Genetic Studies of Familial Myeloproliferative Disorders

INAUGURALDISSERTATION

ZUR

ERLANGUNG DER WÜRDE EINES DOKTORS DER

PHILOSOPHIE

VORGELEGT DER

PHILOSOPHISCH-NATURWISSENSCHAFTLICHEN

FAKULTÄT DER UNIVERSITÄT BASEL

VON

KUN LIU

AUS BENXI, CHINA

BASEL, 2007

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Professor Radek Skoda
Professor Markus Affolter
Professor Andreas Papassotiropoulos
Professor Mike Hall

Basel, den 13th November 2007

Professor Dr. Hans-Peter Hauri
Dekan

To my parents

报谁
得言
三寸
春草
晖心

Acknowledgements

I am very grateful to my supervisor Prof Radek Skoda for giving me the opportunity to do my PhD study in his laboratory and for his patient instructions and constant encouragement throughout this course of research. I thank Dr. Robert Kralovics for his supervision and invaluable suggestions on my projects. I thank Prof Markus Affolter, Prof Andreas Papassotiropoulos and Prof Mike Hall for joining my thesis committee.

My thanks go to all the lab members during my PhD years: Teo Soon Siong for introducing me to the lab and his friendship in all these years; Ralph Tiedt for correcting my thesis and many help at work; Hui Hao-Shen, Franz Schaub, Pan Dejing, Li Sai for their kind help and company.

I specially thank my friends Wang Xuejuan, Liu Kenan, Zhou Haiyan and Zhang Xin for their kind support in my time of need, and I thank Philip Fung for the happy time.

My deepest appreciations belong to my parents, my sister and brother-in-law for their loving supports through all these years.

Finally I thank the Swiss National Foundation for supporting my study.

TABLE OF CONTENTS

SUMMARY

Genetic studies using families have successfully identified many disease genes causing Mendelian diseases. Familial myeloproliferative disorders (MPD) offer interesting opportunities to identify disease genes involved in the thrombopoiesis and erythropoiesis, and some undiscovered genetic components might also contribute to the etiology of the sporadic MPD.

Hereditary thrombocythemia (HT) is an autosomal dominant disorder with clinical features resembling sporadic essential thrombocythemia. HT families share similar clinical symptoms caused by heterogeneous genetic alterations. Inherited germ-line mutations in the thrombopoietin (TPO) gene and its receptor MPL have been found causing thrombocytosis in a number of HT families. Five reported mutations in the thrombopoietin gene are all located in the 5 prime untranslated region (5'UTR) and cause overproduction of Tpo protein by the same mechanism: increased translation efficiency for the mutant mRNAs. One mutation identified in the *MPL* gene is located at the transmembrane domain and results in a hyperactive receptor, thereby leading to thrombocytosis. All these germ-line mutations have not been found in sporadic patients and are only responsible for the etiology of some HT families, indicating that the occurrence of these germ-line mutations is a rare event. The disease-causing genes for many HT families remain unknown. Identifying genetic lesions in these families will increase our knowledge of the physiology of thrombopoiesis and some of these unknown genetic components may contribute to the pathogenesis in sporadic MPD patients.

In the first part of the project for genetic studies of HT families, the *TPO* and *MPL* genes were analyzed by genomic DNA exon sequencing and linkage analysis. A splice donor mutation in the *TPO* gene was identified in a Polish family. This mutation was previously identified in a Dutch family and the reoccurrence of this rare mutation has not been reported to date. In order to determine whether the

mutation in these two families arose *de novo* or from a founder effect, haplotype analysis was performed to examine polymorphic DNA sequences in the vicinity of the mutation using microsatellites and single nucleotide polymorphism (SNP) in these two families. Six microsatellite markers on the affected allele showed different sizes in PCR products and 3 SNPs close to the mutation differed in their sequences between the two families. We therefore concluded that the mutation in these two families occurred de novo. The previously reported MPL mutation at the transmembrane domain of MPL protein was identified in one of the HT families studied here. Recently, 5 additional HT families were found carrying this mutation. We conducted haplotype analysis using microsatellite markers in the *MPL* gene locus for the 6 HT families. Four microsatellite markers surrounding the MPL mutation showed identical sizes in the PCR products on the affected allele, suggesting that the MPL mutation occurred from a single founder event. This may explain the high frequency of this mutation in HT families.

In a large US family with HT, where the *TPO* and *MPL* genes were excluded as disease causes, genome-wide linkage analysis was performed aiming to identify novel genetic component for the thrombocytosis phenotype. Two genetic regions with significant logarithm of odds (LOD) score values have been located using microsatellites and SNP chip arrays. Candidate gene sequencing revealed one novel polymorphism in the gelsolin gene, which encodes an actin-binding protein abundant in platelets. Gelsolin has multiple biological functions in addition to cytoskeletal actin modulation. Functional studies in cell proliferation assays and mouse bone marrow transplantation did not validate this polymorphism as an active disease causing mutation. Further studies on this polymorphism in platelet biogenesis are planned for the future. In addition, sequencing of all the candidate genes in the segregating regions is in progress.

In a second project, genome-wide linkage analyses were performed using microsatellites and SNP chip arrays in a family with secondary polycythemia inherited in an autosomal recessive mode. Both parametric and nonparametric

linkage analysis were conducted for this family. Five genetic regions were found linked to the disease phenotype. A few candidate genes were sequenced and studied, however no genetic variation was found so far. Additionally, no mutations were found in several genes involved in erythropoiesis and oxygen sensing pathway. Burst forming units-Erythroid cultures in hypoxia condition showed high expression of the *EPO* gene in 3 out of 4 affected family members, suggesting a potential unknown defect in the oxygen-sensing pathway.

GENERAL INTRODUCTION

Hematopoiesis

Hematopoiesis is the process of blood cell formation that lasts the entire lifetime. In humans, hematopoiesis occurs initially in the yolk sac of an embryo at the first few weeks and then moves to fetal liver and spleen until 6 to 7 months, when bone marrow develops into the main site of blood production and remains the major source of new blood cells throughout normal life.

One type of self-renewable and pluripotent hematopoietic stem cell (HSC) gives rise to all cell types in blood.[1] Initial differentiation of HSC is along two major pathways: myeloid or lymphoid. Common myeloid precursors develop into multiple cell types including erythrocytes, megakaryocytes (sources of platelets), granulocytes, and macrophages.[2] Common lymphoid progenitors give rise to B and T lymphocytes and natural killer cells.[3] Each cell type has specific functions and a distinct development procedure, which is regulated tightly through interactions between progenitor cells and various growth factors.

Megakaryopoiesis and Thrombopoiesis

Megakaryocytes are giant cells in bone marrow with a single, mutilobulated, polyploidy nucleus. Like other hematopoietic cell types, mature megakaryocytes are derived from pluripotent HSCs through a process called megakaryopoiesis. The major function of the megakaryocytes is to generate platelets in peripheral blood (thrombopoiesis). Platelets are anucleate cells formed by fragmentation of megakaryocyte cytoplasma and play important roles in hemostasis and thrombosis of peripheral blood vasculature. The regulation of megakaryopoiesis involves numerous cytokines, transcription factors and their target genes.

Thrombopoietin and thrombopoietin receptor c-MPL

The key regulatory cytokine in megakayopoiesis is thrombopoietin (TPO), which binds to its receptor c-MPL on megakaryocyte surfaces and supports the entire

process of megakaryocyte development, maturation and platelet production.[4] Knocking out of either the *TPO* gene or the *c-MPL* gene leads to 85% reduction of platelet counts in mice,[5,6] highlighting the central role of this ligand-receptor reaction in megakaryopoiesis.

Although being sought for long, the *TPO* gene was not identified molecularly until its receptor *c-MPL* gene was discovered from a study on the murine myeloproliferative leukemia virus (MPLV).[7] A novel virus oncogene *v-MPL* encoding part of an unknown member of the cytokine receptor superfamily was identified[8] and homolog searching in human and mouse led to cloning of the corresponding cellular protooncogene *c-MPL*.[9-11] Using c-Mpl as an essential tool, the gene encoding its ligand Tpo was soon cloned by several groups simultaneously.[12-15]

Human *TPO* gene is located on chromosome 3q27[16] with 5 coding exons and 2 upstream noncoding exons,[17-19] which result in a long 5 prime untranslated region (5'UTR) of *TPO* mRNA with additional 7 upstream initiation codons. These upstream start codons can bind to 40S ribosomal subunits and initiate premature translation of upstream open reading frames (uORF), which encode short and functionless polypeptides. These translations thereby suppress the binding of ribosome to the physiological start codon.[20,21] Five different mutations in the 5'UTR of the *TPO* gene have been described in families with hereditary thrombocythemia that cause increased efficiency of mRNA translation of TPO protein.[22-27] These types of uORF only exist in 10% of mRNA transcripts in human, but are often found in highly regulated genes.[28] The *TPO* cDNA predicts a polypeptide of 353 amino acids including a 21 amino acid secretary leader sequence.[29] The TPO protein consists of 2 domains: the N-terminal and the C-terminal domains. The N-terminal portion of 154 amino acids has striking homology with erythropoietin (EPO) and represents the receptor-binding domain of hormone. This domain, like EPO and other members of the hematopoietic growth factor family, is expected to fold into an anti-parallel four-helix bundle protein oriented up-up-down-down with two overhand loops connecting the first

two and last two helices.[30] The C-terminal part of TPO protein bears no resemblance to any known proteins.

Liver is the main site to produce TPO though a tiny amount of it can also be generated in kidneys, bone marrow stromal cells and other organs. The production of TPO is constitutive and the regulation of TPO levels in circulation is dependant mainly on the amount of its receptor c-MPL that is specifically expressed on the surface of platelets. By binding to c-MPL, TPO is internalized and consumed.[31,32] TPO serum concentration was shown being correlated inversely with platelet count.[33,34] An alternative mechanism for regulating TPO is that *TPO* mRNA production, instead of being constitutive, is up-regulated by low platelets events. Increased *TPO* mRNA in bone marrow was observed in mice made severely thrombocytopenic.[35] Although TPO production in liver was not altered in response to changes in platelet counts in peripheral blood, hepatic *TPO* mRNA levels increased in the presence of the inflammatory mediator interleukin-6 (IL-6).[36] Thus, the regulation of TPO production and levels may be more complex than only one mechanism.

The gene encoding the TPO receptor, c-MPL, is located on human chromosome 1p34 consisting of 12 exons.[37] Unlike the *TPO* gene, the *c-MPL* has a relatively short 5'UTR. The transcript starts sites are located between 4 and 45 bp upstream from the first ATG codon.[37] The promoter of *c-MPL* has consensus binding sequences for Ets and GATA transcription factors, which are vital for regulating many megakaryocytic specific genes. Variation in splicing results in 4 distinct *MPL* mRNA species in human. The predominant form (P-form) encodes the full-length protein. MPL-K (K-form) is due to a read through beyond the exon10 splice donor site, resulting a protein different from native MPL since the ninth cytoplasmic amino acid with additional 66 amino acids encoded by intron 10.[9] The third variant, MPL-del, arises from alternative splicing between exon 8 and 9 and encodes a protein with an inframe deletion of 24 amino acids.[38] The only variant found in both human and mouse is due to elimination of exon10 and thus the juxtamembrane WSXWS motif, the transmembrane domain and the initial cytoplasmic domain.[9,11,37] Multiple splice forms of the receptor have been

described as altered in their biological functions, and could exert regulatory effects on the receptors. The isoform c-MPL without exon 10 was found able to promote degradation of the P-form receptor rapidly when coexpressing both receptors.[39] However, physiological functions in the regulation of TPO signaling have not yet been demonstrated for the various isoforms.

As a member of the type I hematopoietic growth factor receptor family, c-MPL is a homodimer of a single receptor which is composed of a cytokine receptor motif, a transmembrane domain and an intracellular domain containing short sequences that bind intracellular kinases and other signal-transducing molecules. Upon ligand binding, c-MPL receptor undergoes conformational changes by bringing the cytoplasmic domains near to each other to initiate many biochemical activities. Members of the Janus kinase (JAK) family bind the cytoplasmic domains of c-MPL constitutively even in the inactive state. Upon TPO binding, JAK kinases, predominately JAK2, phosphorylate tyrosine residues in the receptor itself as well as downstream signal transducers and activators of transcription (STATs), phosphoinositide-3 kinase (PI3K), and the mitogen-activated protein kinases (MAPKs) to promote cell survival and proliferation. Meanwhile JAK2 also activates molecles that limit cell signaling such as the SH2-domain-containing protein tyrosine phosphatase 1 (SHP1), SH2-containing inositol phosphatase 1 (SHIP1) and suppressors of cytokine signaling (SOCSs). An acquired somatic mutation V617F in JAK2 was found in the majority of patients with chronic myeloproliferative disorders (MPD), especially in patients with polycythemia vera (PV).[40-45] This mutation is located at the pseudokinase domain of JAK2, which is the negative regulatory domain for the kinase activity.[46] Mutant JAK2 causes cytokine-independent activation of the downstream pathways including JAK-STAT and PI3K, MAPKs pathways, which lead to overproduction of platelets and erythrocytes in MPD patients.[41-45]

Numerous mutations including either homozygous or heterozygous, missense or nonsense mutations in the *c-MPL* gene were identified in congenital amegakaryocytic thrombocytopenia, which lead to the loss of function of the receptor.[47-49] Patients with severe thrombocytopenia carrying homozygous

mutant alleles can develop aplastic anemia due to stem cell exhaustion. The only treatment for the disease is stem cell transplantation.[50] In addition to loss-of-function, an activating missense mutation in the transmembrane domain of c-MPL leads to TPO-independent signaling activation, which is the cause for hereditary thrombocytopenia in several families.[51,52] Interestingly, this mutation was first discovered by random mutagenesis of mouse *c-MPL*.[53] Mutations in the juxtamembrane domain of c-MPL have been found in patients with MPD, in particular idiopathic myelofibrosis (IMF) and ET.[54,55] These findings support the concept that the membrane-proximal and transmembrane domains are constitutively active but blocked by the membrane-distal domains of the receptor. TPO binding relieves the blocking and therefore activates TPO/c-MPL signaling transduction.

Other growth factors in megakaryopoiesis

Besides the major regulator TPO, many other growth factors are involved in megakaryocyte growth, such as granulocyte-macrophage colony-stimulating factor (GM-CSF), IL-3, IL-6, IL-11, stem cell factor (SCF), FLT ligand, fibroblast growth factor (FGF), and EPO. These factors either stimulate megakaryocyte growth alone or cooperate with TPO. Unlike TPO, which supports the entire process of megakaryopoiesis, these cytokines play roles mostly in the early stage of megakaryocyte lineage development. Several cytokines are known to inhibit megakaryocyte development, such as IL-4,[56] transforming growth factor-β1[57] and Src kinase inhibitors.[58,59]

Transcription factors regulating megakaryopoiesis

Several transcription factors, including GATA-1, acute myeloid leukemia/ runt-related transcription factor 1 (AML/RUNX1) and nuclear factor-erythroid 2 (NF-E2), have been shown to play an important role in megakaryocyte development by activating megakaryocyte-specific genes and repressing genes for other cell lineages. GATA-1 recognizes and binds the sequence (A/T)GATA(A/G) in the cis-regulatory elements of many lineage-restricted genes. It interacts with the

cofactor friend of GATA-1 (FOG-1) and plays an essential role in early stage of megakaryocyte development, where it is involved in lineage commitment of megakaryocytes as well as erythrocytes. GATA-1 also functions later in megakaryocyte development and proliferation.[60] The megakaryocyte-specific knockout of GATA-1 results in thrombocytopenia and increased number of immature megakaryocytes with small sizes and decreased polyploidization.[61] Located on the X chromosome in humans, missense mutations of GATA-1 lead to severe congenital X-linked thrombocytopenia due to disruption between GATA-1 and FOG resulting in arrest of megakaryocyte maturation.[62,63]

RUNX-1 is an important transcription factor for the development of all hematopoietic lineages. It forms a complex with the core binding factor β subunit, which binds the N-terminal domain of GATA-1 and enables the programming of megakaryocyte lineage commitment.[64] Chromosomal translocations in the RUNX-1 gene are frequently found in leukemia. Germ-line mutations of RUNX-1 resulting in monoallelic loss of RUNX complexes are identified in families with autosomal dominant familial platelet disorder with multiple platelet defects, reduced c-MPL and predisposition to AML.[65]

The transcription factor NF-E2 is a heterodimeric leucine zipper transcription factor expressed in the erythroid and megakaryocytic lineage as well as in mast cells. It has been shown that NF-E2 controls terminal megakaryocyte maturation, proplatelet formation and platelet release by regulating target genes such as β-tubulin, thromboxane synthase, Rab27b and possibly some as yet unknown target genes.[66-69] NF-E2 null mice have neonatal lethality due to severe thrombocytopenia with arrest of megakaryocyte maturation, disorganized internal membranes and reduced granule numbers in the cells.[70] Compound homozygous mutation of NF-E2 in mice result in profound impairment of megakaryopoiesis.[71]

Platelet biogenesis

The terminal phase of megakaryocyte maturation is to generate platelets, which is a complex process orchestrated by numerous transcription factors, signaling

molecules and cytoskeletal elements. In order to tailor their cytoplasm and membrane systems for platelet production, megakaryocytes enlarge themselves by multiple rounds of endomitosis, a process that amplifies the DNA contents as much as 64 fold without any cell division.[72] During endomitosis, mitotic spindles assemble but fail to separate. Nuclear envelopes break down and reform to generate a polyploid, multilobed nucleus with up to 128n DNA content in mature megakaryocytes.[73,74] Besides DNA content expansion, internal membrane systems, granules and organelles are assembled in bulk during megakaryocyte development. An expansive and interconnected membrane network of cisternae and tubules called the demarcation membrane system (DMS) is formed and serves as a membrane reservoir for the formation of proplatelets, the precursor of platelets. The initial model of platelets coming from fragmentation of megakaryocyte extensions was brought up a century ago,[75] but little was known about the details of platelet assembly. The cloning of *TPO* and its receptor *c-MPL* has allowed major advances in the study of thrombopiesis.[76] With stimulation of TPO megakaryocytes cultured in vitro become a useful system to study the platelet formation process. It has been shown that megakaryocytes cultured in the presence of TPO extend numerous proplatelets, from which platelets are shed. Both proplatelets and platelets produced in vitro show similar structure and function with those generated in vivo.[77,78] The proplatelet formation process normally starts from one point on the megakaryocyte where pseudopodia form and continue to elongate until they become narrowed into proplatelets. A proplatelet appears as a thin cytoplasmic string containing multiple bulges similar in size to a platelet. Megakaryocytes continue to generate proplatelets from the original site and spread this event throughout the rest of the cell until the megakaryocyte cytoplasm is totally transformed into a network of interconnected proplatelets.[79] Platelets are assembled and released from the swellings at the proplatelet ends, but the details involved in this event have not been characterized.

Microtubules are the major structural component that drives the elongation of proplatelets, which were shown to be filled with microtubules bundles.[69]

Transgenic mice lacking β1-tubulin, the most abundant platelet β-tubulin isoform, develop thrombocytopenia and have spherocytic circulating platelets.[80] A double-nucleotide mutation resulting in the substitution of a conserved glutamine with a proline (Q43P) has been described in the β1-tubulin gene.[81] Individuals carrying a heterozygous Q43P mutation showed reduced expression of β1-tubulin in platelets, which were enlarged and spherocytic due to defects in the microtubule marginal band. The Q43P mutation is present in about 10% of the normal population and has a protective effect against cardiovascular disease.[81]

Several disorders of inappropriate platelet production have been identified harboring genetic mutations within genes that are active during the process of platelet biogenesis. For example, mutations in genes encoding glycoproteins have been identified in Bernard-Soulier syndrome, which is an autosomal dominant disorder characterized by macrothrombocytopenia, increased bleeding time and impaired platelet agglutination. Since glycoprotein complexes are linked to the membrane skeleton, it is believed that the genetic defects may alter normal cytoskeletal dynamics during platelet formation.[82] In a series of MYH9-related disorders, numerous mutations have been found in the MYH9 gene, which encodes nonmuscle myosin heavy chain IIA, the key myosin isoform expressed in platelets. The thrombocytopenia phenotype in MYH9-related disorders is thought to be due to defective platelet production as well because both megakaryocyte numbers and platelet clearance are normal.[83] It has been shown that in MYH9-related disorders, a greater amount of myosin is associated with the actin cytoskeleton in resting platelets and upon activation mutant platelets have altered cytoskeletal dynamics.[84] Although these disorders are very rare in the general population, identifications of the precise genetic lesions for such disorders give further insights into the mechanisms of platelet formation. Recently, mice with MYH9 knockout in megakaryocytes were generated and displayed macrothrombocytopenia with strong increase in bleeding time and absence of clot retraction.[85]

Erythropoiesis

Erythrocytes are non-nucleated cells without any organelles or RNA, whose major function is to carry oxygen to the tissues and bring carbon dioxide back to the lung. It is a biconcave-like disk in shape and flexible to pass through the microcirculation that is even narrower than the cell's diameter. The major component of erythrocytes is a highly specialized protein, hemoglobin, which loads and unloads oxygen. The average lifespan of an erythrocyte in circulation is about 120 days, and each day around 10^{12} erythrocytes need to be made to maintain the homeostasis in cell number by the complex and well-regulated process of erythropoiesis.

Erythropoietin and Erythropoietin receptor

Erythropoietin (EPO) is the major cytokine involved in the control of erythropoiesis. The receptor for EPO (EPOR) is increasingly expressed with the development of erythroid progenitors and the interaction with its ligand EPO activates erythrocytes differentiation and proliferation. Knockout of *EPO* or *EPOR* in mice leads to lethal anemia at day 12.5 of embryonic life, indicating the crucial role of the ligand and receptor interaction that is indispensable for erythropoiesis in vivo. However a normal number of erythroid progenitors is present in the fetal liver, suggesting that EPO and its receptor are not required for the commitment of hematopoietic stem cells to the erythroid lineage.[86]

Although the idea of a hormonal substance existing in serum for regulation of erythropoiesis[87] was long known, purification of EPO was extremely difficult due to its low concentration in serum.[88] Instead, Goldwasser *et al* purified EPO from human urine successfully.[89] The pure urinary EPO enabled the identification of amino acid sequences for the protein and subsequently the isolation of the human *EPO* gene.[90,91] Human EPO is an acidic glycoprotein with a molecular weight of 30.4 kDa that is secreted mainly from kidney and fetal liver.

