessi MC, Bertone P, Jordan G, Kettleborough RN, Kiddle G, Kostadima M *et al*: **Exome sequencing identifies NBEAL2 as the causative gene for gray platelet syndrome**. *Nat Genet* 2011, **43**(8):735-737.
16. Bottega R, Pecci A, De Candia E, Pujol-Moix N, Heller PG, Noris P, De Rocco D, Podda GM, Glembotsky AC, Cattaneo M *et al*: **Correlation between platelet phenotype and NBEAL2 genotype in patients with congenital thrombocytopenia and alpha-granule deficiency**. *Haematologica* 2013, **98**(6):868-874.
17. Song WJ, Sullivan MG, Legare RD, Hutchings S, Tan X, Kufrin D, Ratajczak J, Resende IC, Haworth C, Hock R *et al*: **Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia**. *Nat Genet* 1999, **23**(2):166-175.
18. Liew E, Owen C: **Familial myelodysplastic syndromes: a review of the literature**. *Haematologica* 2011, **96**(10):1536-1542.
19. Pippucci T, Savoia A, Perrotta S, Pujol-Moix N, Noris P, Castegnaro G, Pecci A, Gnan C, Punzo F, Marconi C *et al*: **Mutations in the 5' UTR of ANKRD26, the ankirin repeat domain 26 gene, cause an autosomal-dominant form of inherited thrombocytopenia, THC2**. *Am J Hum Genet* 2011, **88**(1):115-120.
20. Zhang MY, Churpek JE, Keel SB, Walsh T, Lee MK, Loeb KR, Gulsuner S, Pritchard CC, Sanchez-Bonilla M, Delrow JJ *et al*: **Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy**. *Nat Genet* 2015, **47**(2):180-185.

21. Noetzli L, Lo RW, Lee-Sherick AB, Callaghan M, Noris P, Savoia A, Rajpurkar M, Jones K, Gowan K, Balduini CL *et al*: **Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia**. *Nat Genet* 2015, **47**(5):535-538.

22. Berndt MC, Andrews RK: **Bernard-Soulier syndrome**. *Haematologica* 2011, **96**(3):355-359.

23. Savoia A, Pastore A, De Rocco D, Civaschi E, Di Stazio M, Bottega R, Melazzini F, Bozzi V, Pecci A, Magrin S *et al*: **Clinical and genetic aspects of Bernard-Soulier syndrome: searching for genotype/phenotype correlations**. *Haematologica* 2011, **96**(3):417-423.

24. Berrou E, Adam F, Lebret M, Fergelot P, Kauskot A, Coupry I, Jandrot-Perrus M, Nurden A, Favier R, Rosa JP *et al*: **Heterogeneity of platelet functional alterations in patients with filamin A mutations**. *Arterioscler Thromb Vasc Biol* 2013, **33**(1):e11-18.

25. Kunishima S, Kobayashi R, Itoh TJ, Hamaguchi M, Saito H: **Mutation of the beta1-tubulin gene associated with congenital macrothrombocytopenia affecting microtubule assembly**. *Blood* 2009, **113**(2):458-461.

26. Kunishima S, Okuno Y, Yoshida K, Shiraishi Y, Sanada M, Muramatsu H, Chiba K, Tanaka H, Miyazaki K, Sakai M *et al*: **ACTN1 mutations cause congenital macrothrombocytopenia**. *Am J Hum Genet* 2013, **92**(3):431-438.

27. Bottega R, Marconi C, Faleschini M, Baj G, Cagioni C, Pecci A, Pippucci T, Ramenghi U, Pardini S, Ngu L *et al*: **ACTN1-related thrombocytopenia: identification of novel families for phenotypic characterization**. *Blood* 2015, **125**(5):869-872.

28. Nurden AT, Pillois X, Wilcox DA: **Glanzmann thrombasthenia: state of the art and future directions**. *Semin Thromb Hemost* 2013, **39**(6):642-655.

29. Savoia A, De Rocco D, Panza E, Bozzi V, Scandellari R, Loffredo G, Mumford A, Heller PG, Noris P, De Groot MR *et al*: **Heavy chain myosin 9-related disease (MYH9 -RD): neutrophil inclusions of myosin-9 as a pathognomonic sign of the disorder**. *Thromb Haemost* 2010, **103**(4):826-832.