The human *EPO* gene is located at chromosome 7q11-q22 with 5 exons encoding a 193 amino acid protein for the prohormone.[91-94] 27 amino acids are cleaved before secretion,[95] and the C-terminal arginine that is expected from the

mRNA sequence is absent in the circulating EPO. Therefore the peptide core of human mature EPO is composed of 165 amino acids, which form two bisulphide bridges. *EPO* gene expression is regulated by several transcription factors. The promoter region contains binding sites for GATA and nuclear factor κB (NF-κB).[96,97] GATA-4 is assumed to recruit chromatin-modifying activity and promote EPO expression,[98] while GATA-2 and NF-κB are thought to be responsible for the inhibition of *EPO* gene expression in inflammatory diseases.[99-101] Another important regulator for EPO expression is hypoxia-inducible transcription factor (HIF), which binds to the hypoxia response elements (HRE) in the 3' enhancer region of the *EPO* gene. HIF and HRE not only regulate EPO expression but also many other hypoxia inducible genes. More than 100 genes have been identified that are regulated by HIF-binding to HRE, including the vascular endothelial growth factor (VEGF), the glucose transporters and a few glycolytic enzymes.[102-104]

EPO was found binding to a transmembrane receptor of its target cells, EPO receptor (EPOR),[105,106] which was identified as a member of cytokine class I receptor family characterized by an extracellular N-terminal domain, a single hydrophobic transmembrane segment and a cytosolic domain without enzymatic activity.[107,108] The human *EPOR* gene is located on chromosome 19 with 12 exons encoding a 484-amino acid glycoprotein. The inactive EPOR is believed to be in an unliganded, dimeric state. Upon EPO binding, two EPOR molecules connect tightly to each other and undergo a conformational change,[109-111] Two JAK2 molecules associated with the cytoplasmic region of EPOR are activated and thereby several tyrosine residues of the EPOR are phosphorylated to provide docking sites for signaling molecules including PI-3K/Akt, STAT5, MAP kinase and protein kinase C. De-phosphorylation of JAK2 by the hematopoietic cell phosphatase terminates the effect of EPO. *EPOR* gene mutations resulting in the trunction of the receptor at its C-terminal lead to erythrocytosis.[112] The trunctated EPOR contains the binding site for SHP-1 phosphatase, which is a negative regulator of EPOR signaling.[112] SHP-1 dephosphalates a number of cytoplasmic substrates such as JAK2 and STAT5 to terminate the proliferative

signal. The lack of SHP-1 binding results in prolonged STAT5 and JAK2 activation observed in cell lines expressing truncated EPORs.[112]

Hypoxia induction of EPO

The primary function of circulating EPO is to maintain the hemoglobin concentration in the normal range. However the levels of hemoglobin or erythrocytes do not directly regulate the concentration of circulating EPO. Instead, the controlling variable is the tissue oxygen pressure (pO_2), which is dependent on the hemoglobin concentration, the arterial pO_2, the oxygen affinity of the hemoglobin and the rate of blood flow. In kidney, the major organ secreting EPO, pO_2 is hardly affected by the rate of blood flow as the renal oxygen consumption changes in proportion with the glomerular filtration rate.[113] The molecular mechanism of oxygen sensing became clearer since the identification of HIF-1, which is the major transcription factor controlling *EPO* gene transcription.[114,115] HIF-1 is a heterodimeric protein containing a α subunit and a β subunit. Both HIF-1α and HIF-1β are constitutively translated, but HIF-1α is not detectable in normoxic cells due to ubiquitin-mediated degradation of the protein in the presence of oxygen.[116,117] A critical residue of proline in the C-terminal part of HIF-1α is hydroxylated, which is catalysed by specific prolyl hydroxylases (PHD1, 2 and 3).[118-122] Hydoxylated HIF-1α combines with the von Hippel-Lindau tumor suppressor (VHL) to form a protein complex called E3 ligase that polyubiquitinate HIF-1α protein, [117,123,124] and then it undergoes proteasomal degradation rapidly in the condition of normoxia.[125] The PHDs are oxygen sensors because their activities depend on the availability of oxygen. Under hypoxia, PHDs' activity is greatly decreased and the proline residues are no longer hydroxylated, therefore HIF-1α can accumulate to activate the *EPO* gene and many other genes' expression. The transcriptional activity of HIFs is also regulated by another oxygen-liable hydroxylation at an asparagine residue catalysed by a HIF-specific asparaginyl hydroxylase that is termed factor-inhibiting HIF-1 (FIH-1).[126,127] Upon hydroxylation, binding of the transcriptional co-activator CAMP response element-binding protein (CREB)-binding protein

(CBP)/p300 to the C-terminal transactivation domain of HIF-1α is prevented, thus the transcriptional activity of HIF-1 is suppressed.[128] Like the PHDs, FIH-1 also requires oxygen for its function, together these hydroxylases serve as cellular oxygen sensors.

Two isoforms of HIF-α have been identified, namely HIF-2α and HIF-3α.[129-132] Both isoforms are oxygen-dependent and can form the heterodimer with HIF-1β in hypoxia but are different with respect to their tissue-specific mRNA expression pattern.[133] HIF-2α is expressed in renal fibroblasts while tubular cells express HIF-1α.[134] Recent studies indicate that HIF-2α is the primary transcription factor inducing EPO gene expression.[135] In contrast to HIF-1α and HIF-2α, HIF-3α is lacking a transcriptional activation domain and can suppress the expression of hypoxia inducible genes.[136]

Genetic defects in any components in the hypoxia sensing pathway could lead to overproduction of EPO protein and thereafter erythrocytosis. An arginine to tryptophon substitution at amino acid position 200 of VHL protein causes Chuvash polycythemia, which is an inherited endemic polycythemia occurred mainly in patients from Chuvashia.[137] The mutation in VHL protein impairs the interaction of VHL and HIF-1α, reducing the degradation rate of HIF-1α thus leads to increased expression of EPO and many other downstream target genes. A germ-line mutation in PHD2 was also found in a family with moderate erythrocytosis.[138]

Familial Myeloproliferative disorders

Myeloproliferative disorders (MPDs) are a group of clonal hematological malignancies characterized by aberrant proliferation of one or more myeloid lineages, which contains three subtypes: polycythemia vera (PV), essential thrombocythemia (ET) and idiopathic myelofirosis (IMF). Though the etiology for MPD remained unclear for long, it was believed that acquired somatic mutations occurred in hematopoietic progenitors or stem cells lead to over proliferation of the mutant cells thereby causing the disease. Recently-identified somatic

mutations in JAK2, which are prevalent in sporadic MPD patients,[40-45] are reshaping the molecular studies for MPD.

Familial forms of MPD, though showing similar symptoms that are sometimes indistinguishable from sporadic patients clinically, have different pathogenesis from sporadic MPD. Familial MPDs are polyclonal and caused by germ-line mutations inherited in the Mendelian way among family members. Searching for molecular alterations using linkage analysis has been very successful in some families with hereditary thrombocythemia (HT) and with primary familial and congenital polycythemia (PFCP). Five mutations in the *TPO* gene have been identified in HT families and all are located in the 5'UTR of the gene, causing increased translation efficiency of TPO protein.[21-27] One missense mutation was identified in the MPL transmembrane domain in 5 HT families, which leads to a hyperactive receptor activating downstream signals to stimulate cell proliferation.[51,52] Nine mutations have been described causing PFCP (reviewed in [139]). All these mutations lead to truncations of the cytoplasmic domain of the EPOR, which is a negative regulatory domain for the receptor. Truncated receptor therefore is hypersensitive to EPO and stimulates EPO-mediated signal transduction. None of these mutations in *TPO*, *MPL* or *EPOR* are found in sporadic MPD patients,[140-142] indicating these mutations are relatively rare events that only occur to certain families. In addition, these mutations only account for a small part of families with familial MPD, and the disease-causing gene(s) in many other families remain to be clarified.[143-145] It is reasonable to assume that some unknown germ-line mutations or polymorphisms could facilitate somatic mutations that lead to developing MPD in sporadic patients. Familial MPDs offer us a useful tool for searching for such molecular alterations using genetic approaches.

Identification of the germ-line mutations in familial MPDs has made the disease-causing genes such as *TPO* and *MPL* top candidate genes for sequencing in newly studied MPD families with unknown pathogenesis. In addition, studies on familial MPDs contribute to our understanding of the regulatory mechanism of thrombopoiesis and erythropoiesis.

Genetic approaches to identify human disease genes

Genetic studies have made remarkable contributions to the understanding of human diseases. A candidate gene approach led to discovery of many disease-causing genes based on previous knowledge and educated guesses. More often, when the pathogenesis for one disease is unknown or candidate genes have been excluded, a genome-wide screen for the disease gene is indispensable. However this genome-wide screen had rare successes for a long time because there was limited information on the human genome. With the development of the human and other genome projects, a vast range of resources of maps, clones, sequences, expression data and phenotypic data became available. Together with the technology for genotyping large amount of markers and mutation screening, identifying novel disease genes has become commonplace and is currently occurring frequently. The basic procedure for identifying human disease genes in family linkage studies is summarized in Figure 1, even though in practise there is no standard route and every step is affected by experimental results.

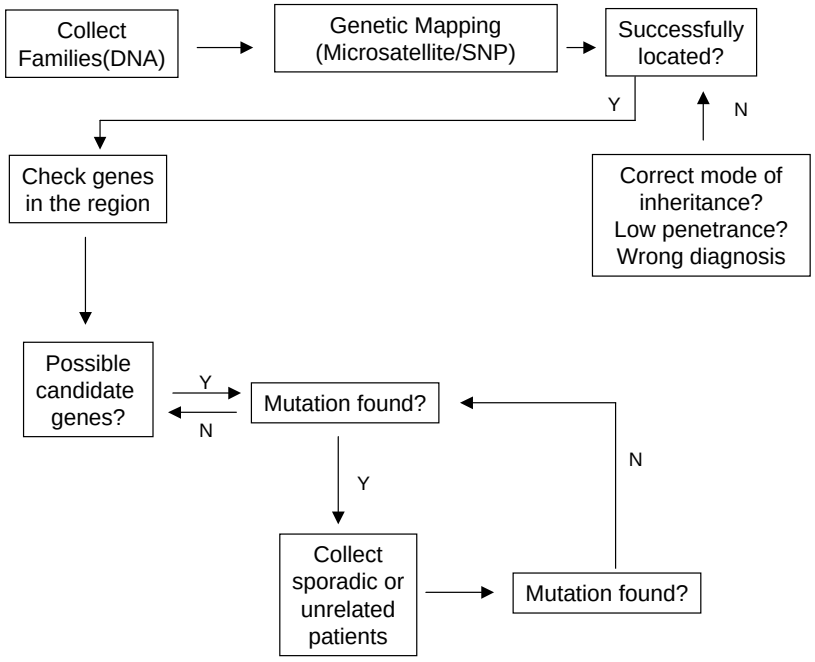

Figure 1 How to identify a human disease gene. Though no standard pathway to success, the key step is to arrive at certain plausible candidate gene, which can then be tested for functional alterations in affected individuals. Interplay between clinical work, laboratory benchwork and computer analysis is important for a success in identifying a disease gene.

Linkage studies

Many genes causing Mendelian diseases have been discovered by linkage studies using families with multiple affected individuals. With the early progress of the Human Genome Project, over ten thousand highly polymorphic genetic markers have been found and placed on framework maps,[146,147] which enable a higher resolution genome-wide mapping in family linkage studies.

The principle of linkage studies is based on the fact that recombination rarely occurs between two loci that are close to each other on the same chromosome. By searching for genetic markers segregating with the disease phenotype inherited through generations in a family, the disease gene can be tracked down

since it is close to the marker's location. Theoretically any Mendelian character that can distinguish the paternal and maternal allele in one individual can be used as a genetic marker. It is crucial, however, for genetic markers to be sufficiently polymorphic and densely located throughout the whole genome (<20cM). The first generation of genetic markers was restriction fragment length polymorphism (RFLP). The limitation of RFLPs is that they have only two alleles: the site is present or it is absent. Disease mapping using RFLPs could be frustrating because very often a key meiosis in a family turns out to be uninformative. Microsatellite markers were a great improvement since they have multiple alleles and high heterozygosity. Most meioses are informative within a family. PCR using microsatellites make linkage analysis fast and easy. Moreover, many compatible sets of microsatellite markers have been developed that can cover the whole genome with 400 markers. Microsatellite markers remained the most commonly used genetic markers for linkage studies until the recent launch of gene chip technology that can integrate thousands of SNPs in one single chip and genotype them at once. SNPs, as bi-allelic markers, have lower informativeness than microsatellite markers. However the large amount of SNPs that can genotype in chips make them very powerful and they are replacing the microsatellites rapidly in genetic linkage and association studies nowadays.

There are two types of linkage analyses: parametric and nonparametric linkage analysis. Parametric linkage analysis requires a precise genetic model including the mode of inheritance, the disease penetrance and the allele frequencies. As long as a valid model is available, parametric linkage should be applied since it provides the most powerful method to locate a disease gene. For some Mendelian characters, to ascertain a clear-cut pedigree is normally easy. For nonmendelian conditions, however, it is much less tractable. An incorrect genetic model can corrupt linkage analysis and mislead the follow-up studies. Therefore nonparametric linkage analysis is recommended in such cases. Nonparametric linkage analysis, also named model-free or allele-sharing analysis, ignores unaffected people, and looks for alleles or chromosomal segments that are shared by affected individuals within a family. Without making any assumptions

about the genetics of the disease, it has been used as the main tool for studying common nonmendelian diseases such as diabetes and schizophrenia. However nonparametric methods decrease the power of mapping, candidate regions defined by this method are usually large.

Linkage studies using families discovered many genes for Mendelian diseases, but have only limited success in finding genes for complex diseases such as diabetes, asthma and heart disease. With the rapid progress of genotyping technology, association studies become a powerful and preferable method for mapping complex diseases.

Association studies

In association studies, a large number of SNPs, either for the whole genome or in candidate linkage regions, are genotyped in a large group of unrelated people and a statistical analysis is performed to detect the co-occurrence of certain alleles and the disease. For example allele A is associated with disease D if people who have D also have A more (or maybe less) often than would be predicted from the individual frequencies of D and A in the population. Two different designs can be used in association studies: the population-based method that uses unrelated individuals and the family-based method that uses numerous families sharing the same disease. Genetic association studies have increasingly reported positive results, but many of them could not be replicated.[148] With the huge amount of information from genome-wide association studies, it becomes more and more challenging to interpret the data genetically and statistically.

RESULTS I:

Mutations of thrombopoietin and MPL gene in hereditary thrombocythemia

A de novo splice donor mutation in the thrombopoietin gene causes hereditary thrombocythemia in a Polish family

Kun Liu[1], Robert Kralovics[1], Zbigniew Rudzki[2], Barbara Grabowska[3], Andreas S. Buser[4], Ralph Tiedt[1], Patricia Frank[1], Krzysztof Okon[2], Anthonie P.C. van der Maas[5], Radek C. Skoda[1]

[1]Experimental Hematology, Department of Research, Basel University Hospital, Basel, Switzerland, [2]Department of Pathomorphology, Collegium Medicum, Jagiellonian University, Kraków, Poland, [3]Department of Hematology, Ludwik Rydygier Memorial District Hospital, Kraków, Poland, Clinical Hematology, [4]Basel University Hospital, Basel, Switzerland, and the [5]Department of Internal Medicine, Medical Centre Haaglanden, 2512 VA The Hague, The Netherlands

Correspondence: Radek C. Skoda, MD, Department of Research, Experimental Hematology, Basel University Hospital, Hebelstrasse 20, 4031 Basel, Switzerland, radek.skoda@unibas.ch

Running title: Splice donor mutation in THPO gene causes HT

Keywords: TPO, Hereditary thrombocythemia, de novo mutation, founder effect, MPL

Abstract

Background and Objectives. Hereditary thrombocythemia (HT) is an autosomal dominant disorder with clinical features resembling sporadic essential thrombocythemia (ET). Germ line mutations in HT families have been identified in the genes for thrombopoietin (*THPO*) and its receptor, *MPL*. Here we characterized a *THPO* gene mutation in a HT pedigree with 11 affected family members.

Design and Methods. Linkage analysis was performed and the *THPO* gene was sequenced. Thrombopoietin (TPO) serum concentrations were determined by ELISA. The mRNA and protein levels for MPL were assessed by real-time PCR and Western blotting, respectively. Haplotype analysis using microsatellites and single nucleotide polymorphisms (SNPs) were performed.

Results. We identified a G→C transversion in the splice donor of intron 3 of the *THPO* gene that co-segregated with thrombocytosis within the pedigree. We have previously described the identical mutation in a Dutch HT family. Haplotype analysis using microsatellites and SNPs surrounding the mutation provided no evidence for a founder effect and indicated that the mutations have arisen independently in the two families. MPL protein levels, but not mRNA levels, were low in platelets from affected family members. Bone marrow histology showed features compatible with those of ET, but the megakaryocytes were unusually compact, as assessed by planimetric analysis. Impaired microcirculation resulting in brief episodes of fainting and dizziness that responded well to aspirin were the predominant clinical features in a total of 23 affected family members studied. Compared to sporadic ET patients, familial patients have earlier onset ages but similar severity of symptoms.

Conclusions and interpretations. This is the fifth report of HT family caused by a mutation in the *THPO* gene, yet the first one sharing the same mutation with a family reported previously. Patients with overproduction of TPO have a mild phenotype not requiring cytoreductive treatment.

Introduction

Hereditary thrombocythemia (HT), also known as familial thrombocytosis or familial essential thrombocythemia, is an autosomal dominant disorder with clinical features resembling sporadic essential thrombocythemia (ET). HT is characterized by active proliferation of megakaryocytes and overproduction of platelets. The key regulators for platelet production are thrombopoietin (TPO) and its receptor, MPL. To date, four different germ line mutations in the thrombopoietin (*THPO)* gene have been identified and all of them alter the 5 prime untranslated region (5'-UTR) of the *THPO* mRNA, which contains open reading frames (ORF) that inhibit the translation of *THPO* mRNA.[21-27] The mutations remove the inhibitory ORFs and lead to increased translation of the *THPO* mRNA causing elevated TPO serum levels and overproduction of platelets.[22,24] A missense mutation in the transmembrane domain of *MPL* has been identified in a HT family.[51] This mutation generates a hyperactive MPL protein and results in excessive platelet production. Recently, mutations in the juxtamembrane domain of MPL have been found in patients with chronic myeloproliferative disorders (MPD), in particular idiopathic myelofibrosis (IMF) and ET,[54,55] but *THPO* mutations have not been detected in patients with sporadic ET.[140] In some HT families, both *THPO* and *MPL* genes can be excluded as the cause of thrombocytosis and thus, other as yet unknown genes can be involved in causing an identical phenotype.[143,144]

In this study, we analyzed a Polish family with HT and identified a G→C transversion at the intron 3 splice donor in *THPO* that co-segregated with the thrombocythemia phenotype. This mutation is identical with the mutation we previously described in a Dutch family with HT.[22] Here we present a comparison of the clinical and pathomorphological features of 23 affected family members with 17 sporadic ET patients published previously.[149]

Design and methods

Patients and clinical features

The proposita (PL09) was referred to the hematology clinic at the Ludwik Rydygier Memorial District Hospital in Kraków, Poland, in the year 2000 (age at diagnosis = 19 years) because of significant thrombocytosis (1'455 x 10^9/L) found in a routine blood test. At presentation the patient was asymptomatic and without physical signs. Abdominal ultrasound revealed minimal splenomegaly (124 mm in the long axis). The peripheral blood values were: platelets 1'032 x 10^9/L, white blood cells 7 x 10^9/L; red blood cells 5.1 x 10^{12}/L; hemoglobin 13.6 g/dl, hematocrit 39.7 %, MCV 78.5 fl; MCH 26.9 pg; MCHC 34.2 g/dl. No cause of reactive thrombocytosis was found and the histology of the bone marrow was compatible with the diagnosis of MPD other than chronic myeloid leukemia, most probably ET. Since the patient fulfilled the Polycythemia Vera Study Group (PVSG) and the World Health Organization (WHO) criteria for ET,[150-154] and her platelet levels on follow up constantly exceeded 1'000 x 10^9/L, treatment with hydroxyurea at a dose of 1 g/day was initiated and continued for two years. During the two years of hematological follow-up she had been suffering from episodes of brief transient unconsciousness, initially interpreted with the aid of electroencephalography as epileptic in origin, and treated with carbamazepine for one month. After the familial background of the disease became evident, hydroxyurea was stopped and treatment was changed to low-dose aspirin (75 mg/d). Currently, she is maintained on low-dose aspirin, her platelets are stabilized at levels around 800 x 10^9/L. The spleen is not palpable, and she does not manifest any other signs or symptoms of disease. At the end of 2001, thrombocytosis was diagnosed in her two sisters (PL07 and PL08). The older sister, PL07, suffered from a Raynauld phenomenon and brief episodes of fainting and dizziness. In addition, a persistent pain in her right elbow was noted, without any detectable local radiological and vascular abnormalities. She manifested mild splenomegaly (130 mm in the long axis on ultrasound). She was treated with low-dose aspirin and ticlopidine, and currently is asymptomatic. At presentation, PL08, the dizygotic twin of the proposita, complained of bilateral paresthesia in her fingers, and reported an episode of superficial vein thrombosis in her left hand. Treatment with low-

dose aspirin resulted in complete remission of paresthesia. Her platelets are stable at the levels below 700 x 10^9/L. Soon after, another young thrombocythemic female patient (PL04) treated in another institution for headaches, arterial hypertension and obesity, was identified as their great-grandparental cousin. Similar to the other family members, her symptoms responded to low-dose aspirin. The clinico-pathological picture found in several members of the youngest generation prompted a wide screening of their extended family, revealing altogether 11 affected family members.

Blood cells separation, DNA and RNA extraction

Blood cells were separated by standard protocols using Histopaque (Sigma, St. Louis, MO, USA) gradient centrifugation. Granulocytes and peripheral blood mononuclear cells were collected respectively. Platelets were collected using the Sepharose (Amersham Pharmacia Biotech AB, Uppsala, Sweden) gel filtration method.[155] DNA was extracted using a standard proteinase K (Promega, Madison, WI, USA)/phenol (Fluka Chemie AG, Buchs, Switzerland) extraction protocol. RNA isolation was carried out using the TRIfast reagent (peqLab Biotechnology GmbH, Erlangen, Germany).

Pathology of bone marrow

Trephine bone marrow biopsies of anterior superior iliac crest were obtained for the diagnostic purposes in five members of the family after their informed consent. Tissue cores were fixed in 4% buffered formaldehyde, decalcified in hydrochloric acid-based commercial solution (Shandon TBD-1 Rapid Decalcifier, Anatomical Pathology International, Runcorn, UK), embedded in paraffin and cut into 4 μm sections. The dewaxed slides were stained with hematoxylin and eosin (H&E), periodic acid-Schiff, Giemsa, and Gomori silver for reticulin fibers. Bone marrow fibrosis was quantitated in a scale ranging from 0 to +4.[156] Accessory immunostaining encompassed CD61 and CD34 (both from DakoCytomation, Glostrup, Denmark) used to highlight atypical and particularly small megakaryocytes, and to visualize blasts, respectively. Objective, computer-assisted analysis of megakaryocyte planimetric parameters was performed on the H&E sections as described previously.[157] Briefly, the high-power/high-resolution electronic images of representative megakaryocytes were transformed into two-color

bitmaps depicting the cytoplasmic and nuclear shapes. These were analyzed in respect to a series of standard planimetric parameters (linear sizes, areas, shape factors, etc) using a computer image analysis system Analysis pro v. 3.2 (Soft Imaging System GmbH, Münster, Germany). The results were compared to the analogous parameters characterizing representative megakaryocytes in 10 control trephines demonstrating normal marrows and 20 cases of classical sporadic essential thrombocythemia, diagnosed according to the WHO criteria.[152,153]

Quantitative PCR for PRV-1and MPL

Total RNA (2 μg) was reverse transcribed after random hexamer priming. The primers for ribosomal protein L19 (*RPL19*), *MPL* and polycythemia rubra vera-1 (*PRV-1*) were designed across exon-intron junctions. The primers for *RPL19* were GATGCCGGAAAAACACCTTG, TGGCTGTACCCTTCCGCTT, CCTATGCCCATGTGCCTGCCCTT (probe); for *PRV-1*: CCCCAGCAGACCCAGGA, TTGTCCCCTCCAGACAGCC, CCATAGACAAGCAGACTGGGCACCTCAA (probe). The probes were dual-labeled with 5'-6-carboxyfluorescein (FAM) and 3'-tetramethyl-6-carboxyrhodamine (TAMRA). The SYBR detection primers for *MPL* were AGCCCTGAGCCCGCC and TCCACTTCTTCACAGGTATCTGAGA. The ΔCT values were derived by subtracting the threshold cycle (C_T) values for PRV-1 and MPL from the C_T value for RPL19, which served as an internal control.[158] A non-affected family member PL15 was chosen as a calibrator for calculating the ΔΔCT values.[159,160] All reactions were run in duplicates using the ABI 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA).

EEC assay

The clonogenic cultures for erythropoietin-independent colony formation (EEC assay) were performed as previously described using commercial reagents Methocult H4531 (Stem Cell Technologies Inc, Vancouver, BC, Canada). [158]

Analysis of genetic linkage

DNAs were PCR amplified using dye-labeled primers for microsatellite markers. The PCR program includes at 94ºC for 15 sec, 55ºC for 15 sec, 72ºC for 30 sec for 10 cycles, 89ºC for 15 sec, 55º for 15 sec and 72ºC for 30 sec for 20 cycles. The PCR products were analyzed using ABI 3100 genetic analyzer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol. Genotypes were scored using the Genemapper software package version 3.5 (Applied Biosystems, Foster City, CA, USA) and linkage analysis was carried out with FASTLINK software package version 4.1p. Equal allele frequencies for the marker alleles were assumed. An autosomal dominant inheritance model with a 100% penetrance was used.

Genomic DNA sequencing

The entire coding region including intron/exon boundaries of the *THPO* gene was sequenced from PCR fragments, amplified from genomic DNA of the affected family member PL10. The primer sequences for PCR are shown in Supplementary Table 1. The PCR conditions were 95ºC for 2 min, 94ºC for 30 sec, 58ºC for 30 sec and 72ºC for 1 min for 35 cycles. Sequencing was performed on an Applied Biosystems 3700 DNA sequencer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocols.

RFLP analysis.

For co-segregation analysis, a 951 bp PCR fragment was amplified using the primers AGCCTAAGCCGCCTCCATG (exon 3, sense) and GGTGGCCAAGCTGAAGGTG (intron 5, antisense) from genomic DNA of all family members and digested with *Bsr* I restriction enzyme at 65ºC overnight. Fragments of 460 bp for the mutant allele and 359 bp for the normal allele were visualized by ethidium-bromide staining after agarose gel electrophoresis.

Haplotype analysis

To determine the founder effect, 6 microsatellite markers located in the vicinity of the *THPO* gene were chosen for linkage analysis. Sequences of the primers are provided in Supplementary Table 2. The PCR products were analyzed as described above. The

haplotypes were determined based on the segregation within the pedigrees and the sizes of the PCR products of the co-segregating microsatellite markers were compared between affected members of the two families. Ten SNPs located within the *THPO* gene (Supplementary Table 3) were selected from the dbSNP at the NCBI homepage (http://www.ncbi.nlm.gov/projects/snp/) and genotyped by sequencing.

Human TPO ELISA and immunoblot assay of MPL

TPO serum levels were measured using the TPO-Quantikine ELISA (enzyme-linked immunosorbent assay) kit (R&D Systems, Minneapolis, MN, USA) according to the manufacturer's protocol. MPL protein expression in platelets was determined by immunoblot analysis using the polyclonal rabbit antibody (CTP7) specific for the C-terminus of human MPL (kindly provided by Dr. Jerry L. Spivak, John Hopkins University, Baltimore). The membranes were re-probed using a monoclonal antibody against human CD61 (BD Biosciences, San Jose, CA, USA) serving as a loading control.

Results

The clinical features of 11 family members with thrombocytosis are summarized (Table 1). Thrombocytosis in many of the patients was detected in childhood or adolescence. Five of the 11 affected family members had symptoms that are potentially related to thrombocytosis, including hypertension, headaches, Raynauld phenomenon, limb paresthesia, venous thrombosis, transient ischemic attacks, miscarriage and Buerger disease. Most of these symptoms, except Buerger disease, were manageable using low-dose aspirin. In contrast, attempts to relieve the symptoms by cytoreductive therapy with hydroxyurea were ineffective. Patient PL13 died of thromboembolic complications of his Buerger disease at age 57.

Table 1 Summary of clinical data of 11 members of the Polish family with thrombocytosis.

UPN	Sex	Date of Birth	Date of Dx	Date of last follow up or DOD	Platelets 150-450 x10^9/L	WBC 3.5-10.0 x10^9/L (2003)	RBC 4.2-6.3 x10^{12}/L (2003)	Hemo-globin 12-14 F 14-18 M g/L (2003)	Spleno-megaly (last follow up)	Hepato-megaly (last follow up)	Thrombocytosis-associated symptoms	Important co-morbidity
PL02	F	1956	3/200 3	3/2003	545-560	5.9	4.5	131	na	na	none	not known
PL04	F	1986	8/2001	6/2006	595-1300	8.1	4.7	123	(+)	(-)	hypertension headaches	obesity
PL06	M	1950	3/2003	3/2006	408-420	6.5	5.0	145	(-)	(-)	none	none
PL07	F	1978	10/2001	9/2005	760-960	6.1	4.7	132	(+)	(-)	Raynauld phenomenon transient ischemic attacks miscarriage	persistent pain in the right elbow
PL08	F	1982	10/2001	92004	750-890	7.1	4.7	135	(-)	(-)	limb paresthesia venous thrombosis	none
PL09	F	1982	11/2000	12/2005	740-1340	6.7	4.1	127	(-)	(-)	transient ischemic attacks	
PL10	F	1991	3/2003	3/2003	960	10.6	5.2	142	na	na	none	
PL11	F	1922	3/2003	3/2003	510	7.7	5.0	150	na	na	na	
PL12	M	1947	10/1998	3/2003	550-560	5.3	5.5	156	(+)	na		melanoma
PL13	M	1948	3/2003	4/2003	910 - 1190	10.7	4.7	139	(+)	(-)	Buerger disease, died in 2005	viral hepatitis type B
PL14	M	1992	3/2003	3/2003	460	6.2	4.6	123	na	na	na	

UPN, unique patient number; Dx, diagnosis; DOD, date of death; WBC, white blood cells; RBC, red blood cells; lowest and highest values are given where available; Dx, diagnosis; na, data not available

Linkage analysis revealed co-segregation of thrombocytosis with two microsatellite markers (THPO1 and THPO2) located in close vicinity of the *THPO* locus with a logarithm of odds (LOD) score of 3.3 at theta = 0 (data not shown). Sequencing of the *THPO* coding region as well as the intron-exon boundaries revealed a G→C transversion in the splice donor of intron 3 (Figure 1A). This mutation alters a BsrI restriction site and produces a restriction fragment length polymorphism (RFLP), which was used to confirm

the co-segregation of the mutation within the pedigree (Figure 1B). We previously described the identical mutation in a Dutch family with thrombocytosis. [22] This mutation destroys the splice donor site in intron 3 and results in exon 3 skipping (Figure 1A). We have shown that the resulting shortened 5'-UTR leads to overproduction of TPO protein by a mechanism of increased efficiency of *THPO* mRNA translation.[22] The mutation was not found in a previously published series of 50 sporadic ET patients.[140] Here we screened for this mutation in additional 76 sporadic ET patients, but did not find any new case (data not shown).

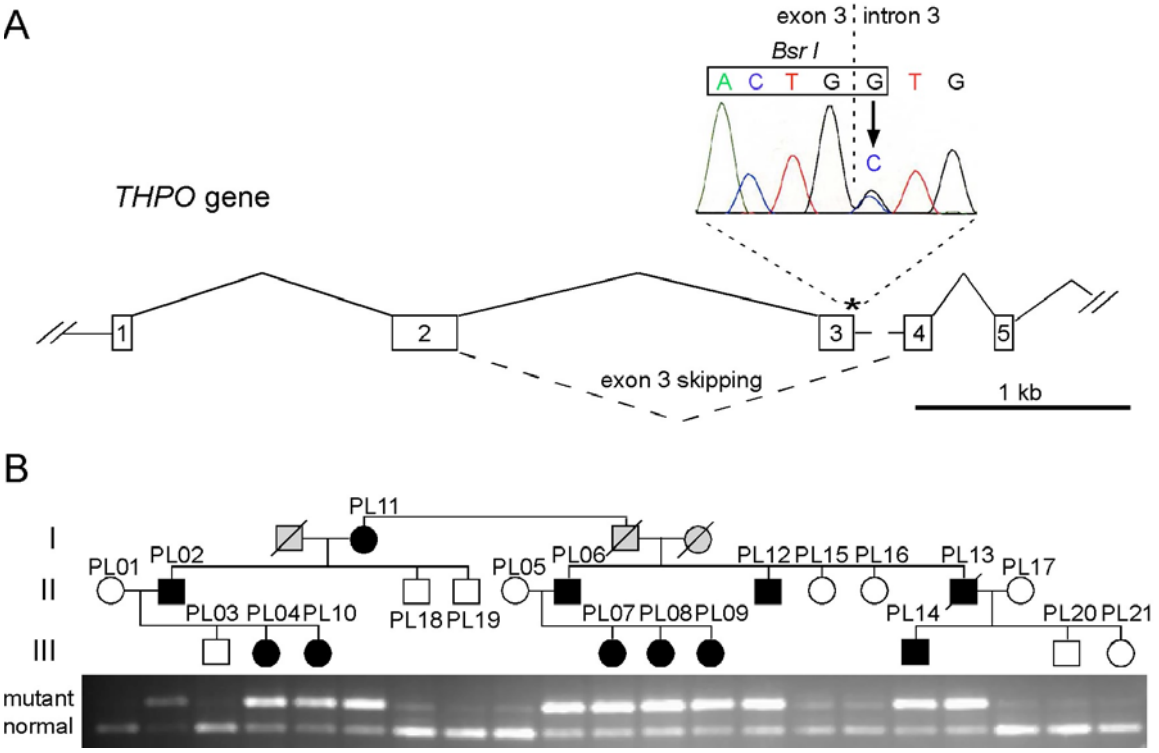

Figure 1 The *THPO* gene mutation. A) The sequencing chromatogram of the boundary between *THPO* exon 3 and intron 3 (dashed vertical line) from an affected individual is shown. Arrow points to the G→C transversion in the sequence. The recognition sequence for the *Bsr* I restriction endonuclease is boxed. This recognition sequence is destroyed by the G→C transversion. The *THPO* gene locus is shown below. The asterisk marks the position of the G→C transversion. Open boxes represent exons. Exons connected by solid lines represent normal splicing and dashed lines indicate expected consequence of the *THPO* mutation on splicing. **B)** Co-segregation of the *THPO* mutation and thrombocythemia within the pedigree. The *Bsr* I restriction fragment length polymorphism, caused by the presence or absence of the G→C transversion, was used to follow the inheritance of the *THPO* mutation. Individuals within the pedigree are positioned above the corresponding lanes.

To determine whether the mutation in the two families represents a founder effect or has independently arisen de novo, we examined polymorphic DNA sequences in the vicinity of the mutation. A founder effect, i.e. descent of both families from a common affected ancestor, is expected to result in sharing of allelic sequence polymorphisms in the vicinity of the *THPO* mutation in affected members from both families. First, we compared six microsatellite markers located between 4 kb to 40 kb from the *THPO* mutation, but all PCR products that represent the haplotype of the affected allele in the two families showed different sizes (Figure 2A), suggesting that the mutation occurred independently in these two families. Since the mutation rate of microsatellites is relatively high, in the range of 10^{-3} to 10^{-4} per locus per generation,[161] we cannot exclude that some of the differences could be due to the inaccuracy in the replication of repetitive elements. We therefore genotyped SNPs, which are genetically more stable and display a lower mutation rate (10^{-8} per generation).[162] By screening 10 SNPs located within the *THPO* gene we found that 3 informative SNPs, representing the haplotype of the co-segregating mutant allele, differed in their sequences between the two families (Figure 2B). One SNP (rs956732) is located only 150 nucleotides upstream of the mutation, while the other 2 SNPs (rs2280740 and rs10513797) are located 507 and 1'553 nucleotides downstream of the mutation, respectively (Figure 2B). Due to the very short physical distance between these SNPs and the G→C mutation, it is very unlikely that the differences in the sequence between these two families are due to recombination. Therefore, we conclude that the mutation in these two families occurred independently and *de novo*.

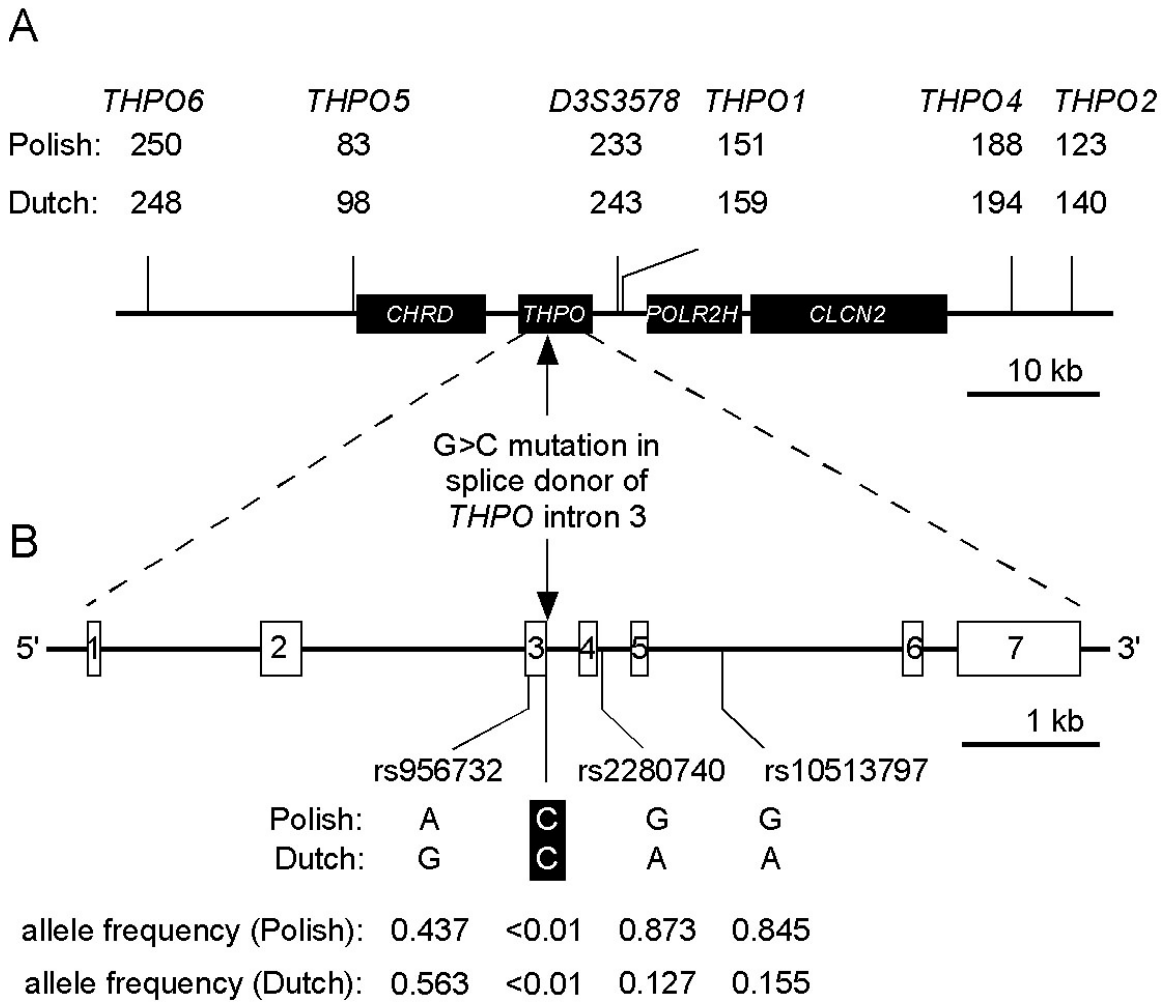

Figure 2 Haplotype analysis for the THPO locus in the Polish family and the Dutch family. A) The chromosomal locus containing the THPO gene is shown. Black boxes represent genes: THPO, thrombopoietin; CHRD, chordin; POLR2H, polymerase RNA II DNA directed polypeptide H; CLCN2, chloride channel 2. Microsatellite markers are shown above the locus, numbers indicate the sizes in nucleotides of the PCR products of the co-segregating mutated alleles in the Polish and Dutch families. Note that none of the allele sizes are identical in the two families. **B)** The THPO gene locus with the positions of three informative single nucleotide polymorphisms (SNPs) is shown. The sequences at each of the SNP positions are shown for the co-segregating mutated allele only. The allele frequencies for each SNP are listed below. Note that the sequences of the 3 SNPs located in the vicinity of the G→C transversion (black box with white letters) differ in the two families.

To explore how the *THPO* mutation affects the regulation of platelet production, we measured the TPO serum concentrations and MPL protein expression levels on platelets and compared them with the platelet counts in all family members (Figure 3A). Two affected family members showed highly elevated serum TPO levels (PL12 and PL13), the other 9 affected family members had only slightly elevated or normal TPO serum levels. TPO concentrations showed no clear correlations with the platelet counts (Figure 3A). The MPL protein expression levels were determined in platelet lysates and normalized ratios against CD61 were used to determine the relative MPL protein amount. Nine of 11 patients showed decreased expression of MPL protein amount compared to the normal. Interestingly the individual with the highest TPO serum level had the lowest MPL expression (PL12). There were significant differences in mean values for platelet counts ($p<0.001$), serum TPO concentration and MPL protein expression ($p<0.05$) between affected and non-affected family members (Figure 3B). The low amount of MPL protein was not due to decreased mRNA levels, as shown by real-time PCR. Rather, there was a slight, but non-significant increase in MPL mRNA in the affected individuals (Figure 3B). Interestingly PL12, who had the lowest MPL protein level, showed the highest *MPL* mRNA level (Figure 3A), confirming similar data obtained in the Dutch family with the same *THPO* mutation.[158] All 11 affected family members had normal levels of *PRV-1* mRNA in granulocytes and did not display growth of EECs (data not shown).

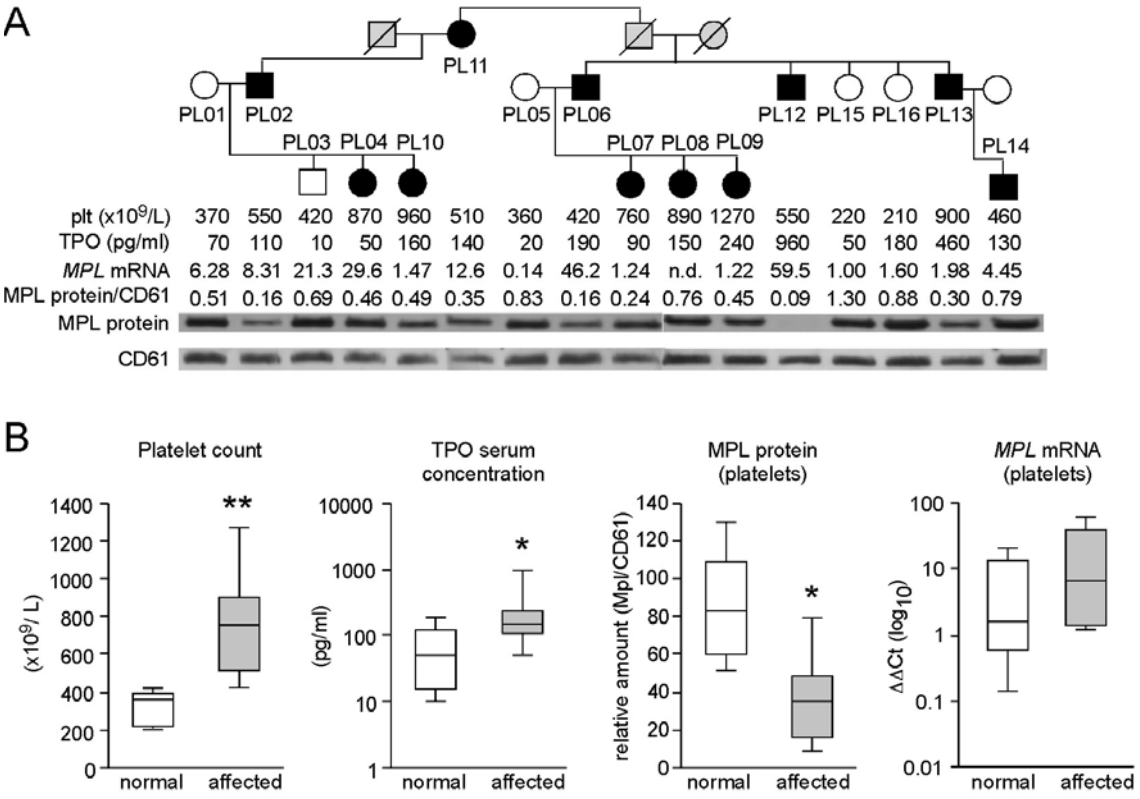

Figure 3 The correlations between the platelet count, TPO serum concentration and MPL expression. A) The Western blotting of MPL and CD61 protein in platelets are shown. The individuals within the pedigree are placed above the corresponding lanes. The platelet counts $x10^9$/l (plt), the TPO serum concentrations in pg/ml (TPO), the *MPL* mRNA expression determined by real time PCR, and the ratios of MPL protein against CD61 determined by densitometry are shown. **B)** Boxes represent the interquartile range that contains 50% of the values, the horizontal line in the box marks the median and bars indicate the range of values. P values are calculated by one-side t-test for independent samples (* P value < 0.05, ** P value < 0.001). The relative expressions of *MPL* mRNA (ΔΔCT values) are shown on a logarithmic scale. The values are relative to a non-affected family member PL15. The horizontal bar indicates the median.

Since in this family thrombocytosis is caused by a known mechanism, i.e. overproduction of TPO protein, it is interesting to compare the histopathology of bone marrow trephines from affected family members with trephines from sporadic ET patients. The results of the analysis of archival material from 5 affected members of the Polish family and 1 member of the Dutch family are summarized in Table 2. The megakaryocyte densities were comparable to the lower limits of the values encountered in chronic MPDs, but were much higher than the appropriate age norms.[163,164] The density of megakaryocytes in the individual cases did not correlate with the platelet count. The tendency for clustering was noticeable, but less then 5% of megakaryocytes participated in the cluster formation. Only rarely did the clusters contain more than 3 cells. There was no tendency for intrasinusoidal location of megakaryocytes. None of the patients displayed obvious atypia, but the megakaryocytes were strikingly compact when compared to normal megakaryocytes and particularly to the megakaryocytes of essential thrombocythemia (Figure 4). This impression was corroborated by the objective planimetric analysis of megakaryocytes, which demonstrated significant differences between these 3 diagnostic groups (ANOVA data not shown). Specifically, the cellular regularity may be quantified as the cellular compactness factor (CF), representing the ratio of the minimal diameter to the maximal diameter of an object, thus ranging from 1 (for a circle) to 0. On the post-hoc analysis, cellular CF of megakaryocytes in familial thrombocytosis (0,752 ± 0,102) significantly surpassed both the controls (0.726 ± 0.111, $p = 0.0078$) and the essential thrombocythemia cases (0.706 ± 0.122, $p < 1x10^{-6}$). Furthermore, the megakaryocyte nuclear CF (0.697 ± 0.116) was significantly higher than in the controls (0.675 ± 0.123, $p = 0.026$), but only marginally and insignificantly higher than in essential thrombocythemia (0.686 ± 0.118, $p = 0.15$). Most other planimetric parameters, like the descriptors of a cellular size, nucleo-cytoplasmic ratio, the degree of nuclear segmentation expressed as a number of distinct nuclear fragments per cell, and nuclear sizes, corresponded to the values typical for normal megakaryocytes, but not for the essential thrombocythemia (data not shown).

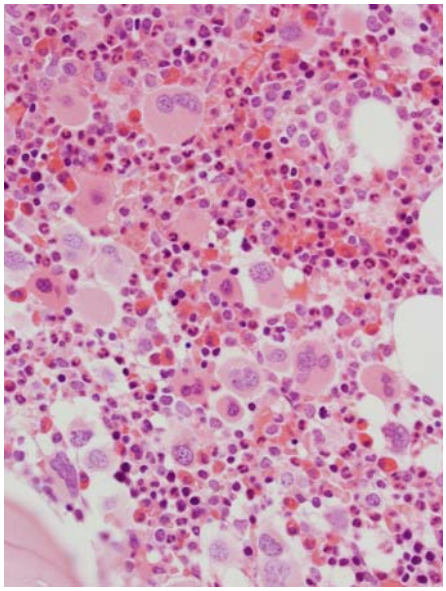

Figure 4 Trephine bone marrow biopsy of patient PL08 stained with hematoxylin & eosin, showing high marrow cellularity, markedly increased number of megakaryocytes, their occasional loose clustering and moderate increase in dispersed mature eosinophils. Note very regular shapes of the megakaryocytes. Original objective magnification 60x.

In 2/5 patients a minor fraction of megakaryocytes expressed CD34 (Table 2). This phenomenon is rare in a non-neoplastic setting, although its significance still remains unclear. There was no increase or clustering of blasts. Minimal intensification of the reticulin fiber meshwork was compatible with the upper limit of the semiquantitative norm.[156] The distention of sinuses was mild. All the cases showed a slight increase in eosinophil granulocytes, a phenomenon not reflected in peripheral blood. The erythroid cells in the four younger patients did not form clear-cut islands (erythrons) and in the case PL09, the intensified erythropoiesis was dispersed or organized in small clusters. There were no features of dyserythropoiesis or dysmyelopoiesis. A paratrabecular rim of promyelocytes was preserved. The bone structures, particularly the cortical bone, were quite massive (although not sclerotic), which may explain problems in obtaining representative trephines in the members of this family. In summary, bone marrow histology in the family demonstrated a range of features mimicking a "true" MPD, particularly megakaryocytic hyperproliferation and clustering. However the megakaryocytic morphology, with their very compact ("hyper-normal") shapes and rather normal sizes notably differed from the cytology of this hematopoietic lineage characterizing classical ET.

Table 2 Quantitative aspects of bone marrow histology in familial thrombocytosis.