30. Morison IM, Cramer Borde EM, Cheesman EJ, Cheong PL, Holyoake AJ, Fichelson S, Weeks RJ, Lo A, Davies SM, Wilbanks SM *et al*: **A mutation of human cytochrome c enhances the intrinsic apoptotic pathway but causes only thrombocytopenia**. *Nat Genet* 2008, **40**(4):387-389.

31. Massaad MJ, Ramesh N, Geha RS: **Wiskott-Aldrich syndrome: a comprehensive review**. *Ann N Y Acad Sci* 2013, **1285**:26-43.

32. Sjoblom B, Salmazo A, Djinovic-Carugo K: **Alpha-actinin structure and regulation**. *Cell Mol Life Sci* 2008, **65**(17):2688-2701.

33. Schwer HD, Lecine P, Tiwari S, Italiano JE, Jr., Hartwig JH, Shivdasani RA: **A lineage-restricted and divergent beta-tubulin isoform is essential for the biogenesis, structure and function of blood platelets**. *Curr Biol* 2001, **11**(8):579-586.

34. Italiano JE, Jr., Lecine P, Shivdasani RA, Hartwig JH: **Blood platelets are assembled principally at the ends of proplatelet processes produced by differentiated megakaryocytes**. *J Cell Biol* 1999, **147**(6):1299-1312.

35. Hartwig JH, Italiano JE, Jr.: **Cytoskeletal mechanisms for platelet production**. *Blood Cells Mol Dis* 2006, **36**(2):99-103.

36. Kunishima S, Nishimura S, Suzuki H, Imaizumi M, Saito H: **TUBB1 mutation disrupting microtubule assembly impairs proplatelet formation and results in congenital macrothrombocytopenia**. *Eur J Haematol* 2014, **92**(4):276-282.

37. Balduini CL, Pecci A, Noris P: **Diagnosis and management of inherited thrombocytopenias**. *Semin Thromb Hemost* 2013, **39**(2):161-171.

38. Balduini CL, Cattaneo M, Fabris F, Gresele P, Iolascon A, Pulcinelli FM, Savoia A, Italian Gruppo di Studio delle P: **Inherited thrombocytopenias: a proposed diagnostic algorithm from the Italian Gruppo di Studio delle Piastrine**. *Haematologica* 2003, **88**(5):582-592.

39. Sanger F, Nicklen S, Coulson AR: **DNA sequencing with chain-terminating inhibitors**. *Proc Natl Acad Sci U S A* 1977, **74**(12):5463-5467.

40. Johnsen JM, Nickerson DA, Reiner AP: **Massively parallel sequencing: the new frontier of hematologic genomics**. *Blood* 2013, **122**(19):3268-3275.

41. Albers CA, Newbury-Ecob R, Ouwehand WH, Ghevaert C: **New insights into the genetic basis of TAR (thrombocytopenia-absent radii) syndrome**. *Curr Opin Genet Dev* 2013, **23**(3):316-323.

42. Manchev VT, Hilpert M, Berrou E, Elaib Z, Aouba A, Boukour S, Souquere S, Pierron G, Rameau P, Andrews R *et al*: **A new form of macrothrombocytopenia induced by a germ-line mutation in the PRKACG gene**. *Blood* 2014, **124**(16):2554-2563.

43. Levin C, Koren A, Pretorius E, Rosenberg N, Shenkman B, Hauschner H, Zalman L, Khayat M, Salama I, Elpeleg O *et al*: **Deleterious mutation in the FYB gene is associated with congenital autosomal recessive small-platelet thrombocytopenia**. *J Thromb Haemost* 2015, **13**(7):1285-1292.

44. Fletcher SJ, Johnson B, Lowe GC, Bem D, Drake S, Lordkipanidze M, Guiu IS, Dawood B, Rivera J, Simpson MA *et al*: **SLFN14 mutations underlie thrombocytopenia with excessive bleeding and platelet secretion defects**. *J Clin Invest* 2015, **125**(9):3600-3605.

45. Gueguen P, Rouault K, Chen JM, Raguenes O, Fichou Y, Hardy E, Gobin E, Pan-Petesch B, Kerbiriou M, Trouve P *et al*: **A missense mutation in the alpha-actinin 1 gene (ACTN1) is the cause of autosomal dominant macrothrombocytopenia in a large French family**. *PLoS One* 2013, **8**(9):e74728.