UPN	Sex	Date of birth	Year of biopsy	Platelets x10^9/L	Cellularity [%]		Megakaryocytes per mm^2		Myeloid/ erythroid ratio	CD34 blast cells	Clusters of megakaryo-cytes	Increase in eosinohils	Reticulin fibres**	Distended sinuses	CD34+ megakaryo-cytes
					Cases	Age norm*	Cases	Age norm*							
PL08	F	1982	2001	887	90	72 ± 11	90	16 ± 4	4:1	<1%	+	+	0	+	+
PL09	F	1982	2000	1340	90	72 ± 11	58	16 ± 4	2:1	<1%	+	+	0/+1	+	-
PL07	F	1978	2001	956	80	70 ± 12	60	15 ± 6	3:1	<1%	+	+	+1	+	-
PL04	F	1986	2003	868	95	72 ± 11	78	16 ± 4	4:1	2%	+	+	0	+	+
PL13	M	1948	2003	905	45	59 ± 13	38	14 ± 5	3:1	nd	+	+	+1	+	nd
II-3	M	1934	1991	701-1200	60	59 ± 13	88	14 ± 5	4:1	nd	+	+	+1	+	nd

*age norms based on reference (162)
** according to reference (153)

We performed a retrospective analysis of the clinical course in affected family members of the Polish and Dutch families (Table 1 and Supplementary Table 4). Clinical characteristics of the 23 familial patients were compared with 17 sporadic ET patients previously published in 2000 (Table 3).[149] General characteristics between the two groups were comparable, except for age at diagnosis. As expected, familial thrombocytosis patients were diagnosed at an earlier age than sporadic ET patients (35 median, range 5-81 versus 67 median, range 16-84, p = 0.007). Rates on localization of hemostatic complications are summarized in Table 4. Hemorrhage (given as number of events per 100 patient years) was the only complication that differed significantly between the two groups (familial thrombocytosis patients 0.6 events per 100 patient years, essential thrombocytosis patients 11.8 events per 100 patient years, p 0.009). Venous thromboembloic events (0.6 versus 3.2, p= 0.20) and major vasomotor complications (1.3 versus 1.1, p= 0.45) were comparable in two groups. Two patients in each group developed secondary fibrosis (p value is not significant).

Table 3 Comparison of characteristics of patients with familial and sporadic thrombocythemia

	Familial thrombocythemia (n= 23)	**Sporadic essential thrombocythemia (n= 17)**	**P value**
Sex (male/ female)	8/15	8/9	0.43
Age at diagnosis, median (range)	35 (5- 81)	67 (16-84)	0.007
Time of follow up (months)	63 (0-242)	46 (0-180)	0.47
Platelets ($x10^9$/L); median (range)	890 (414-1400)	876 (452-1390)	0.78
Leukocytes ($x10^9$/L); median (range)	7.6 (5.3-10.7)	9.0 (5.3-10.7)	0.07
Number of patients with palpable splenomegaly	6	4	0.59
Rate and localization of clinical complications:			
Total number of events	11	7	0.06
Total number of events (per 100 patient years)	8.3	22.6	0.06
Thrombotic events (per 100 patient years)	6.4	8.6	0.29
Venous thromboembolic events	0.6	3.2	0.20
Vasomotor symptoms/ functional symptoms	1.3	1.1	0.45
Arteriovascular events	4.5	4.3	0.92
Hemorrhage (per 100 patient years)	0.6	11.8	0.009
Myelofibrosis (per 100 patient years)	0	2.2	0.09
Transformation to acute leukemia	0	0	
Causes of death:			
Total numbers of deaths	4	1	-
Thrombocytosis-related deaths	2	0	
Stroke	2	0	-
Hemorrhage	0	1	-
Unknown causes of death	2	0	-

When multiple events per patients occurred, every event was scored.
Vasomotor symptoms: only Raynaud's phenomenon and erythromelalgia were scored. Cold tip feeling, acral paresthesia
and headache were excluded. Limb paresthesia was included in arterial events

Discussion

We described a Polish family carrying the same splice donor mutation in intron 3 of the *THPO* gene as previously identified in a Dutch HT family.[22] The mechanism by which this mutation increases TPO protein production is loss of translational inhibition from the 5'-UTR of *THPO* mRNA.[21,22] Although there is no obvious relatedness between members of the two families, it was conceivable that the mutation may have passed on from a common ancestor (founder effect). However, haplotype analysis for the two families showed differences in 3 genetic polymorphisms located near the mutation. Due to the very small physical distance between the mutation and the three SNPs (150, 507 and 1'553 bp, respectively) a founder effect is highly unlikely. Therefore we conclude that the mutation in the two families arose de novo.

The TPO serum concentrations in most affected family members were only slightly elevated or even normal, similar as in the Dutch family.[22] A possible explanation is that the increase in platelet count and megakaryocyte mass lowers the serum concentration by binding TPO through its receptor MPL,[33,34] reaching a new equilibrium at levels close to normal. Perhaps as a consequence of the increased internalization and degradation of the MPL-TPO complex,[32,165] the MPL protein was decreased in platelets from most of the affected family members. The *MPL* mRNA levels in platelets were normal or even slightly elevated in affected individuals, indicating that the low MPL protein levels were not due to decease in mRNA expression.

Histological appearances of bone marrow in HT showed some similarities to chronic myeloproliferative disorders. Marked increase and clustering of megakaryocytes, marrow hypercellularity and occasional mild reticulin fibrosis were seen. However; in contrast to ET, the megakaryocytes of HT assumed very round compact shapes and were even more regular than megakaryocytes in normal controls. These histological findings were confirmed by our objective planimetry-based analysis and may be potentially useful in resolving diagnostic dilemmas in thrombocythemias presenting with an unclear or missing familial history. Interestingly, the bone marrow histology in an affected member of the Dutch family (II/3) we had an opportunity to review, was very similar, with

numerous and rather compact megakaryocytes showing tendency to clustering, mildly increased eosinophil granulocytes and minimal reticulin fibrosis. The megakaryocytic nuclei appeared slightly more complex, although not overtly atypical. The overall impression of bone marrow morphology corroborated close genotype-phenotype associations in this form of HT.

A total of 23 affected family members with a median follow up of 5 years (range 0-20) were studied. The clinical consequences of increased TPO production and elevated platelet count were comparable with the clinical course of 17 patients reported previously.[149] Venous thrombotic events, major vasomotor events and arterio vascular events occurred at a comparable rate as in ET patients. Hemorrhage occurred more frequently in ET patients. Main site of bleeding in these patients was the gastrointestinal tract, whereas no such event was reported in the familial thrombocytosis patients. As in ET patients, Minor vasomotor symptoms such as cold tip feeling and acral paresthesia respond well to aspirin in all affected patients.

The clinical course in familial thrombocytosis patients is generally believed to be milder than in ET. However, in our patients we could not see a milder phenotype. This might in part be due to the small number of patients studied; these findings should be verified in a larger cohort of familial thrombocytosis patients. Nevertheless, to our knowledge, this is the largest cohort of patients reported so far.

Mutation analysis in families with hereditary thrombocythemia and identification of a founder effect for a MPL mutation

Running title: Mutation analysis in HT families

Keywords: Hereditary thrombocythemia, MPL, mutation, founder effect, TPO

Abstract

Background and Objectives. Hereditary thrombocythemia (HT) is an autosomal dominant disorder characterized by sustained megakaryocytopoiesis with overproduction of platelets. Germ line mutations in HT families have been identified in the genes for thrombopoietin (*THPO*) and its receptor, *MPL*. Recently four families have been reported carrying the Ser505Asn mutation in *MPL* gene. Here we performed mutation analysis in 11 HT families and haplotype analysis for families sharing the identical *MPL* mutation.

Design and Methods. *THPO* and *MPL* gene were examined by genomic DNA sequencing. Haplotype analysis using microsatellites and single nucleotide polymorphisms (SNPs) were performed.

Results. In 11 HT families we discovered one family carrying the Ser505Asn mutation in *MPL* gene whereas the other 10 families have normal *THPO* and *MPL* gene sequences. Haplotype analysis for a total of 6 HT families carrying the same *MPL* mutation suggests a found effect.

Conclusions and interpretations. The frequent occurrence of the *MPL* Ser505Asn mutation in HT families is due to a founder effect. Unknown genetic variations cause similar clinical phenotype in other HT families.

Introduction

Hereditary thrombocythemia (HT) is an autosomal dominant disorder characterized by sustained megakaryocytopoiesis with overproduction of platelets, whose clinical features resemble sporadic essential thrombocythemia (ET). Thrombopoietin (TPO) and its receptor MPL regulate proliferation and maturation of megakaryocytes and thus platelet production. Germ-line mutations have been described in thrombopoietin gene (*THPO*) and *MPL* gene in HT families. All the five mutations identified in the *THPO* gene to date are located in the 5 prime untranslated region (5'-UTR) of the mRNA sequence, which contains upstream open reading frames (uORF) that inhibit the translation of *THPO* mRNA.[21-27] The mutations remove the inhibitory uORFs and lead to increased translation of the *THPO* mRNA causing elevated TPO serum levels and overproduction of platelets.[22,24] A missense mutation in the transmembrane domain of *MPL* has been identified in a Japanese family with HT.[51] This G → A transition changes a serine to an asparagine at amino acid position 505 (Ser505Asn). Mutant MPL protein is hyperactive and activates downstream signals to stimulate cell proliferation, which result in excessive platelet production. All the germ-line mutations in *THPO* and *MPL* gene have not been identified in sporadic patients with thrombocythemia,[140] though mutations in a different location within juxtamembrane domain of MPL have been found in patients with chronic myeloproliferative disorders (MPD), in particular idiopathic myelofibrosis (IMF) and ET,[54,55] One splice donor mutation in *THPO* gene was identified in a second HT family and arose *de novo* (Liu *et al* submitted). Recently four out of five families analyzed have shown carrying the Ser505Asn mutation in *MPL* gene,[52] suggesting a relatively high frequency of this mutation in HT etiology. In some HT families *THPO* and *MPL* genes were excluded as the cause of thrombocytosis.[143,144] In this study, we analyzed *THPO* and *MPL* gene in 11 HT families and discovered one family carrying the Ser505Asn mutation in *MPL* gene. The other 10 HT families have normal *THPO* and *MPL* gene sequences, indicating some unknown genetic lesion resulting in the similar phenotype. The frequent reoccurrence of this mutation in different HT families raise the question whether the mutation is originated from a single founder or multiple mutational events. We performed haplotype analysis for a total of 6 HT families carrying the Ser505Asn mutation in *MPL* gene.

Design and Methods

Patients and clinical features

The proposita, a 5-year-old came to our observation for asymptomatic thrombocytosis since birth. The patient was sent to our Institution for a suspect of essential thrombocythemia. However, clonogenic assays were essentially negative, and we investigated family history and blood cell counts, discovering that other 2 siblings in the same family over 3 generations had high platelet counts, i.e. the proposita's father and grand-mother.

THPO and MPL gene sequencing

Genomic DNA was extracted from blood or buccal swab using a standard proteinase K (Promega, Madison, WI, USA)/phenol (Fluka Chemie AG, Buchs, Switzerland) extraction protocol. Exons including intron/exon boundaries of the *THPO* and *MPL* gene were sequenced from PCR fragments, amplified from genomic DNA of one affected family member for each family. The primer sequences are shown in Supplementary Table 1 and 2. The PCR conditions were 95ºC for 2 min, 94ºC for 30 sec, 60ºC for 30 sec and 72ºC for 1 min for 35 cycles. Sequencing was performed on an Applied Biosystems 3130 DNA sequencer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocols.

Haplotype analysis

To determine the founder effect, 4 informative microsatellite markers located in the vicinity of the *MPL* gene locus were used for haplotype analysis. Sequences of the primers are provided in Supplementary Table 3. The PCR products were analyzed using the Applied Biosystems 3130 genetic analyzer and the Genemapper software package version 3.5 (Applied Biosystems, Foster City, CA, USA). The haplotypes were determined based on the segregation within the pedigrees and the sizes of the PCR products of the co-segregating microsatellite markers were compared between affected members of the six families.

Results and Discussion

The clinical features of 3 family members with thrombocytosis in our newly identified Italian family with the Ser505Asn mutation are summarized (Table 1). As compared to non-affected family members, platelets from affected siblings were significantly smaller (MPV of affected 8.7±0.4 fL, n=3; non-affected 11.0±0.7, n=5, p=0.0006) and had a narrower distribution of the population (PDW of affected: 10.6±0.8 versus 13.7±1.6, $p<0.05$). More interestingly, both MPV and PDW were higly-significantly inversely correlated to platelet counts within the affected and non-affected members. Thus, it appears that HT platelets, in spite of their increased number, are in fact smaller, which might compensate for the whole platelet circulating mass. Furthermore, platelet function was analyzed by the Platelet-Function-Analyzer (PFA)-100 methods which is a shear-dependent test in whole blood. By this method the affected family members had consistently prolonged closure times in response to both collagen-ADP (CADP) and collagen-epinephrine (CEPI) cartridges.

We did genomic DNA sequencing for the entire encoding sequences, the 5'-UTR, and the intron/exon boundaries of *THPO* and *MPL* gene for one affected family member from 11 HT families (Figure 1). In one Italian family we identified a G→A transition in exon10 of *MPL* gene (Figure 2A). This mutation changes a serine into an asparagine in the amino acid sequence (Figure 2B). The mechanism by which this mutation increases platelet production is that the mutant MPL receptor is activated and stimulate downstream signaling pathway in a ligand independent fashion.[51] The same mutation was found in 5 additional HT families (Figure 3), and four of them have been reported recently.[52] The frequent reoccurrence of this mutation in different families indicate that the mutation may have originated from a common ancestor (founder effect). We therefore performed haplotype analysis with microsatellite markers in the vicinity of the *MPL* mutation in affected members from the six families. A founder effect, i.e. descent of all families from a common allele, is expected to result in sharing of allelic sequence polymorphisms in the vicinity of the *MPL* mutation. We compared 3 microsatellite markers located between 42 kb to 1290 kb from the *MPL* mutation, and all PCR products that represent the haplotype of the affected allele in the six families showed the idential sizes (Figure 4), suggesting that the mutation occurred from a single founder event.

Table 1 Summary of clinical data of 3 members of the Italian family with thrombocytosis.

UPN	Sex	Age at Dx	Platelets 150-450 x10^9/L	WBC 3.5-10.0 x10^9/L	RBC 4.2-6.3 x10^{12}/L	Hemo-globin 12-14 F 14-18 M g/L	Spleno-megaly (last follow up)	PFA-100 assay CEPI (up to 195 sec)	PFA-100 assay CADP (up to 130 sec)	Hepato-megaly (last follow up)	Thrombocytosis-associated symptoms	Important co-morbidity
RC02	F	64	926	7.2	4.20	12.9	No	> 300	> 300	No	None	none
RC05	M	35	625	5.1	5.59	16.2	No	180	158	No	None	smoker
RC06	F	5	1092	7.7	4.75	13.4	No	> 300	190	No	No	none

UPN, unique patient number; Dx, diagnosis; WBC, white blood cells; RBC, red blood cells

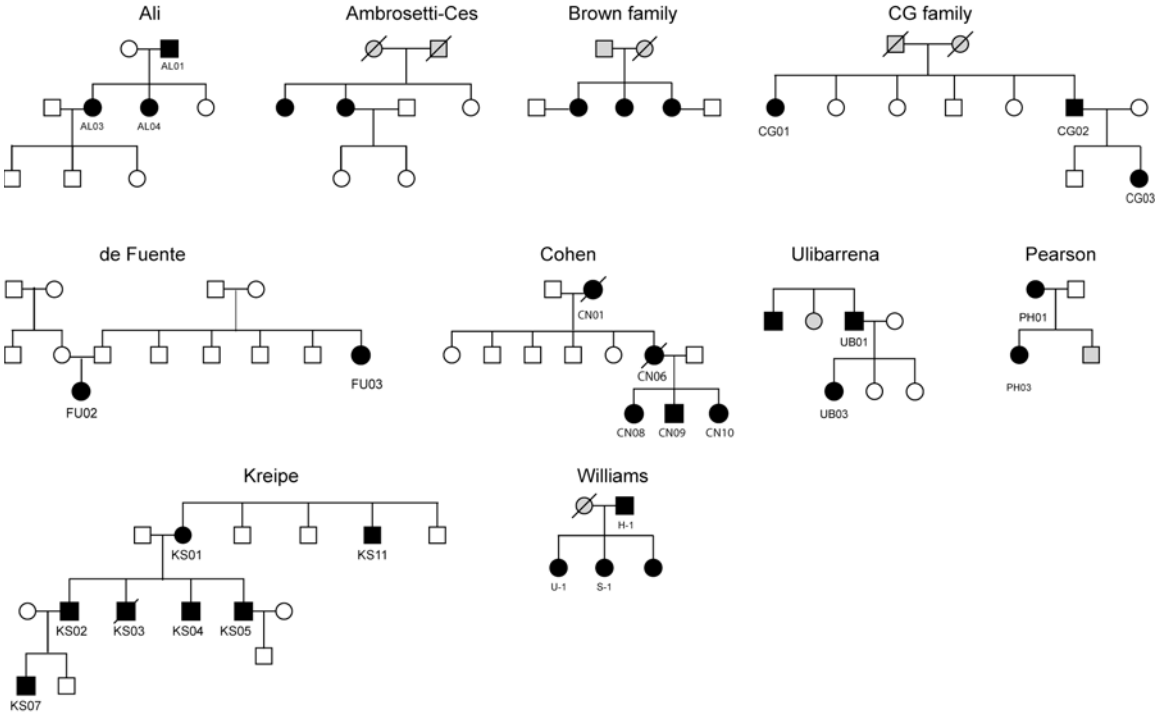

Figure 1 Family pedigrees of the 9 HT families with normal *THPO* and *MPL* gene. Square represents male and circle female; solid symbol indicates affected, open unaffected, grey not studied and slash marks deceased. Unique patient numbers are shown below the symbols which correspond to the clinical data in Table 1.

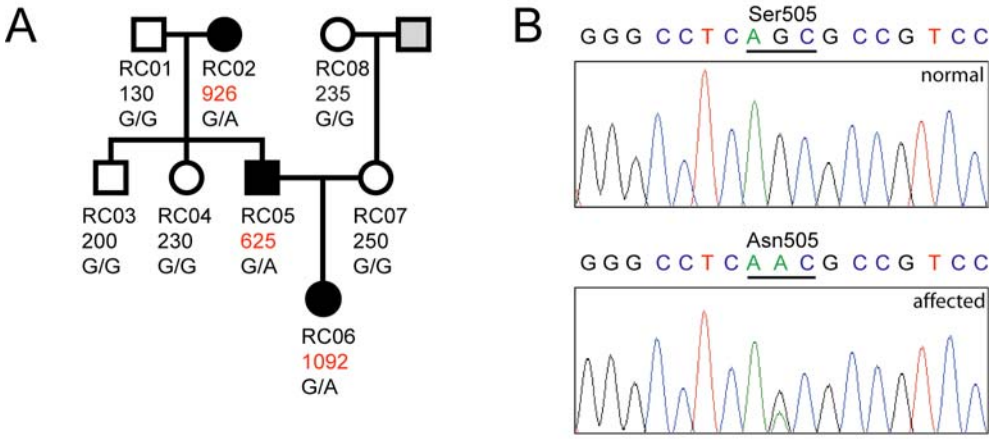

Figure 2 The Italian family with Ser505Asn mutation. A) Pedigree of the Italian family. Square represents male and circle female; solid symbol indicates affected, open unaffected and grey not studied. Unique patient numbers are shown below the symbols. Platelet count number is shown below the unique patient number. Number in red is above the normal range. The genotype at the mutation point is indicated below the platelet counts. Noted that affected individual is heterozygous for the position, whereas unaffected are homozygous. **B)** The sequencing chromatograms of the mutation region in *MPL* exon10 from an unaffected and an affected family member are shown. Noted that the G→A transition is a double peak. Bases and changed amino acid are indicated above the chromatograms.

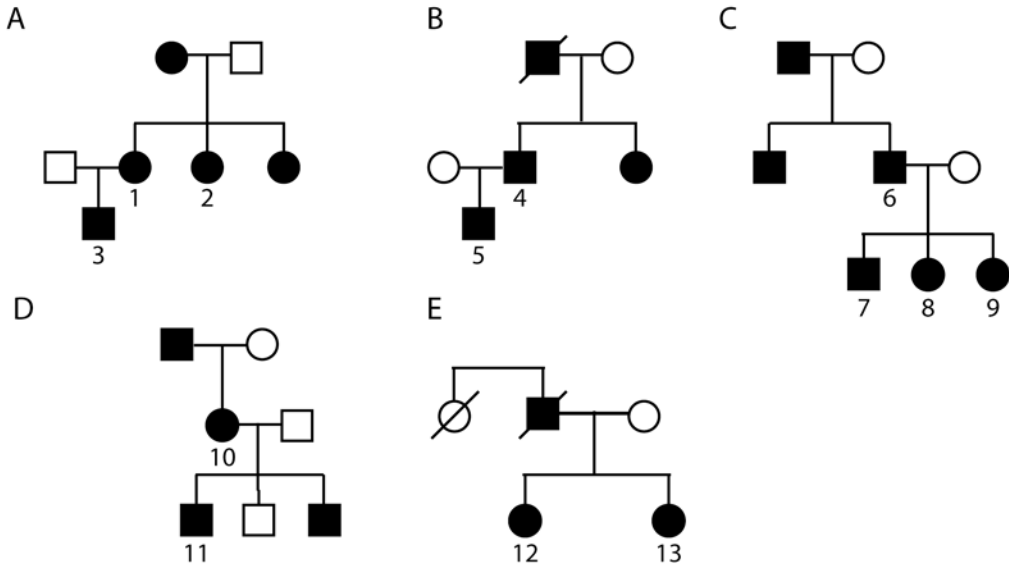

Figure 3 Family pedigrees of the 5 HT families with Ser505Asn mutation in *MPL* gene. Square represents male and circle female; solid symbol indicates affected, open unaffected and slash marks deceased. Unique patient numbers that correspond to the ones in Table 3 are given below the symbols.

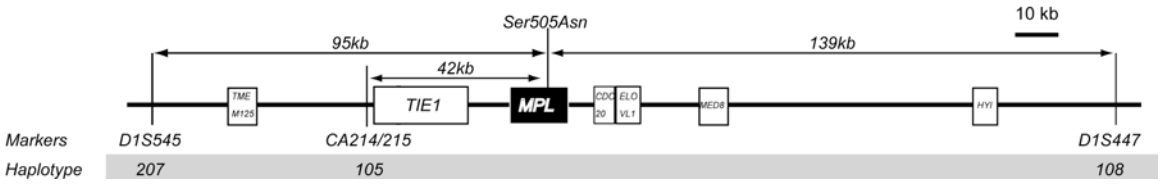

Figure 4 Haplotype analysis for the *MPL* locus in the five Italian HT families carrying the same Ser505Asn mutation. The chromosomal locus containing the *MPL* gene is shown. Black box represents the *MPL* gene and open boxes genes close to it: *MPL*, myeloproliferative leukemia virus oncogene; *TMEM125*, transmembrane protein 125; *TIE*, tyrosine kinase with immunoglobulin-like and EGF-like domains 1; *CDC20*, cell division cycle 20 homolog (S. cerevisiae); ELOVL1, enlongation of very long chain fatty acids (FEN/E102, SUR4/E103, yeast)-like 1; MED8, mediator of RNA polymerase II transcription subunit 8 homolog (S. cerevisiae); HIY, hydroxypyruvate isomerase homolog (E. coli). The distances of microsatellite markers to the location of mutation are shown above the locus. Names of markers are shown below the locus, and numbers below the markers indicate the sizes in nucleotides of the PCR products of the co-segregating mutated alleles in the five families. Note that all the allele sizes are identical in the six families.

RESULTS II:

Genetic studies of a hereditary thrombocythemia family with normal thrombopoietin and MPL gene

Abstract

Background and Objectives. Hereditary thrombocythemia (HT) is an autosomal dominant disorder with clinical features resembling sporadic essential thrombocythemia (ET). Although germ line mutations causing HT have been identified in the genes for thrombopoietin (*THPO*) and its receptor, *MPL*, genetic lesions for many HT families remain unknown. Here we studied a HT pedigree with 10 affected family members.

Design and Methods. Linkage analysis was performed using microsatellites and SNP chip arrays. Candidate genes were selected for genomic DNA sequencing. A candidate mutation was cloned into retroviral vectors to infect cell lines and mouse bone marrow cells for transplantation. Cell proliferation rate was tested in cell lines transfected with candidate gene. Peripheral blood counts were determined in transplanted mice.

Results. We identified two genetic regions that co-segregated with thrombocytosis within the pedigree. Candidate gene sequencing revealed one novel polymorphism in the gelsolin gene, which changed a glycine into a cysteine in the protein sequence. Cell lines transfected with the candidate mutant gelsolin did not show any proliferation advantage over the wild type. Blood counts were normal in mice transplanted with bone marrow expressing the candidate mutant gelsolin. The other 33 candidate genes in the co-segregating region were also sequenced but no variant was found.

Conclusions and interpretations. Genome-wide linkage analysis for this HT family leads to the identification of two genetic regions. Sequencing candidate genes in these two regions is under way. One candidate mutation in gelsolin needs futher functional studies to validate it as disease causing mutation.

Introduction

Hereditary thrombocythemia (HT), also known as familial thrombocytosis or familial essential thrombocythemia is characterized by sustained proliferation of megakaryocytes and overproduction of platelets, which is often clinically indistinguishable from the sporadic essential thrombocythemia (ET). HT is inherited as an autosomal dominant trait with a nearly hundred percent penetrance. The pathogenesis for some HT families has been elucidated with the discovery of germ-line mutations in genes encoding the primary regulators for platelet production thrombopoietin (TPO) and its receptor, MPL. To date, five different germ-line mutations in the thrombopoietin (*THPO)* gene have been identified and all of them alter the 5 prime untranslated region (5'-UTR) of the *THPO* mRNA, which contains upstream open reading frames (uORF) that inhibit the translation of *THPO* mRNA.[21-27] The mutations remove the inhibitory uORFs and lead to increased translation of the *THPO* mRNA causing elevated TPO serum levels and overproduction of platelets.[22,24] A missense mutation in the transmembrane domain of *MPL* has been identified in several HT familes.[51,52] This mutation changes a serine to an asparagine in the MPL protein sequence. The mutant MPL protein is hyperactive and results in excessive platelet production. Recently, mutations in the juxtamembrane domain of MPL have been found in patients with the chronic myeloproliferative disorders (MPD), in particular idiopathic myelofibrosis (IMF) and ET,[54,55] but *THPO* mutations have not been detected in patients with sporadic MPD.[140] In some HT families, both *THPO* and *MPL* genes can be excluded as the cause of thrombocytosis and thus, other as yet unknown genes must be involved in causing an identical phenotype.[143,144] Here we performed genetic study on a previously described HT family,[166] in which *THPO* and *MPL* genes were excluded as the disease-causing genes,[144] to search for novel genetic lesion causing the thrombocytosis.

Design and methods

Patients and clinical features

The clinical features of this family have been described previously.[166] Briefly, the propositus had a persistent elevation of platelet counts in the range of 1000 x 10^9 per liter, splenomegaly, a normal hemoglobin and white blood cell count. Her mother and her maternal aunt, and her two sons were later found to display high platelet counts. In the propositus, her mother and her older son platelet aggregation tests were abnormal including disaggregation to ADP, no response to epinephrine, delayed aggregation to collagen, and delayed submaximal response to arachidonic acid. Bleeding time was normal in the propositus, whereas her mother has a prolonged bleeding time. For our study, more family members were recruited. In total, 10 affected and 11 unaffected family members participated in the genetic study.

Blood cells separation, DNA and RNA extraction

Blood cells were separated by standard protocols using Histopaque (Sigma, St. Louis, MO, USA) gradient centrifugation. Granulocytes and peripheral blood mononuclear cells (PBMC) were collected respectively. Platelets were collected using the Sepharose (Amersham Pharmacia Biotech AB, Uppsala, Sweden) gel filtration method.[155] DNA was extracted using a standard proteinase K (Promega, Madison, WI, USA)/phenol (Fluka Chemie AG, Buchs, Switzerland) extraction protocol. RNA isolation was carried out using the TRIfast reagent (peqLab Biotechnology GmbH, Erlangen, Germany).

Genotyping

Linkage mapping set v2.5-MD10 (Applied Biosystems, Foster City, CA, USA) was used for genome-wide analysis. Genomic DNAs were PCR amplified and analyzed with an ABI 3100 genetic analyzer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol. The PCR program was at 94ºC for 15 sec, 55ºC for 15 sec, 72ºC for 30 sec for 10 cycles, 89ºC for 15 sec, 55º for 15 sec and 72ºC for 30 sec for 20 cycles. The Affymetrix GeneChip Human Mapping 50K Xba 240 was used to genotype single nucleotide polymorphisms (SNPs) in genomic DNA of 12 family members according to

the Affymetrix GeneChip Mapping Assay Manual (Affymetrix Inc., Santa Clara, CA, USA)

Linkage analysis

Genotypes from microsatelite mapping were scored using the Genemapper software package version 3.5 (Applied Biosystems, Foster City, CA, USA) and linkage analysis was carried out by FASTLINK software package version 4.1p. The SNP calls were generated by GeneChip DNA Analysis Software. dCHIP version 2005 was used to perform parametric linkage analysis.[167] An autosomal dominant inheritance model with a 100% penetrance was used for both analyses.

Genomic DNA sequencing

The entire coding region including intron/exon boundaries of candidate genes was sequenced from PCR fragments amplified from genomic DNA of one affected family member, EY01. The primer sequences are available upon request. The PCR conditions were 95ºC for 2 min, 94ºC for 30 sec, 60ºC for 30 sec and 72ºC for 1 min for 35 cycles. Sequencing was performed on an Applied Biosystems 3700 DNA sequencer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocols.

cDNA synthesis, RT-PCR and Quantitative RT-PCR

Total RNA (2 μg) was reverse transcribed to generate cDNA using Superscript II reverse transcriptase (Life Technologies). Real-time PCR primers for ribosomal protein L19 (*RPL19*) and gelsolin were designed across exon-intron junctions. In addition, primers for gelsolin were designed specific for human sequence to distinguish the expression of human gelsolin in mouse bone marrow. The primer sequences are available upon request. The ΔCT values were derived by subtracting the threshold cycle (C_T) values for target genes from the C_T value for RPL19, which served as an internal control.[158] Relative expression level for each reaction was calculated as $2^{-\Delta CT}$. Folds were calculated with one calibrator set as 1. All reactions were run in duplicates using the ABI 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA).

Western blotting

Protein lysate was denatured at 95°C for 2 minutes and separated by 10% sodium dodecyl sulfate-plyacryamide gel electrophoresis (SDS-PAGE). The gel was transferred electrophoretically onto 0.45-μM nitrocellulose membrane (Whatman GmbH, Dassel, Germany).[168] The membrane was blocked with 5% nonfat milk in 1 x PBS with 1% Tween-20 for 2 hours at room temperature and incubated overnight at 4°C with diluted primary antibody and then probed with a mouse monoclonal antibody (BD Biosciences, San Jose, CA, USA) or a goat polyclonal antibody (Santa Cruz Biotechnology, CA, USA). The membranes were re-probed using a monoclonal antibody against human CD61 (BD Biosciences, San Jose, CA, USA) serving as a loading control.

DNA construct

A human gelsolin cDNA clone in pOTB7 vector was purchased from RZPD (Germany). The full length gelsolin cDNA was PCR amplified using primers with XhoI and EcoRI restriction enzyme cutting sites and then cloned in pSP72 vector (Promega). The candidate mutant gelsolin cDNA fragment was amplified by RT-PCR using platelet RNA isolated from one affected family member (EY03) and cloned into the pCR 2.1-TOPO vector (Invitrogen). The candidate mutant fragment was cut out and inserted into the PSP72 vector containing gelsolin cDNA. The wild type and candidate mutant gelsolin were finally cloned into the retroviral vector pMSCVpuro and MSCV2.2IRESGFP encoding green fluorescent protein (kindly provided by Dr. J. Cools, Flanders Interuniversity Institute for Biotechnology, Leuven, Belgium). The gelsolin cDNA in the resulting expression construct was verified by sequencing.

Cell proliferation assay

Stably transfected BaF3 cells, 32D cells and UT7 cells were selected with puromycin (2.5 μg/ml for BaF3 and 32D cells and 0.625 μg/ml for UT7 cells). The proliferation of the cells grown for 3 days in presence of cytokines at different concentrations was assessed using the Cell proliferation kit II XTT (Roche Molecular Biochemicals).

Retrovirus production and Bone marrow transplantation in mice

Equal amounts of retroviral vector with gelsolin and packaging plasmids (Ecopak) were incubated with FuGene 6 reagent (Roche Applied Science) for 10 minutes, and then added to the human embryonic kidney-cell line, 293T. The supernatants were harvested 48 and 72 hours later and used to transfect murine bone marrow cells. BALB/c donor mice (6-10 weeks old) were injected with 5-fluorouracil at 150 mg per kg of body weight 5 days before harvesting bone marrow from femurs and tibias. Cells were cultured for 24 hours in transplantation medium (RPMI-1640 medium, 10% fetal-calf serum, 6 ng/ml of murine interleukin-3, 10 ng/ml of human interleukin-6, and 10 ng/ml of murine stem-cell factor) and then transfected with retrovirus supernatant. Briefly, 4 x 10^6 bone marrow cells were centrifuged at 2500 rpm for 90 minutes in the presence of 1 ml of retroviral supernatant and 2 μl of Polybrene Infection Transfection Reagent (American Bioanalytical). Cells were infected for a second time on the day after and subsequently resuspended in Hank's balanced salt solution (Gibco) and then injected into lethally irradiated BALB/c mice. Peripheral-blood counts were evaluated for each recipient 4 weeks after transplantation.

Results

Genome-wide linkage analysis for 21 family members using microsatellite markers revealed one region on chromosome 9q that co-segregated with the thrombocytosis with the highest logarithm of odds (LOD) score of 3.9 at theta = 0 (Figure1).

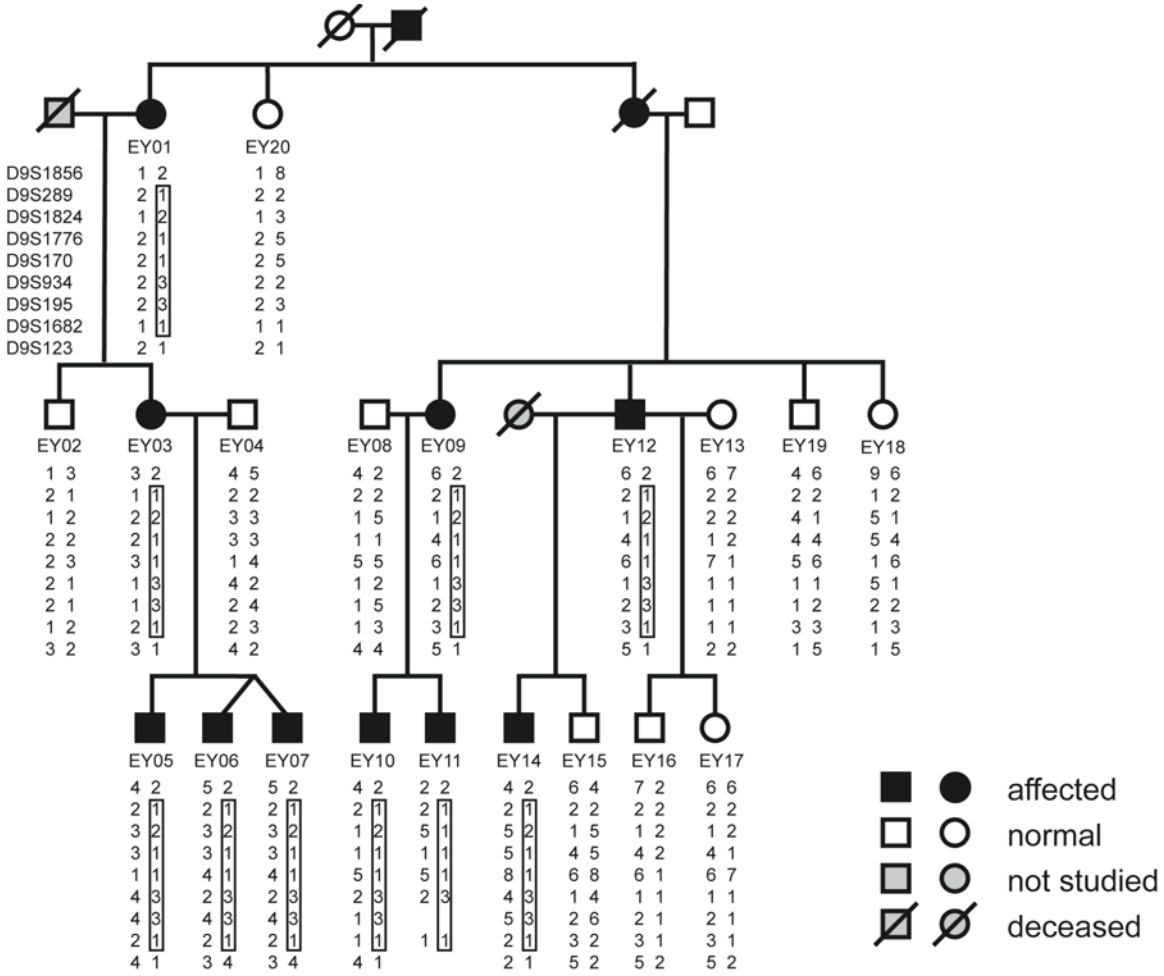

Figure 1 Segregation of microsatellite markers on chromosome 9. Filled black symbols, affected individuals; open symbols, normal individuals; filled grey symbols, not studied; crossed symbols, deceased. Unique patient numbers are placed below the symbols. Haplotypes for nine polymorphic microsatellite markers from D9S1856 through D9S123 are shown with numbers defined by different sizes of PCR fragments. The unique haplotype linked with the disease is boxed.

The 11 Mb region contains 106 genes, including a core region with LOD score higher than 3 (17 genes) and two flanking regions at each side with LOD score between –2 and 3, which could not be exclude statistically. SNP chip arrays for 13 family members confirmed the genetic region on chromosome 9q and identified an additional region on chromosome 20 (Figure 2). The highest LOD score from SNP chips data of 13 family members was 2.68. The core region on chromosome 9q from SNP arrays was slightly larger than the one from microsatellite mapping containing 29 genes (Table 1). The candidate genes on chromosome 20q are summarized in Table 2.

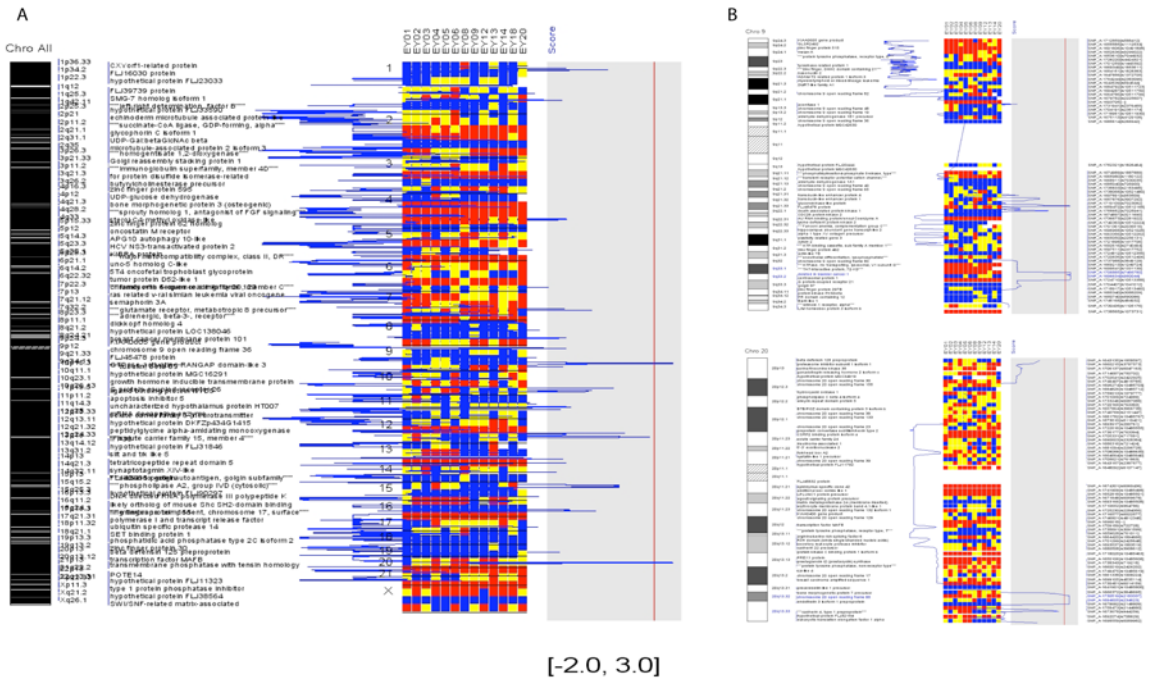

Figure 2 Parametric linkage analysis of 50K SNP array Xba 240 for 13 family members. A) The whole genome Lod score curve was plotted on the right side with a threshold from –2 to 3, Lod score of 2 is indicated by a red line. Genotypes of each family member are shown on the left side with different colors, red represents AA, blue BB, and yellow AB. **B)** The LOD score curve of the segregating region on chromosome 9q and chromosome 20q. The region on chromosome 9q is consistent with the one from microsatellite markers with the highest LOD score of 2.68.

Table 1 Overview of genes in the region of chromosome 9q

Gene Symbols	Gene Full Names
TRIM32	tripartite motif-containing 32
TLR4	toll-like receptor 4
LOC340477	hypothetical LOC340477
LOC389787	similar to Translationally controlled tumor protein (TCTP) (p23) (Histamine-releasing factor) (HRF)
LOC442434	similar to beta-tubulin 4Q
DBC1	deleted in bladder cancer 1
CDK5RAP2	CDK5 regulatory subunit associated protein 2
EGFL5	EGF-like-domain, multiple 5
LOC392387	similar to Adenosylhomocysteinase (S-adenosyl-L-homocysteine hydrolase) (AdoHcyase)
FBXW2	F-box and WD-40 domain protein 2
LOC402377	similar to beta-1,3-N-acetylglucosaminyltransferase bGnT-5; beta 1,3 N-acetyglucosaminyltransferase Lc3 synthase
PSMD5	proteasome (prosome, macropain) 26S subunit, non-ATPase, 5
PHF19	PHD finger protein 19
TRAF1	TNF receptor-associated factor 1
C5	complement component 5
CEP1	centrosomal protein 1
RAB14	RAB14, member RAS oncogene family
GSN	gelsolin (amyloidosis, Finnish type)
LOC441464	LOC441464
LOC441465	hypothetical gene supported by AK130710
STOM	stomatin
GGTA1	glycoprotein, alpha-galactosyltransferase 1
LOC441466	similar to HMG-1
DAB2IP	DAB2 interacting protein
C9orf20	chromosome 9 open reading frame 20
C9orf148	chromosome 9 open reading frame 148
NDUFA8	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8, 19kDa
C9orf18	chromosome 9 open reading frame 18
LHX6	LIM homeobox 6
Total number of genes	29

Table 2 Overview of genes in the region of chromosome 20q

Gene Symbols	Gene Full Names
TMEPAI	transmembrane, prostate androgen induced RNA
C20orf85	chromosome 20 open reading frame 85
C20orf86	chromosome 20 open reading frame 86
PPP4R1L	protein phosphatase 4, regulatory subunit 1-like
RAB22A	RAB22A, member RAS oncogene family
VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C
APCDD1L	adenomatosis polyposis coli down-regulated 1-like
LOC149773	hypothetical protein LOC149773
MGC4294	hypothetical protein MGC4294
STX16	syntaxin 16
NPEPL1	aminopeptidase-like 1
LOC391258	hypothetical LOC391258
SANG	GNAS1 antisense
GNAS	GNAS complex locus
TH1L	TH1-like (Drosophila)
CTSZ	cathepsin Z
LOC729181	hypothetical protein LOC729181
TUBB1	tubulin, beta 1
ATP5E	ATP synthase, H+ transporting, mitochondrial F1 complex, epsilon subunit
C20orf45	chromosome 20 open reading frame 45
MRPS16P	mitochondrial ribosomal protein S16 pseudogene
C20orf174	chromosome 20 open reading frame 174
EDN3	endothelin 3
LOC645605	similar to Protein FAM38A
PHACTR3	phosphatase and actin regulator 3
Total number of genes	25

We did genomic DNA exon sequencing for all the 29 candidate genes on chromosome 9q and 5 candidate genes on chromosome 20. One novel polymorphism in exon 6 of thegelsolin gene was found in all affected family members (Figure 3A). This G→T transversion changed a glycine into a cysteine in the second of its six gelsolin-like domains. Expression levels of gelsolin mRNA in platelets were determined in 3 affected and 2 unaffected family members and did not show significant differences (Figure 3B). Gelsolin protein levels were also measured in platelets protein lysates and no variation was observed between the affected and unaffected family members (Figure 3C).

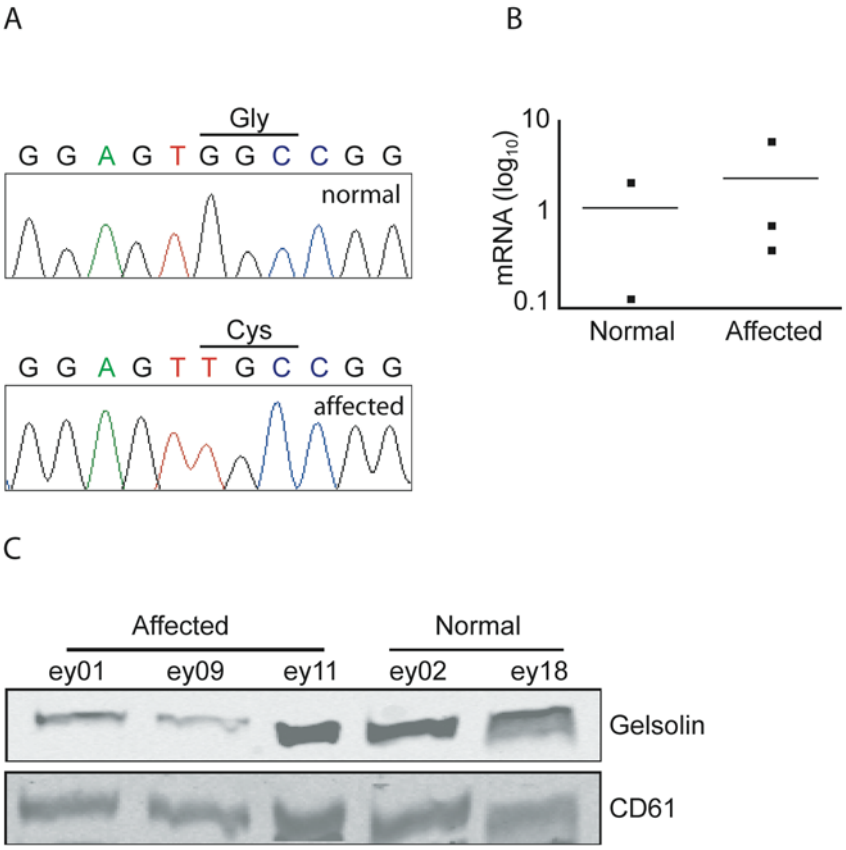

Figure 3 The Candidate mutation in the gelsolin gene. A) Sequencing chromatograms of the polymorphism in gelsolin which are cloned in expression vectors. In all affected family members there is a G→T transversion causing a change of glycine to cysteine. **B)** Gelsolin mRAN levels in platelet cDNA in 2 unaffected and 3 affected family members available. Relative mRNA level were calculated from ΔCT values and plotted in logarithmic scale. The horizontal line indicates the mean for each group. **C)** Western blotting of gelsolin in platelet protein of unaffected and affected family members. CD61 is reprobed to the membrane as a loading control.

The candidate mutant and wild type human gelsolin gene were cloned into retroviral expression vectors to generate stably transfected cell lines. In the mouse interleukin-3–dependent cell line BaF3, the over expression of candidate mutant gelsolin did not show any proliferation advantage compared with the wild type gelsolin or empty vector (Figure 4). Similar results were observed in UT7 cells and 32D cells (data not shown).

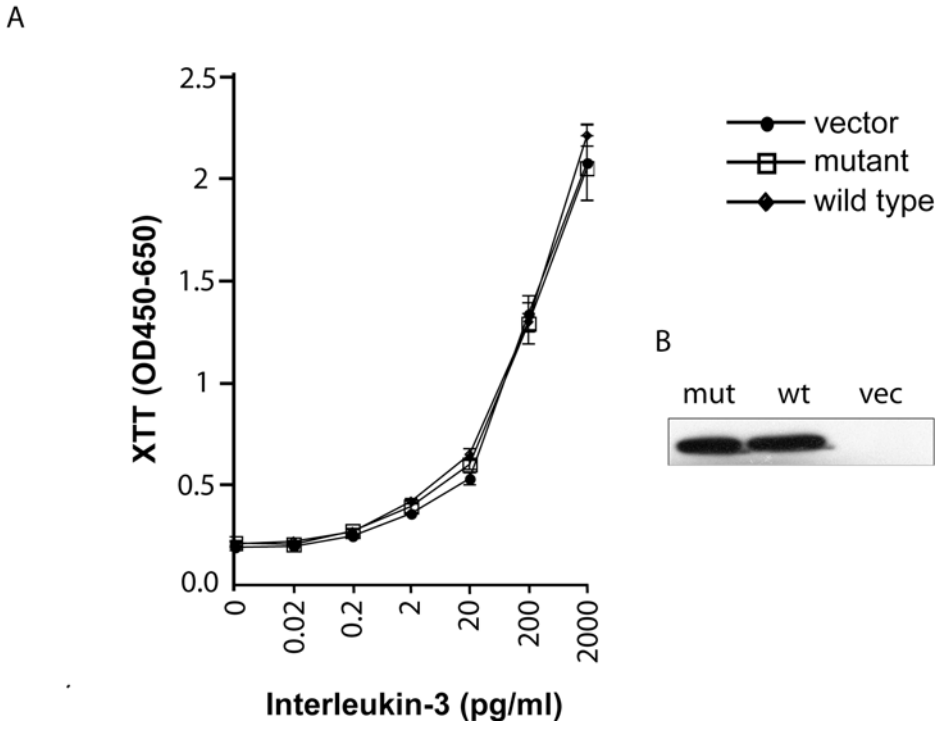

Figure 4 Cell proliferation XTT assay. A) Proliferation of BaF3 cells transfected with the gelsolin candidate mutant, wild-type, or the empty vector in the absence of interleukin-3 (concentration of 0) and the presence of increasing concentrations of interleukin-3, as determined by the tetrazolium salt (XTT) assay. The mean (±SD) of triplicate results is shown. Increased optical density (OD) of the XTT dye corresponds to increased numbers of cells. **B)** Over-expression of gelsolin protein. Western blotting was performed using an antibody against human gelsolin in cell lysates. Over-expression of gelsolin was detected in cells transfected with candidate mutant, the wild type, but not in cells with the empty vector.

Retroviral vectors expressing mutant and wild type human gelsolin were used to infect mouse bone marrow cells and then transplant these cells into lethally irradiated mice. Transplanted mice were physically normal. Peripheral blood counts at 4 weeks, 10 weeks and 20 weeks after bone marrow transplantation did not show variations in platelet numbers (Figure 5A), hemoglobin concentrations, or white blood cells numbers (data not shown). The efficiency of transplantation in mouse bone marrow cells determined by flow cytometry based on the green fluorescent protein (GFP) was 13.7% in the wild type gelsolin and 28.5% in the candidate mutant gelsolin (Figure 5B). The mRNA levels of human gelsolin were detected in sorted cells by real-time PCR using primers specific for the human gelsolin sequence. Up to 100 folds of increase in human gelsolin mRNA in FACS-sorted GFP positive mouse bone marrow cells compared to GFP negative cells was detected (Figure 5C). The over expression of human gelsolin protein were detected in the GFP positive cells. Though the antibody was against both mouse and human gelsolin, mouse gelsolin protein was undetectable in the GFP negative cells (Figure 5D).
In total, 30 HT families, 172 sporadic MPD patients as well as 102 normal controls were screened for the polymorphism in gelsolin and none of them showed the same variation.

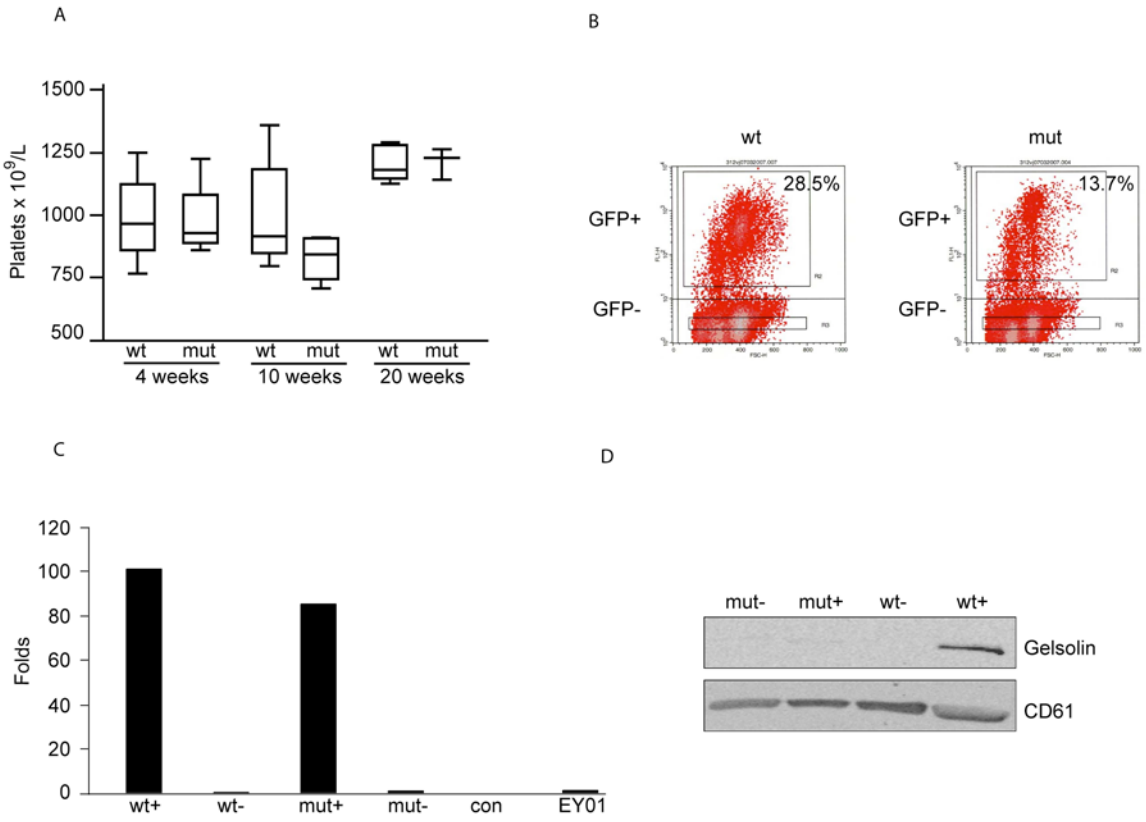

Figure 5 Recipient mice after transplantation with bone marrow cells transduced with gelsolin. A) The platelet counts of recipient mice 4 weeks, 10 weeks and 20 weeks after bone marrow transplantation are shown with boxes that represent the interquartile range that contains 50% of the values, the horizontal line in the box marks the median and bars indicate the range of values. **B)** FACS based on GFP in mice bone marrow cells transplanted with candidate mutant or wild type gelsolin. **C)** The mRNA levels of human gelsolin in sorted mouse bone marrow cells are shown. The mRNA levels of gelsolin in the platelets from one affected family member EY01 was set as 1 fold. **D)** Western blotting for gelsolin protein in sorted mice bone marrow cells. CD61 was used as a loading control.

Discussion and perspectives

The disease-causing gene mutations are unknown in many HT families.[139] Here we performed genome-wide linkage analysis for a HT family with normal TPO and MPL using two methods: microsatellite marker mapping and SNP chip array mapping. One co-segregating region on chromosome 9q containing 29 candidate genes was revealed by both methods with the highest LOD score of 3.9. By sequencing all the candidate genes we identified one novel polymorphism in the gelsolin gene. Gelsolin is an actin-binding protein abundant in platelets. Gelsolin-null mice have normal embryonic development and longevity but decreased platelet shape changes and prolonged bleeding time.[169] A point mutation at the nucleotide position 654 of the gelsolin cDNA causes familial amyloidosis of the Finnish type (FAF). This mutation causes an amino acid substitution of aspartic acid at residue 187 by either asparagine or tyrosine.[170-172] Though there is no evidence showing gelsolin involved in thrombocytosis, it has been shown that gelsolin has multiple biological properties in addition to cytoskeletal actin modulation. Gelsolin is cleaved by caspase-3 and may operate as an executor of apoptosis.[173] Additionally, it has been long regarded as tumor suppressor in breast cancer and other carcinomas because of its down-regulation during tumor progression.[174] Gelsolin appears to act down-stream of Ras and phosphoinositides to promote motility and invasive potential of transformed cell lines.[175]

Over-expression of candidate mutant gelsolin in mouse and human cell lines did not show any proliferation advantages compared to the wild type gelsolin or to the empty vector, which suggests that gelsolin might not be involved in the pathways regulation cell proliferation. Bone marrow transplantation of candidate mutant gelsolin did not generate thrombocytosis or any other obvious phenotype in mice. The data so far do not support gelsolin as the disease-causing gene in this family though inheriance of the polymorphism in this gene is compatible with such a hypothesis. The candidate mutation was not detected in other HT families, sporadic MPD patients and normal controls, indicating that this variation is rare. A second allele in a HT family would be a strong indicator that mutant gelsolin is the disease causing gene.

Several other functional studies for the candidate mutant gelsolin are planned. First is to study the process of platelet releasing in megakaryocytes cultured in vitro. Since our data showed that over expression of gelsolin did not increase the proliferation rate of cells, the high platelet counts might be due to a defect in the platelet releasing process, which remains poorly understood to date. Platelet formation and releasing from megakaryocytes requires complicated changes in the cytoskeletal organization including microtubule enlongation and actin assembling.[66,176] Actin is the most abundant protein in platelets. Gelsolin, as a member of the actin-binding protein family, is thought to function in capping, severing, nucleating or bundling actin filaments. Expression of gelsolin protein was detected in platelets in our study and many other studies. Besides, abnormal platelets functions were reported in affected family members. Therefore it is of interest to see if the candidate mutant gelsolin plays a role in the over production of platelets in this family. A second approach is to perform bone marrow transplantation using mouse gelsolin. Human and mouse gelsolin share 95% similarity at the amino acid sequences. Though transplanted human gelsolin failed to produce any phenotype in mice, the reason could be that human gelsolin is unable to interact with mouse signal molecules or too low infection rate.

Although gelsolin is considered as a top candidate in this family, it is possible that another gene might carry the disease-causing mutation. Sequencing of genes in the segregating region on chromosome 9q has not been fully completed due to some PCR-related technical problems. Additionally, only 5 candidate genes in the region on chromosome 20q identified by SNP chips array were sequenced. There are still 20 genes in this area that remain uncharacterized for this study. Searching for disease-causing gene in this family still needs complete sequencing for all genes.

RESULTS III:

Genetic analysis of a family with congenital secondary polycythemia

Abstract

Background and Objectives. Polycythemia is a condition with an increased red blood cell mass. Erythropoietin (EPO) receptor gene and VHL gene mutations have been described in families with primary familial and congenital polycythemia and Chuvash polycythemia, respectively. Here we performed genetic analysis for a family with congenital secondary polycythemia.

Design and Methods. Linkage analysis was performed using microsatellite markers and Affymetrix GeneChips. Top candidate genes were studied in detail. EPO gene expression was measured under normoxia and hypoxia conditions.

Results. We identified 5 genetic regions containing 325 genes that co-segregated with erythrocytosis in the pedigree. Top candidate genes in the regions and genes involved in erythropoiesis and oxygen sensing pathway were sequenced but no mutation was found. Cells from the Burst forming unit-Erythroid colonies showed higher EPO mRNA levels in hypoxia condition in 3 out of 4 affected family members.

Conclusions and interpretations. Genetic linkage study revealed genetic regions and possible genes linked with the polycythemia, however genetic lesions causing the polycythemia and associated phenotypes in this family remain unknown.

Introduction

Polycythemia is a condition with a net increased red blood cell mass. It can be classified into primary and secondary polycythemia based on its pathogenesis. In primary cases, the overproduction of red blood cells is due to an innate defect to the erythroid progenitor cells resulting in abnormal response to circulating hormones, mostly erythropoietin (EPO). In contrast, secondary polycythemias are caused by external factors, which act on normal erythroid progenitors.

Both primary and secondary polycythemias may arise as acquired or congenital. Congenital polycythemias are primary when inherited mutations lead to increased responsiveness of red blood cell precursors to EPO. Primary familial and congenital polycythemia (PFCP), an autosomal dominant familial disorder, is a well-studied example. Numerous mutations in the endoplasmic domain of the EPO receptor gene (EPOR) have been described in pedigrees with PFCP, which all result in truncations of the EPOR and thus loss of the negative regulatory domain of the protein.[139] The retained positive domain is reinforced and associated with JAK2/STAT5 proteins to stimulate erythropoiesis.[177] None of these mutations were found in sporadic patients,[141,142] and EPOR was excluded as the disease-causing gene in some PFCP families.[145]

Chuvash polycythemia, on the contrary, is secondary to a defect in the oxygen sensing pathway due to germline mutations in the von Hippel-Lindau tumor suppressor (VHL) gene. As the only known endemic congenital polycythemia, Chuvash polycythemia is an autosomal recessive disorder caused by a homozygous mutation in VHL gene.[178] The main regulator of oxygen homeostasis is hypoxia-inducible factor (HIF1), which is a transcription factor regulating many downstream target genes such as EPO, vascular endothelial growth factor (VEGF), transferring (TF), transferring receptor (TFRC), etc.[179] HIF1 is a heterodimer of two subunits, HIF1α and HIF1β.[180] Only HIF1α can be regulated by oxygen level. At normoxia HIF1α level is maintained low through ubiquitin-mediated degradation of the protein.[181] Hypoxia condition reduces the degradation rate of HIF1α protein, thereby increasing the cellular protein level of HIF1α

and stimulating transcription of target genes. [114] The polyubiquitination of HIF1α requires VHL protein and proline hydroxylase (PHD) activity.[118,119] A missense mutation in the VHL gene in Chuvash polycythemia impairs the interaction between VHL and HIF1α and therefore reduces the ubiquitination and degradation of HIF1α protein. The excessive HIF1α protein induces the transcription of downstream genes including EPO.[137] A large portion of congenital polycythemia is due to altered oxygen sensing, but does not carry any VHL gene mutation. Genes linked to oxygen sensing pathway became top candidates for mutation screening in these polycythemias. A missense mutation in PHD2 gene has been described as the cause for an autosomal dominant inherited polycythemia in a family.[138]

In this study, we analyzed a family with congenital secondary polycythemia. Although similar to Chuvash polycythemia to some extent such as the autosomal recessive inheritance pattern and high serum EPO levels, this family has some unique clinical and laboratory features. Five out of 6 siblings are affected with unaffected parents and 3 unaffected persons in the third generation. The number of affected individuals is against the odds as a recessive trait, possibly surggesting undiscovered consanguinity or incomplete penetrance in this family. In addition to high red blood cell counts, 3 affect males also have low platelet numbers. Genome-wide linkage mapping was performed and 5 segregating regions were identified. Interestingly, the VHL gene is located in one of the regions. Sequencing and expression analysis did not reveal any variation in the VHL gene. Several other candidate genes in the mapped regions as well as genes associated with oxygen sensing and erythropoiesis were sequenced but none of them show polymorphisms linked to the polycythemia phenotype in this family. Burst forming units-Eryroid (BFU-E) cultures in hypoxia condition showed high levels of EPO mRNA in 3 out of 4 affected family members, suggesting a potential unknown defect in the oxygen sensing pathway. Genetic lesions causing the polycythemia and associated phenotypes in this family remain to be clarified.

Design and methods

Patients, clinical and laboratory features

Five of 6 siblings (2 woman, twins, and 3 men) display elevated Hemoglobin levels since early years of life, ranging between 20,7-24 g/dL. Hematocrit range is 61-73,4 %. No additional affected relatives were found. Blood cell mass was high (65 mL/Kg) in the propositus (DL02), in whom organic lesions have been ruled out. All 5 affected individuals had normal leucocyte counts, and 3 male patients displayed decreased platelet counts, inversely proportional to Hemoglobin levels. All patients but one had elevated serum EPO levels (range ELISA: 9-68 U/L); all have normal pO2, Hemoglobin electrophoresis, 2,3-diphosphoglycerate (2,3-DPG) levels and p50. BFU-E cultures showed a normal pattern of Epo-dependent growth. 3 male patients had elevated serum EPO levels 24 hours after phlebotomies (5-70 mU/mL over basal values).

Blood cells separation, DNA and RNA extraction

Blood cells were separated by standard protocols using Histopaque (Sigma, St. Louis, MO, USA) gradient centrifugation. Granulocytes and peripheral blood mononuclear cells (PBMC) were collected respectively. Platelets were collected using the Sepharose (Amersham Pharmacia Biotech AB, Uppsala, Sweden) gel filtration method.[155] DNA was extracted using a standard proteinase K (Promega, Madison, WI, USA)/phenol (Fluka Chemie AG, Buchs, Switzerland) extraction protocol. RNA isolation was carried out using the TRIfast reagent (peqLab Biotechnology GmbH, Erlangen, Germany).

Genotyping

The linkage mapping set v2.5-MD10 (Applied Biosystems, Foster City, CA, USA) was used for genome-wide analysis. Genomic DNAs were PCR amplified and analyzed with an ABI 3100 genetic analyzer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocol. The PCR program was at 94ºC for 15 sec, 55ºC for 15 sec, 72ºC for 30 sec for 10 cycles, 89ºC for 15 sec, 55º for 15 sec and 72ºC for 30 sec for 20 cycles. The Affymetrix GeneChip Human Mapping 50K Xba 240 was used to genotype single nucleotide polymorphisms (SNPs) in genomic DNA of 12 family members

according to the Affymetrix GeneChip Mapping Assay Manual (Affymetrix Inc., Santa Clara, CA, USA)

Linkage analysis

Genotypes from microsatelite mapping were scored using the Genemapper software package version 3.5 (Applied Biosystems, Foster City, CA, USA) and linkage analysis was carried out by FASTLINK software package version 4.1p. The SNP calls were generated by GeneChip DNA Analysis Software. dCHIP version 2005 was used to perform parametric linkage analysis.[167] An autosomal recessive inheritance model with a 100% penetrance was assumed for both analyses.

Genomic DNA sequencing

The entire coding region including intron/exon boundaries of candidate genes was sequenced after PCR amplification of genomic DNA from one affected family member, DL04. The primer sequences are available upon request. The PCR conditions were 95ºC for 2 min, 94ºC for 30 sec, 60ºC for 30 sec and 72ºC for 1 min for 35 cycles. Sequencing was performed on an Applied Biosystems 3700 DNA sequencer (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's protocols.

BFU-E culture in hypoxia condition

PBMCs from family members as well as healthy volunteers were cultured in 35-mm Petri dishes using semisolid medium (Methocult H4431; Stem Cell Technologies Inc, Vancouver, BC, Canada) in a hypoxia chamber with 1% oxygen and 5% CO2 at 37°C for 14 days. BFU-E colonies were then harvest and checked for EPO expression.

cDNA synthesis, RT-PCR and Quantitative RT-PCR

Total RNA (2 μg) was reverse transcribed to generate cDNA using Superscript II reverse transcriptase (Life Technologies). cDNAs were PCR amplified to detect *VHL* and *EPO* expression at the condition described above. The quantitative real-time PCR primers for ribosomal protein L19 (*RPL19*), *VHL* and *EPO* were designed across exon-exon junctions. The primer sequences are available upon request. ΔCT values were derived by

subtracting the threshold cycle (C_T) values for target genes from the C_T value for RPL19, which served as an internal control.[158] Relative expression level for each reaction was calculated as $2^{-\Delta CT}$. A non-affected family member DL01 was chosen as a calibrator for calculating the ΔΔCT values.[160,182] All reactions were run in duplicates using the ABI 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA).

Western blotting

Protein lysate was denatured at 95°C for 2 minutes and separated by 10% sodium dodecyl sulfate-olyacryamide gel electrophoresis (SDS-PAGE). The gel was transferred electrophoretically onto 0.45-uM nitrocellulose membrane (Whatman GmbH, Dassel, Germany).[168] The membrane was blocked with 5% nonfat milk in 1 x PBS with 1% Tween-20 for 2 hours at room temperature and incubated overnight at 4°C with diluted primary antibody. The primary rabbit anti-VHL polyclonal antibody was kindly provided by Professor Willy Krek (Swiss Federal Institute of Technology, Zurich). The membranes were re-probed using a monoclonal antibody against human CD61 (BD Biosciences, San Jose, CA, USA) serving as a loading control.

Results

The clinical features of 5 family members with erythrocytosis are summarized (Table 1). Erythrocytosis in all of the patients was detected in their childhood. Four patients have hyperviscosity symptoms. Most of these symptoms are manageable using phlebotomy.

Genome-wide linkage analysis for 12 family members using microsatellite markers revealed 2 genetic regions co-segregating with the erythrocytosis in the family. The highest LOD score for the informative markers is 2.53. One region is on chromosome 3, in which the VHL gene is localized. Fine mapping for this region with additional microsatellite markers in close vicinity reduced the size of this region to 21 Mega bases (Mb) containing 180 genes (Figure 1).

Table 1 Clinical characteristics of 5 family members with polycythemia.

UPN	Sex	Date of Birth	Date of diagnosis	Date of last follow up	Hemo-globin 12-14 F 14-18 M g/L	Hematocrit (%)	Platelets 150-450 $x10^9$/L	WBC 3.5-10.0 $x10^9$/L	Treatments
DL02	M	1975	1977	2005	20-24	62-73	99-160	5.9-9.8	Phlebotomy
DL03	F	1982	1987	2005	19-23	58-65	117-145	5.0-7.4	Phlebotomy
DL04	F	1982	1987	2005	18-20	54-61	101-182	5.0-6.5	Phlebotomy
DL07	M	1969	1977	2005	19-22	61-66	75-113	4.6-7.6	Phlebotomy
DL11	M	1971	1977	2005	20-21	68-73	72-96	7.3-7.4	Phlebotomy

UPN, unique patient number; WBC, white blood cells; RBC, red blood cells; lowest and highest values are given

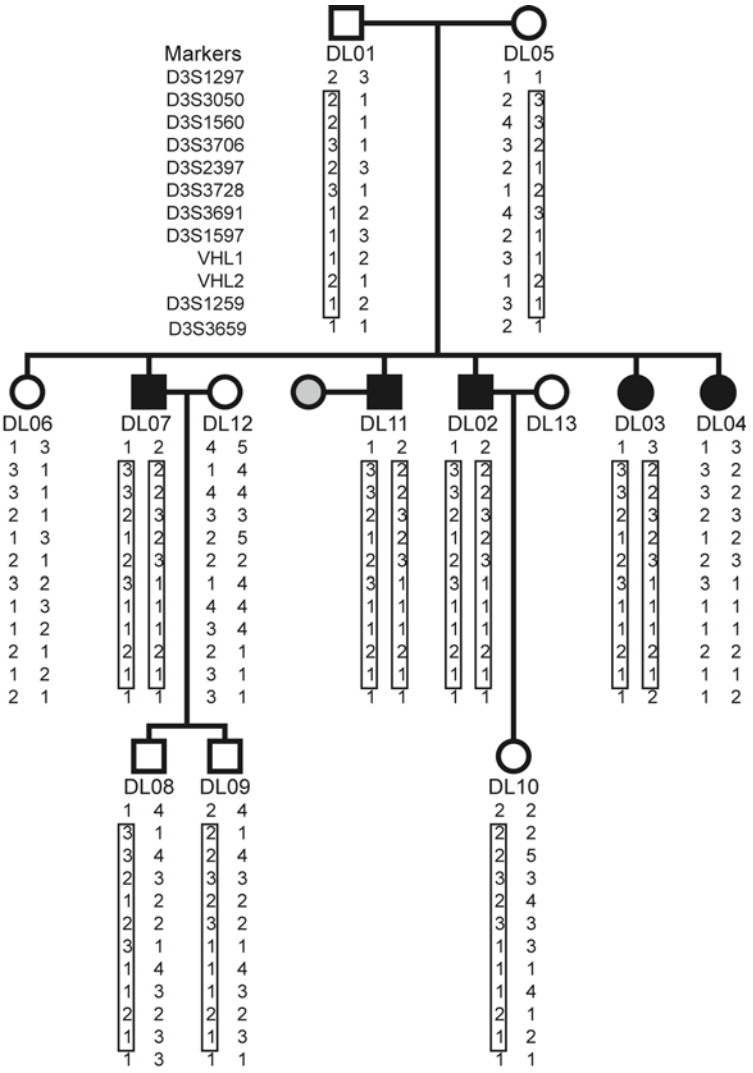

Figure 1 Segregation of microsatellite markers on chromosome 3p. Filled black symbols represent affected individuals; open symbols, normal individuals; filled grey symbols, not studied. Unique patient numbers are placed below the symbols. Haplotypes for twelve polymorphic microsatellite markers from D3S1297 to D3S3659 are shown as numbers defined arbitrarily according to different sizes of PCR fragments. Two haplotypes linked with disease phenotype are boxed. Affected individuals are homozygous and carriers are heterozygous for the disease haplotype.

The other region on chromosome 18 was 2.8 Mb in size without any obvious candidate genes (data not shown). SNP chips array analysis revealed 5 regions in total with the highest LOD score of 2.49 throughout the whole genome (Figure 2A). Among them the region on chromosome 3 (Figure 2B) and chromosome 18 were consistent with those identified by microsatellite mapping. The region on chromosome 3 was reduced to 16 Mb whereas the one on chromosome 18 was slightly larger (3.2 Mb) compared to the one from microsatellite mapping. Co-segregating regions identified by genome-wide linkage microsatelite markers mapping and SNP chip arrays are summarized in Table 2.

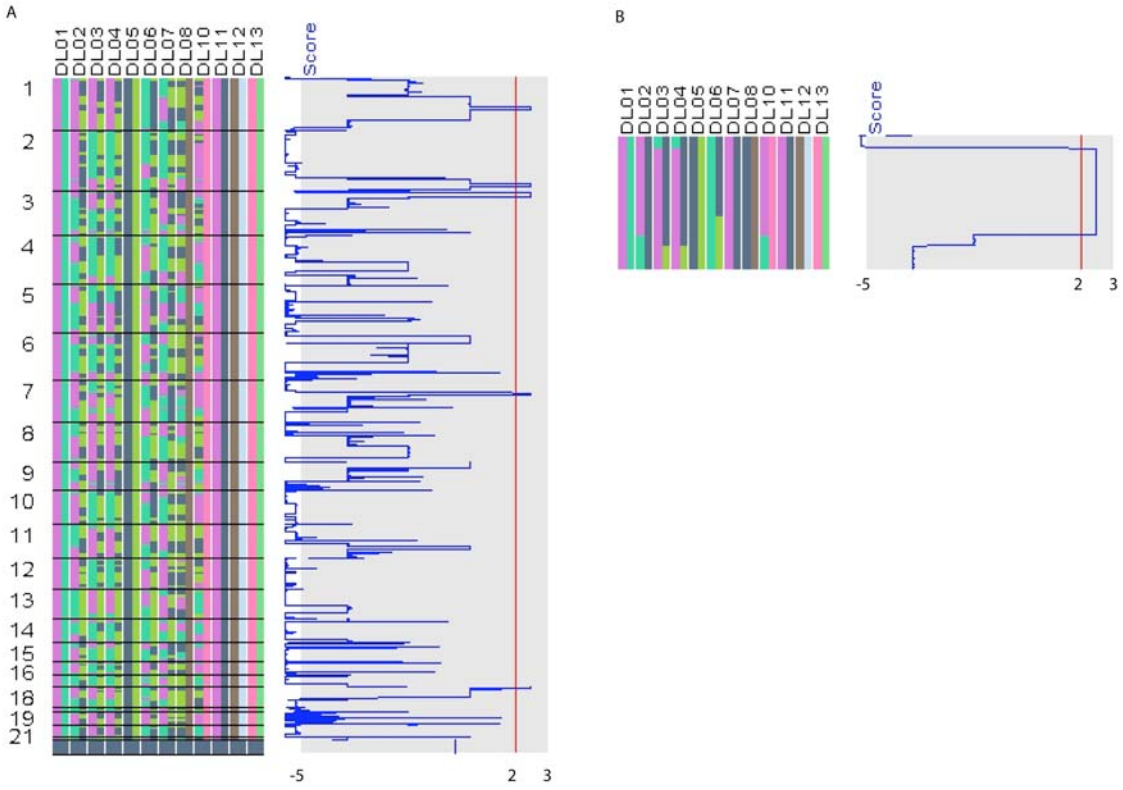

Figure 2 Parametric linkage analysis of 50K SNP array Xba 240 for 13 family members. A) The whole genome LOD score curve was plotted on the right side with a threshold from –5 to 3, LOD score of 2 is indicated by a red line. Haplotypes of each family member are shown on the left side with different colors. For each person the maternal allele is on the right and paternal on the left. **B)** LOD score curve of the segregating region on chromosome 3p. The region is consistent with the one from microsatellite markers with the highest LOD score 2.49.

Table 2 Segregating regions identified by Microsatellite linkage mapping and SNP array analysis.

Microsatellite Markers			SNP arrays		
Locus	Size (Mb)	Genes	Locus	Size (Mb)	Genes
3p	21	180	3p	16	139
18q	2.8	22	18q	3.2	29
			1q	5.7	57
			2q	8.3	49
			7p	5.7	51
Total number of genes		202			325

The top candidate gene VHL was studied intensively. The three exons including intron/exon boundaries were sequenced using genomic DNA from one affected family member (DL04). The untranslated region (UTR) was sequenced as well using cDNA as a template. No mutation was found in the sequence of VHL. Expression of VHL was examined by RT-PCR in granulocyte cDNA. Two mRNA splice variants were detectable (Figure 3A). Total VHL mRNA levels or the two splice variants determined by real-time RT-PCR showed no significant differences between normal and affected family members (Figure 3B). VHL protein in platelets was assessed by western blotting (Figure 3C). Both protein isoforms translated from the two mRNA variants were detectable in platelets. There were no visible variations for the protein sizes or expression levels between the unaffected and affect family members.

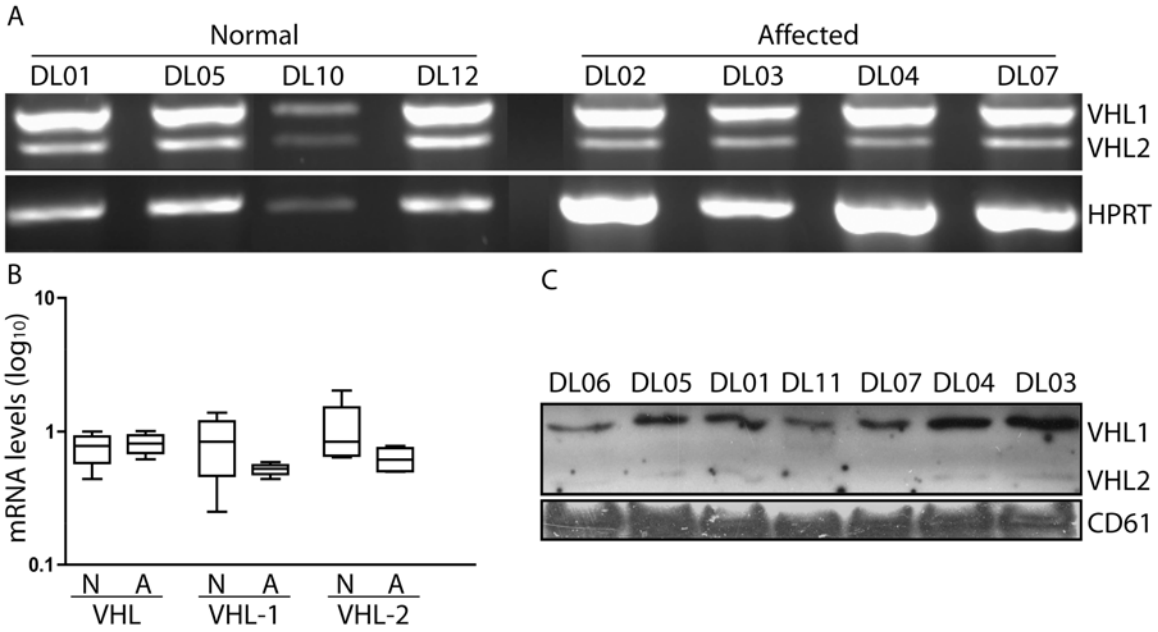

Figure 3 Analysis of VHL gene expression. A) VHL expression in platelet cDNA detected by RT-PCR. Both isoforms could be detected in all 8 family members. **B)** Expressions of total VHL and two isoforms respectively in platelet cDNA detected by real-time PCR. **C)** Western blotting of VHL protein in platelet proteins

Three other candidate genes in segregating regions were chosen for sequencing: Chromosome 3 open reading frame 10 (C3orf10) on chromosome 3, v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1 (YES-1) and TGFB-induced factor homeobox 1 (TGIF) on chromosome 18. However no mutations were found in these three genes. In addition to candidate genes in co-segregating regions, we did exon sequencing for genes known to be involved in hypoxia sensing and erythropoiesis such as HIF1α, PHD 1, 2 and 3, EPO, EPO-R and phosphatase and tesin homolog deleted (PTEN) etc. However none of the genes showed polymorphisms linked to the polycythemia in the family. Information on all sequenced candidate genes is summarized in Table 3.

Table 3 Overview of analyzed candidate genes

Gene Symbol	Gene Full Name	Position	Main Function
VHL	von Hippel-Lindau tumor suppressor	3p26-p25	tumor suppressor; Ubiquitin degradation
C3orf10	chromosome 3 open reading frame 10	3p25.3	uncharacterized hematopoietic stem/progenitor cells protein
YES-1	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	18p11.31-p11.21	tyrosine kinase activity and one of the src family of proteins
TGIF	TGFB-induced factor homeobox 1	18p11.3	transcriptional regulator, mutations associated with holoprosencephaly type 4
EPO	erythropoietin	7q22	encoding EPO protein, regulating red cell differentiation and production
EPO-R	erythropoietin-receptor	19p13.3	encoding EPO receptor, binging with EPO to activate different cellular pathways
PHD1	prolyl hydroxylase domain containing 1	19q13.2	encodes enzymes responsible for prolyl
PHD2	prolyl hydroxylase domain containing 2	1q42.1	hydroxylation of HIF-1α and ubiquitin-mediated
PHD3	prolyl hydroxylase domain containing 3	14q13	degradation
HIF-1a	hypoxia-inducible factor-1a	14q21-24	an essential transcription factor for maintaining oxygen homeostasis
PTEN	phosphatase and tesin homolog deleted	10q23.3	a tumor suppressor, negatively regulating AKT/PKB signaling pathway

EPO gene expression levels were determined in BFU-E culture cells in both normoxia and hypoxia conditions for 4 affected family members. One unaffected sibling and one normal individual were used as controls (Figure 4). In 3 out of 4 affected family members EPO mRNA levels were 4.6 to 7.2 folds higher in hypoxia condition than in normoxia culture. The unaffected siblings showed 1.4 fold increase of EPO mRNA in

hypoxia treated BFU-E culture cells. The normal control showed even lower EPO mRNA level in hypoxia condition. However one affected family member (DL07) did not show significant increase of EPO mRNA levels after hypoxia treatment (1.3 fold).

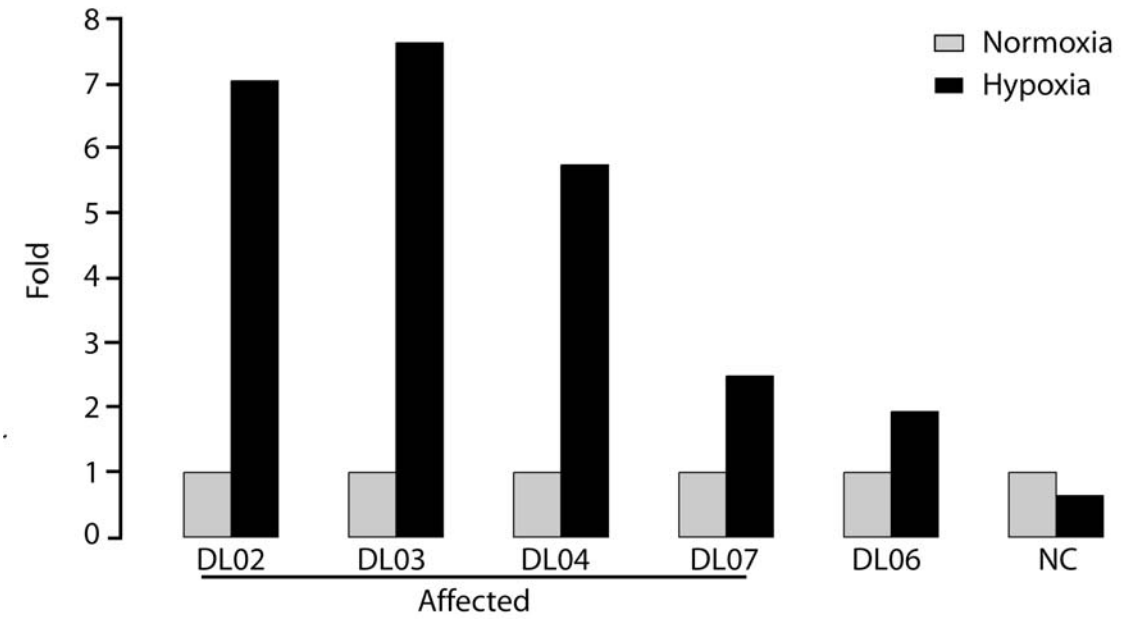

Figure 4 EPO gene expression in BFU-E culture cells in normoxia and hypoxia condition. Four affected family members, one unaffected sibling and one normal control were tested. BFU-E cultures were incubated at normoxia and hypoxia chambers for 14 days before RNA extraction and real-time PCR. EPO expression levels at normoxia are set as 1 to calibrate the expression levels at hypoxia.

Discussion

In this study we analyzed a family with congenital secondary polycythemia with the purpose to identify the genetic lesion associated with the phenotype. Genome wide mapping by microsatellite markers and SNP chip arrays revealed a total of 5 genetic regions linked with the disease, ranging from 2.8 to 21 Mb. The multiple and relatively large genetic regions in this family are mainly due to the limited size of the pedigree and a high degree of homozygosity in certain genomic regions in some family members. SNP arrays are more sensitive than microsatellite marker mapping since the density of SNPs is much higher than microsatellite markers. Indeed 3 extra regions were found in SNP arrays, whereas two regions (chromosome 3 and chromosome 18) were confirmed by both mapping methods. In addition, SNP chip arrays reduced the region on chromosome 3 by 5 Mb and therefore excluded 41 genes as candidate genes. However the region on chromosome 18 was increased by 0.4 Mb with additional 7 genes inside. These discrepancies indicate that although SNP arrays with high density of SNP markers increase resolusion in genetic linkage analysis, it is not necessarily helping in fine mapping because the reduction of the region is totally dependent on recombinations that occurred in family members.

In total, there are 325 genes in the 5 segregating regions. It is feasible to sequence a limited number of candidate genes. The most promising candidate gene within the segregating regions is VHL, in which a homozygous mutation causes the Chuvash polycythemia. We did not find any mutation by sequencing VHL. VHL mRNA and protein levels were comparable in affected and unaffected family members. Although the family has a similar phenotype and inheritance pattern like Chuvash polycythemia, VHL was excluded as the disease-causing gene for this family. Three other genes in the regions were chosen for sequencing and no mutation was found. The criteria for picking candidate genes were involvement in hematopoiesis, kinase activity or transcriptional regulation. In addtion to candidate genes in co-segregating regions, several key genes involved in oxygen sensing and erythropoiesis were also excluded by sequencing. We suppose that the undiscovered mutant gene in the segregating regions must have an uncharacterized function, which leads to the phenotype in this family.

BFU-E cultures in hypoxia showed higher EPO mRNA expression in 3 out of 4 affected family members compared to normal controls indicating a hypersensitivity for hypoxia condition in patients. There might be a defect in oxygen sensing that contributes to the erythrocytosis in this family.

PERSPECTIVES

Genetic studies of familial MPD have identified numerous inherited germ-line mutations in disease causing genes such as TPO, MPL, EPOR and so on. The characterizations of these mutations have helped us to understand not only the etiology of the inherited diseases, but also the physiological processes of erythropoiesis and thrombopoiesis. In many families, however, the disease causing genes remain to be identified to clarify the mechanism underlying the inherited phenotypes, which are often similar to the sporadic form of MPD. Some of these unknown genetic lesions in families are supposed to facilitate or predispose for one or several somatic mutations that cause the sporadic cases of MPD. In the first part of my thesis project, I studied HT families by both candidate gene approach and genome-wide linkage analysis. Previously reported mutations in candidate genes TPO and MPL have been identified in two families of our HT collection. One mutation arose *de novo*, the other from a founder effect. TPO and MPL gene have been excluded in 11 HT families and the genetic variation causing the thrombocytosis inherited in these families still remains unknown. Genome-wide linkage analysis was performed for a large HT pedigree containing 20 family members to search for new genetic components. In a second project, a family with congenital secondary polycythemia was analyzed by both genome-wide linkage and candidate gene approach. Several genetic regions were shown to co-segregate with the disease, but no mutation was found.

Mutations causing hereditary thrombocythemia

TPO and MPL gene mutations

Five germ-line mutations in the *TPO* gene and one in the *MPL* gene have been identified in HT families.[22,23,25-27,51] The mechanisms causing thrombocytosis are different for the TPO and MPL mutations though they both cause the disease via the TPO and MPL ligand-receptor mediated pathway. TPO mutations are located

in the 5'UTR of the *TPO* mRNA and increase the efficiency of TPO protein translation.[21] The *MPL* mutation is in the transmembrane domain of the receptor and results in a hyperactive protein independent of TPO ligation.[51] Families with *TPO* germ-line mutations, despite increased translation and excessive production of TPO protein, show only slightly elevated or even normal serum TPO levels.[22] A possible explanation is that the increase in platelet and megakaryocyte mass activates the feedback regulation system and reaches a new equilibrium at TPO levels close to normal.(Liu *et al* submitted) Therefore the TPO serum level is not necessarily suitable as an indicator for a *TPO* gene mutation. The clinical features of different HT families carrying either *TPO* or *MPL* mutation are indistinguishable. HT, therefore, is a disease with uniform clinical features but heterogeneous molecular genetic causes.

No mutation in the TPO gene has been found in sporadic ET patients, whereas mutations in juxtamembrane domain of MPL have been found in sporadic IMF and ET patients.[54,55]

Due to the multiple genetic variations for different HT families, combination of many small pedigrees for linkage analysis is not feasible since they may have different genetic locations for the disease mutations. For many small HT families, a candidate gene approach by analyzing TPO and MPL gene is still the best way to explore their etiology.

Origins of the mutations

Recently families carrying the identical TPO or MPL mutations have been reported, raising the intriguing question of the origin of these rare germ-line mutations in these families (Liu *et al*. submitted).[52]

One splice donor mutation in the TPO gene has been identified both in a Dutch family and in a Polish family. Although a very rare event, our data suggest that the mutation arose independently in these two families (Liu et al. submitted). Recently a mutation in the transmembrane domain of the MPL gene has been found in 4 HT families.[52] This is the first report of multiple families sharing an identical germ-line mutation. In our HT family group of 11 families, one family

was found carrying the same MPL mutation. The unexpectedly high frequency of this mutation in HT families suggested a founder effect, i.e. all families inherited this mutation from a common ancestor. All the six HT families carrying the identical MPL mutation were examined by haplotype analysis using microsatellites. This analysis revealed that all the families share the same genetic characters in the vicinity to the mutation, indicating that the mutant allele in these six families may share a common ancestor. It is reasonable to assume that more HT families in the general population may carry the same MPL or TPO germ-line mutations that either originated de novo or from a founder effect.

Novel genetic components

In addition to the TPO and MPL gene, unknown disease causing gene(s) remain to be identified in some HT families.[143,144] Genome-wide linkage study is only feasible when the HT pedigree is large enough with a clear inheritance mode and relatively high penetrance. These families provide powerful tools to track down the genetic variations causing the disease, since the genetic regions co-segregating with the disease can be located using genome-wide linkage mapping methods. By examining genes in the co-segregating region, it is possible to identify the novel genetic components involved in the disease.

Genetic regions identified by microsatellites and SNPs

A large HT family with normal TPO and MPL gene was chosen for genome-wide linkage analysis to identify a novel genetic variation causing the thrombocytosis in this family. Two different regions on chromosome 9 and 20 displayed co-segregation with the phenotype. The region on chromosome 9 was confirmed using both microsatellite markers and SNP chip arrays, whereas the region on chromosome 20 was only detectable in SNP chip arrays. The discrepancies might be explained by the much higher density of the SNPs than the microsatellite markers, which enables us to discover regions falling between two microsatellites or those without informative microsatellite markers located within. Similarly, in the linkage analysis for a family with congenital secondary

polycythemia, three additional genetic regions have been identified when using SNP arrays. The high density of SNPs increase the power remarkably in identifying genetic regions but the power in reducing the size of genetic regions is limited since it is mainly determined by the recombinations during meiosis within the pedigree. Among all the genetic regions that show linkage to the disease phenotype, only one region contains the disease-causing gene. The reasons for multiple regions include coincidental segregations and uninformative markers within certain regions. The genetic regions confirmed by two methods are more reliable and should have priority for further studies.

Functional studies on the candidate gene gelsolin

All the 29 genes in the region of chromosome 9 have been sequenced to search for the novel genetic component causing thrombocytosis in this family. A novel polymorphism in the gelsolin gene was found, which changed a glycine into a cysteine in the protein sequence. However cell lines transfected with the candidate mutant gelsolin did not show any proliferation advantage over the wild type. Blood counts were normal in mice transplanted with the candidate mutant gelsolin. These results do not support the hypothesis that gelsolin is the disease-causing gene in this family, even though the polymorphism in this gene is the only variation found so far in the segregating region.

Further functional studies on gelsolin are planned for the future. The first aim is to study the process of platelet releasing in megakaryocytes cultured in vitro. Since our data showed that over expression of gelsolin did not increase the proliferation rate of cells, the high platelet counts might be due to a defect in the platelet releasing process, which remains poorly understood to date. Platelet formation and releasing from megakaryocytes requires complicated changes in the cytoskeletal organization including microtubule elongation and actin assembling.[66,176] Actin is the most abundant protein in platelets. Gelsolin, as a member of the actin-binding protein family, is thought to function in capping, severing, nucleating or bundling actin filaments. Expression of gelsolin protein was detected in platelets in our study and many other studies. Knock-out of

gelsolin in mice showed abnormal platelet functions.[169] Interestingly, abnormal platelet aggregation and prolonged bleeding time were also found in affected family members studied.[166] Therefore it is of interest to see if the candidate mutant gelsolin could lead to over-production of platelets or abnormal platelet functions. Secondly, bone marrow transplantation using mouse gelsolin sequences might be necessary since human and mouse glesolin have 5% difference at amino acid sequence. Human gelsolin might not be able to interact with mouse signaling molecules in the bone marrow transplantation assay. Due to the relatively low efficiency of retrovirus integration in bone marrow transplantation (up to 30%), it might be difficult to mimic a germ-line mutation, which is present in all cells, to generate a disease phenotype in mice. An alternative method for in vivo study could be to generate transgenic mice or knock-in micefor the candidate gelsolin mutation.

Sequencing candidate genes

Although gelsolin is considered as a top candidate gene in this HT family with normal TPO and MPL genes, it is possible that another gene might carry the disease-causing mutation. Sequencing for genes in the segregating region on chromosome 9q has not been fully completed due to some difficulties with some PCR. In addition, only 5 candidate genes in the region on chromosome 20q identified by SNP chips array were sequenced. There are still 20 genes in this area that remain to be sequenced.

For the family with congenital secondary polycythemia, there are a total of 325 genes in the 5 segregating regions identified by microsatellite markers and SNP arrays. It is only feasible to check a limited number of candidate genes. Combination with expression profile data might be helpful for choosing more candidate genes for further study.

REFERENCES

1. Kondo M WA, Manz MG, Prohaska SS, Scherer DC, Beilhack GF, Shizuru JA, Weissman IL. Biology of hematopoietic stem cells and progenitors: implications for clinical application. Annu Rev Immunol. 2003;21:759-806.
2. Akashi K TD, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature. 2000;404:193-197.
3. Kondo M WI, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. Cell. 1997;91:661-672.
4. Kaushansky K DJ. The molecular and cellular biology of thrombopoietin: the primary regulator of platelet production. Oncogene. 2002;21:3359-3367.
5. Gurney AL, Carver-Moore K, de Sauvage FJ, Moore MW. Thrombocytopenia in c-mpl-deficient mice. Science. 1994;265:1445-1447.
6. de Sauvage FJ, Carver-Moore K, Luoh SM, et al. Physiological regulation of early and late stages of megakaryocytopoiesis by thrombopoietin. J Exp Med. 1996;183:651-656.
7. Wendling F, Varlet P, Charon M, Tambourin P. MPLV: a retrovirus complex inducing an acute myeloproliferative leukemic disorder in adult mice. Virology. 1986;149:242-246.
8. Souyri M, Vigon I, Penciolelli JF, Heard JM, Tambourin P, Wendling F. A putative truncated cytokine receptor gene transduced by the myeloproliferative leukemia virus immortalizes hematopoietic progenitors. Cell. 1990;63:1137-1147.
9. Vigon I, Mornon JP, Cocault L, et al. Molecular cloning and characterization of MPL, the human homolog of the v-mpl oncogene: identification of a member of the hematopoietic growth factor receptor superfamily. Proc Natl Acad Sci USA. 1992;89:5640-5644.
10. Vigon I, Florindo C, Fichelson S, et al. Characterization of the murine Mpl proto-oncogene, a member of the hematopoietic cytokine receptor family: molecular cloning, chromosomal location and evidence for a function in cell growth. Oncogene. 1993;8:2607-2615.
11. Skoda RC, Seldin DC, Chiang MK, Peichel CL, Vogt TF, Leder P. Murine c-mpl: a member of the hematopoietic growth factor receptor superfamily that transduces a proliferative signal. Embo J. 1993;12:2645-2653.
12. de Sauvage FJ, Hass PE, Spencer SD, et al. Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-mpl ligand. Nature. 1994;369:533-538.
13. Bartley TD, Bogenberger J, Hunt P, et al. Identification and cloning of a megakaryocyte growth and development factor that is a ligand for the cytokine receptor mpl. Cell. 1994;77:1117-1124.
14. Lok S, Kaushansky K, Holly RD, et al. Cloning and expression of murine thrombopoietin cDNA and stimulation of platelet production in vivo. Nature. 1994;369:565-568.
15. Kaushansky K, Lok S, Holly RD, et al. Promotion of megakaryocyte progenitor expansion and differentiation by the c-mpl ligand thrombopoietin. Nature. 1994;369:568-571.
16. Foster DC, Sprecher CA, Grant FJ, et al. Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal localization. Proceedings of the National Academy of Sciences of the United States of America. 1994;91:13023-13027.
17. Chang MS, Mcninch J, Basu R, et al. Cloning and characterization of the human megakaryocyte growth and development factor (mgdf) gene. Journal of Biological Chemistry. 1995;270:511-514.
18. Sohma Y, Akahori H, Seki N, et al. Molecular cloning and chromosomal localization of the human thrombopoietin gene. Febs Letters. 1994;353:57-61.
19. Kamura T, Handa H, Hamasaki N, Kitajima S. Characterization of the human thrombopoietin gene promoter - A possible role of an Ets transcription factor, E4TF1/GABP. J Biol Chem. 1997;272:11361-11368.
20. Kozak M. Constraints on reinitiation of translation in mammals. Nucleic Acid Res. 2001;29:5226-5232.
21. Cazzola M, Skoda RC. Translational pathophysiology: a novel molecular mechanism of human disease. Blood. 2000;95:3280-3288.

22. Wiestner A, Schlemper RJ, van der Maas AP, Skoda RC. An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. Nature Genet. 1998;18:49-52.
23. Ghilardi N, Wiestner A, Kikuchi M, Oshaka A, Skoda RC. Hereditary thrombocythemia in a Japanese family is caused by a novel point mutation in the thrombopoietin gene. Br J Haematol. 1999;107:310-316.
24. Ghilardi N, Wiestner A, Skoda RC. Thrombopoietin production is inhibited by a translational mechanism. Blood. 1998;92:4023-4030.
25. Kondo T, Okabe M, Sanada M, et al. Familial essential thrombocythemia associated with one-base deletion in the 5'-untranslated region of the thrombopoietin gene. Blood. 1998;92:1091-1096.
26. Ghilardi N, Skoda RC. A single-base deletion in the thrombopoietin (TPO) gene causes familial essential thrombocytosis through a mechanism of more efficient translation of TPO mRNA [letter]. Blood. 1999;94:1480-1482.
27. Jorgensen MJ, Raskind WH, Wolff JF, Bachrach HR, Kaushansky K. Familial thrombocytosis associated with overproduction of thrombopoietin due to a novel splice donor site mutation. Blood. 1998;92, supplement 1:205a (abstract).
28. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 1987;15:8125-8148.
29. Kaushansky K. Thrombopoietin. N Engl J Med. 1998;339:746-754.
30. Wells JA, de Vos AM. Hematopoietic receptor complexes. Annu Rev Biochem. 1996;65:609-634.
31. Fielder PJ, Gurney AL, Stefanich E, et al. Regulation of thrombopoietin levels by c-mpl-mediated binding to platelets. Blood. 1996;87:2154-2161.
32. Fielder PJ, Hass P, Nagel M, et al. Human platelets as a model for the binding and degradation of thrombopoietin. Blood. 1997;89:2782-2788.
33. Kuter DJ, Rosenberg RD. The reciprocal relationship of thrombopoietin (c-mpl ligand) to changes in the platelet mass during busulfan-induced thrombocytopenia in the rabbit. Blood. 1995;85:2720-2730.
34. Stoffel R, Wiestner A, Skoda RC. Thrombopoietin in thrombocytopenic mice: evidence against regulation at the mRNA level and for a direct regulatory role of platelets. Blood. 1996;87:567-573.
35. McCarty JM, Sprugel KH, Fox NE, Sabath DE, Kaushansky K. Murine thrombopoietin mRNA levels are modulated by platelet count. Blood. 1995;86:3668-3675.
36. Kaser A BG, Steurer W, Kaser S, Offner FA, Zoller H, Theurl I, Widder W, Molnar C, Ludwiczek O, Atkins MB, Mier JW, Tilg H. Interleukin-6 stimulates thrombopoiesis through thrombopoietin: role in inflammatory thrombocytosis. Blood. 2001;98:2720-2725.
37. Mignotte V, Vigon I, Decrevecoeur EB, Romeo PH, Lemarchandel V, Chretien S. Structure and transcription of the human c-mpl gene (MPL). Genomics. 1994;20:5-12.
38. Li J SD, Kuter DJ. Cloning and functional characterization of a novel c-mpl variant expressed in human CD34 cells and platelets. Cytokine. 2000;12:835-844.
39. Coers J, Ranft C, Skoda RC. A Truncated Isoform of c-Mpl with an Essential C-terminal Peptide Targets the Full-length Receptor for Degradation. J Biol Chem. 2004;279:36397-36404.
40. Baxter EJ, Scott LM, Campbell PJ, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet. 2005;365:1054-1061.
41. James C, Ugo V, Le Couedic JP, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. Nature. 2005;434:1144-1148.
42. Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. N Engl J Med. 2005;352:1779-1790.
43. Zhao R, Xing S, Li Z, et al. Identification of an acquired JAK2 mutation in polycythemia vera. J Biol Chem. 2005;280:22788-22792.
44. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005;7:387-397.

45. Scott LM TW, Levine RL, Scott MA, Beer PA, Stratton MR, Futreal PA, Erber WN, McMullin MF, Harrison CN, Warren AJ, Gilliland DG, Lodish HF, Green AR. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. N Engl J Med. 2007;356:459-468.
46. Saharinen P TK, Silvennoinen O. Regulation of the Jak2 tyrosine kinase by its pseudokinase domain. Mol Cell Biol. 2000;20:3387-3395.
47. Ihara K, Ishii E, Eguchi M, et al. Identification of mutations in the c-mpl gene in congenital amegakaryocytic thrombocytopenia. Proc Natl Acad Sci U S A. 1999;96:3132-3136.
48. van den Oudenrijn S BM, Folman CC, Peters M, Faulkner LB, de Haas M, von dem Borne AE. Mutations in the thrombopoietin receptor, Mpl, in children with congenital amegakaryocytic thrombocytopenia. Br J Haematol. 2000;110:441-448.
49. Ballmaier M GM, Schulze H, Cherkaoui K, Lang S, Gaudig A, Krukemeier S, Eilers M, Strauss G, Welte K. c-mpl mutations are the cause of congenital amegakaryocytic thrombocytopenia. Blood. 2001;97:139-146.
50. Ballmaier M GM, Krukemeier S, Welte K. Thrombopoietin is essential for the maintenance of normal hematopoiesis in humans: development of aplastic anemia in patients with congenital amegakaryocytic thrombocytopenia. Ann N Y Acad Sci. 2003;996:17-25.
51. Ding J, Komatsu H, Wakita A, et al. Familial essential thrombocythemia associated with a dominant-positive activating mutation of the c-MPL gene, which encodes for the receptor for thrombopoietin. Blood. 2004.
52. Teofili L GF, Martini M, Cenci T, Guidi F, Torti L, Palumbo G, Amendola A, Foà R, Larocca LM. Markers of myeloproliferative diseases in childhood polycythemia vera and essential thrombocythemia. J Clin Oncol. 2007;25:1048-1053.
53. Onishi M, Mui ALF, Morikawa Y, et al. Identification of an oncogenic form of the thrombopoietin receptor MPL using retrovirus-mediated gene transfer. Blood. 1996;88:1399-1406.
54. Pardanani AD, Levine RL, Lasho T, et al. MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. Blood. 2006;108:3472-3476.
55. Pikman Y, Lee BH, Mercher T, et al. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. PLoS Med. 2006;3:e270.
56. Zauli G CL. Human megakaryocyte biology and pathophysiology. Crit Rev Oncol Hematol. 1995;21:135-157.
57. Kuter DJ GD, Rosenberg RD. Transforming growth factor beta inhibits megakaryocyte growth and endomitosis. Blood. 1992;79:619-626.
58. Gandhi MJ DJ, Reems JA, Thorning D, Lannutti BJ. A novel strategy for generating platelet-like fragments from megakaryocytic cell lines and human progenitor cells. Blood Cells Mol Dis. 2005;35:70-73.
59. Lannutti BJ BN, Gandhi MJ, Reems JA, Drachman JG. Induction of polyploidization in leukemic cell lines and primary bone marrow by Src kinase inhibitor SU6656. Blood. 2005;105:3875-3878.
60. Crispino J. GATA1 in normal and malignant hematopoiesis. Semin Cell Dev Biol. 2005;16:137-147.
61. Shivdasani RA FY, McDevitt MA, Orkin SH. A lineage-selective knockout establishes the critical role of transcription factor GATA-1 in megakaryocyte growth and platelet development. Embo J. 1997;16:3965-3973.
62. Nichols KE CJ, Poncz M, White JG, Orkin SH, Maris JM, Weiss MJ. Familial dyserythropoietic anaemia and thrombocytopenia due to an inherited mutation in GATA1. Nat Genet. 2000;24:266-270.
63. Muntean AG CJ. Differential requirements for the activation domain and FOG-interaction surface of GATA-1 in megakaryocyte gene expression and development. Blood. 2005;106:1223-1231.
64. Elagib KE RF, Mogass M, Khetawat R, Delehanty LL, Goldfarb AN. RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation. Blood. 2003;101:4333-4341.
65. Heller PG GA, Gandhi MJ, Cummings CL, Pirola CJ, Marta RF, Kornblihtt LI, Drachman JG, Molinas FC. Low Mpl receptor expression in a pedigree with familial platelet disorder with predisposition to acute myelogenous leukemia and a novel AML1 mutation. Blood. 2005;105:4664-4670.

66. Lecine P IJ, Kim SW, Villeval JL, Shivdasani RA. Hematopoietic-specific beta 1 tubulin participates in a pathway of platelet biogenesis dependent on the transcription factor NF-E2. Blood. 2000;96:1366-1373.
67. Deveaux S C-KS, Shivdasani RA, Andrews NC, Filipe A, Kuzniak I, Orkin SH, Roméo PH, Mignotte V. p45 NF-E2 regulates expression of thromboxane synthase in megakaryocytes. Embo J;16:5654-5661.
68. Tiwari S IJ, Barral DC, Mules EH, Novak EK, Swank RT, Seabra MC, Shivdasani RA. A role for Rab27b in NF-E2-dependent pathways of platelet formation. Blood. 2003;102:3970-3979.
69. Schulze H KM, Bergmeier W, Italiano JE, Wahl SM, Shivdasani RA. Interactions between the megakaryocyte/platelet-specific beta1 tubulin and the secretory leukocyte protease inhibitor SLPI suggest a role for regulated proteolysis in platelet functions. Blood. 2004.
70. Lecine P, Villeval JL, Vyas P, Swencki B, Xu Y, Shivdasani RA. Mice lacking transcription factor NF-E2 provide in vivo validation of the proplatelet model of thrombocytopoiesis and show a platelet production defect that is intrinsic to megakaryocytes. Blood. 1998;92:1608-1616.
71. Onodera K SJ, Motohashi H, Yamamoto M, Engel JD. Perinatal synthetic lethality and hematopoietic defects in compound mafG:mafK mutant mice. Embo J. 2000;19:1335-1345.
72. Ravid K LJ, Zimmet JM, Jones MR. Roads to polyploidy: the megakaryocyte example. J Cell Physiol. 2002;190:7-20.
73. Nagata Y, Muro Y, Todokoro K. Thrombopoietin-induced polyploidization of bone marrow megakaryocytes is due to a unique regulatory mechanism in late mitosis. J Cell Biol. 1997;139:449-457.
74. Vitrat N, Cohen-Solal K, Pique C, et al. Endomitosis of human megakaryocytes are due to abortive mitosis. Blood. 1998;91:3711-3723.
75. Wright J. The origin and nature of blood platelets. Boston Med Surg J. 1906;154:643-645.
76. Kaushansky K. Thrombopoietin: a tool for understanding thrombopoiesis. J Thromb Haemost. 2003;1:1587-1592.
77. Choi ES NJ, Hokom MM, Hornkohl AC, Hunt P. Platelets generated in vitro from proplatelet-displaying human megakaryocytes are functional. Blood. 1995;85:402-413.
78. Cramer EM NF, Guichard J, Breton-Gorius J, Vainchenker W, Massé JM, Debili N. Ultrastructure of platelet formation by human megakaryocytes cultured with the Mpl ligand. Blood. 1997;89:2336-2346.
79. 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;147:1299-1312.
80. Schwer HD LP, Tiwari S, Italiano JE, 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;11:579-586.
81. Freson K DVR, Wittevrongel C, Thys C, Defoor J, Vanhees L, Vermylen J, Peerlinck K, Van Geet C. The TUBB1 Q43P functional polymorphism reduces the risk of cardiovascular disease in men by modulating platelet function and structure. Blood. 2005;106:2356-2362.
82. Balduini CL IA, Savoia A. Inherited thrombocytopenias: from genes to therapy. Haematologica. 2002;87:860-880.
83. Hamilton RW SB, Ottie JN, Storch AE, Saleem A, White JG. Platelet function, ultrastructure, and survival in the May-Hegglin anomaly. Am J Clin Pathol. 1980;74:665-668.
84. Canobbio I NP, Pecci A, Balduini A, Balduini CL, Torti M. Altered cytoskeleton organization in platelets from patients with MYH9-related disease. J Thromb Haemost. 2005;3:1026-1035.
85. Léon C EA, Hechler B, Aleil B, Freund M, Ravanat C, Jourdain M, Nonne C, Weber J, Tiedt R, Gratacap MP, Severin S, Cazenave JP, Lanza F, Skoda R, Gachet C. Megakaryocyte-restricted MYH9 inactivation dramatically affects hemostasis while preserving platelet aggregation and secretion. Blood. 2007;110:3183-3191.
86. Wu H, Liu X, Jaenisch R, Lodish HF. Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. Cell. 1995;83:59-67.
87. Bonsdorff EJ, E. A humoral mechanism in anoxic erythrocytosis. Acta Physiol Scand. 1948;16:150-170.

88. Goldwasser E. Erythropoietin: a somewhat personal history. Perspect Biol Med. 1996;40:18-32.
89. Miyake T KC, Goldwasser E. Purification of human erythropoietin. J Biol Chem. 1977;252:5558-5564.
90. Jacobs K SC, Rudersdorf R, Neill SD, Kaufman RJ, Mufson A, Seehra J, Jones SS, Hewick R, Fritsch EF, et al. Isolation and characterization of genomic and cDNA clones of human erythropoietin. Nature. 1985;313:806-810.
91. Lin FK SS, Lin CH, Browne JK, Smalling R, Egrie JC, Chen KK, Fox GM, Martin F, Stabinsky Z, et al. Cloning and expression of the human erythropoietin gene. Proc Natl Acad Sci U S A. 1985;82:7580-7584.
92. Powell JS BK, Lebo RV, Adamson JW. Human erythropoietin gene: high level expression in stably transfected mammalian cells and chromosome localization. Proc Natl Acad Sci U S A. 1986;83:6465-6469.
93. Watkins PC ER, Hoffman N, Stanislovitis P, Beck AK, Galli J, Vellucci V, Gusella JF, Shows TB. Regional assignment of the erythropoietin gene to human chromosome region 7pter----q22. Cytogenet Cell Genet. 1986;42:214-218.
94. Law ML CG, Lin FK, Wei Q, Huang SZ, Hartz JH, Morse H, Lin CH, Jones C, Kao FT. Chromosomal assignment of the human erythropoietin gene and its DNA polymorphism. Proc Natl Acad Sci U S A. 1986;83:6920-6924.
95. Recny MA SH, Kim Y. Structural characterization of natural human urinary and recombinant DNA-derived erythropoietin. Identification of des-arginine 166 erythropoietin. J Biol Chem. 1987;262:17156-17163.
96. Blanchard KL AA, Galson DL, Bunn HF. Hypoxic induction of the human erythropoietin gene: cooperation between the promoter and enhancer, each of which contains steroid receptor response elements. Mol Cell Biol. 1992;12:5373-5385.
97. Lee-Huang S LJ, Kung HF, Huang PL, Lee L, Huang PL. The human erythropoietin-encoding gene contains a CAAT box, TATA boxes and other transcriptional regulatory elements in its 5' flanking region. Gene. 1993;128:227-236.
98. Dame C SM, Lim KC, Leach KM, Fandrey J, Ma Y, Knöpfle G, Engel JD, Bungert J. Hepatic erythropoietin gene regulation by GATA-4. J Biol Chem. 2004;279:2995-2961.
99. Imagawa S YM, Miura Y. Negative regulation of the erythropoietin gene expression by the GATA transcription factors. Blood. 1997;89:1430-1439.
100. La Ferla K RC, Jelkmann W, Hellwig-Bürgel T. Inhibition of erythropoietin gene expression signaling involves the transcription factors GATA-2 and NF-kappaB. FASEB J. 2002;16:1811-1813.
101. Nakano Y IS, Matsumoto K, Stockmann C, Obara N, Suzuki N, Doi T, Kodama T, Takahashi S, Nagasawa T, Yamamoto M. Oral administration of K-11706 inhibits GATA binding activity, enhances hypoxia-inducible factor 1 binding activity, and restores indicators in an in vivo mouse model of anemia of chronic disease. Blood. 2004;104:4300-4307.
102. White JR HR, Lee SR, Craigon MH, Binley K, Price T, Beard GL, Mundy CR, Naylor S. Genetic amplification of the transcriptional response to hypoxia as a novel means of identifying regulators of angiogenesis. Genomics. 2004;83:1-8.
103. Manalo DJ RA, Lavoie T, Natarajan L, Kelly BD, Ye SQ, Garcia JG, Semenza GL. Transcriptional regulation of vascular endothelial cell responses to hypoxia by HIF-1. Blood. 2005;105:659-669.
104. Allen JW KS, Johnson RS, Bhatia SN. In vitro liver tissue model established from transgenic mice: role of HIF-1alpha on hypoxic gene expression. Tissue Eng. 2006;12:3135-3147.
105. Chang SC SD, Goldwasser E. Evidence for an erythropoietin receptor protein on rat bone marrow cells. Biochem Biophys Res Commun. 1974;57:399-405.
106. Krantz SB GE. Specific binding of erythropoietin to spleen cells infected with the anemia strain of Friend virus. Proc Natl Acad Sci U S A. 1984;81:7574-7578.
107. D'Andrea AD, Zon LI. Erythropoietin receptor. Subunit structure and activation. Journal of Clinical Investigation. 1990;86:681-687.
108. Youssoufian H LG, Neumann D, Yoshimura A, Lodish HF. Structure, function, and activation of the erythropoietin receptor. Blood. 1993;81:2223-2236.

109. Cheetham JC SD, Aoki KH, Stevenson JL, Hoeffel TJ, Syed RS, Egrie J, Harvey TS. NMR structure of human erythropoietin and a comparison with its receptor bound conformation. Nat Struct Biol. 1998;5:861-866.
110. Livnah O SE, Middleton SA, Johnson DL, Jolliffe LK, Wilson IA. Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. Science. 1999;283:987-990.
111. Remy I WI, Michnick SW. Erythropoietin receptor activation by a ligand-induced conformation change. Science. 1999;283:990-993.
112. Kralovics R, Prchal JT. Genetic heterogeneity of primary familial and congenital polycythemia. Am J Hematol. 2001;68:115-121.
113. Erslev AJ CJ, Besarab A. Why the kidney? Nephron. 1985;41:213-216.
114. Semenza GL, GL W. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. Mol Cell Biol. 1992;12:5447-5454.
115. Semenza GL NM, Chi SM, Antonarakis SE. Hypoxia-inducible nuclear factors bind to an enhancer element located 3' to the human erythropoietin gene. Proc Natl Acad Sci U S A. 1991;88:5680-5684.
116. Salceda S CJ. Hypoxia-inducible factor 1alpha (HIF-1alpha) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes. J Biol Chem. 1997;272:22642-22647.
117. Huang LE GJ, Schau M, Bunn HF. Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway. Proc Natl Acad Sci U S A. 1998;95:7987-7992.
118. Jaakkola P, Mole DR, Tian YM, et al. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science. 2001;292:468-472.
119. Ivan M, Kondo K, Yang H, et al. HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. Science. 2001;292:464-468.
120. Epstein AC GJ, McNeill LA, Hewitson KS, O'Rourke J, Mole DR, Mukherji M, Metzen E, Wilson MI, Dhanda A, Tian YM, Masson N, Hamilton DL, Jaakkola P, Barstead R, Hodgkin J, Maxwell PH, Pugh CW, Schofield CJ, Ratcliffe PJ. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. Cell. 2001;107:43-54.
121. Bruick RK MS. A conserved family of prolyl-4-hydroxylases that modify HIF. Science. 2001;294:1337-1340.
122. Yu F WS, Zhao Q, Lee FS. HIF-1alpha binding to VHL is regulated by stimulus-sensitive proline hydroxylation. Proc Natl Acad Sci U S A. 2001;98:9630-9635.
123. Pugh CW ORJ, Nagao M, Gleadle JM, Ratcliffe PJ. Activation of hypoxia-inducible factor-1; definition of regulatory domains within the alpha subunit. J Biol Chem. 1997;272:11205-11214.
124. Maxwell PH WM, Chang GW, Clifford SC, Vaux EC, Cockman ME, Wykoff CC, Pugh CW, Maher ER, Ratcliffe PJ. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature. 1999;399:271-275.
125. Ohh M PC, Ivan M, Hoffman MA, Kim TY, Huang LE, Pavletich N, Chau V, Kaelin WG. Ubiquitination of hypoxia-inducible factor requires direct binding to the beta-domain of the von Hippel-Lindau protein. Nat Cell Biol. 2000;2:423-427.
126. Mahon PC HK, Semenza GL. FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity. Genes Dev. 2001;15:2675-2686.
127. McNeill LA HK, Claridge TD, Seibel JF, Horsfall LE, Schofield CJ. Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the beta-carbon of asparagine-803. Biochem J. 2002;367:571-575.
128. Lando D PD, Whelan DA, Gorman JJ, Whitelaw ML. Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. Science. 2002;295:858-861.
129. Tian H MS, Russell DW. Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. Genes Dev. 1997;11:72-82.
130. Ema M TS, Yokotani N, Sogawa K, Matsuda Y, Fujii-Kuriyama Y. A novel bHLH-PAS factor with close sequence similarity to hypoxia-inducible factor 1alpha regulates the VEGF expression and is potentially involved in lung and vascular development. Proc Natl Acad Sci U S A. 1997;94:4273-4278.

131. Flamme I FT, von Reutern M, Kappel A, Damert A, Risau W. HRF, a putative basic helix-loop-helix-PAS-domain transcription factor is closely related to hypoxia-inducible factor-1 alpha and developmentally expressed in blood vessels. Mech Dev. 1997;63:51-60.
132. Luo G GY, Jain S, Chan WK, Carr KM, Hogenesch JB, Bradfield CA. Molecular characterization of the murine Hif-1 alpha locus. Gene Expr. 1997;6:287-299.
133. Wiesener MS JJ, Rosenberger C, Scholze CK, Hörstrup JH, Warnecke C, Mandriota S, Bechmann I, Frei UA, Pugh CW, Ratcliffe PJ, Bachmann S, Maxwell PH, Eckardt KU. Widespread hypoxia-inducible expression of HIF-2alpha in distinct cell populations of different organs. FASEB J. 2003;17:271-273.
134. Rosenberger C HS, Rosen S, Shina A, Goldfarb M, Griethe W, Frei U, Reinke P, Bachmann S, Eckardt KU. Up-regulation of HIF in experimental acute renal failure: evidence for a protective transcriptional response to hypoxia. Kidney Int. 2005;67:531-542.
135. Warnecke C ZZ, Kurreck J, Erdmann VA, Frei U, Wiesener M, Eckardt KU. Differentiating the functional role of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha (EPAS-1) by the use of RNA interference: erythropoietin is a HIF-2alpha target gene in Hep3B and Kelly cells. FASEB J. 2004;18:1462-1464.
136. Hara S HJ, Kobayashi C, Kondo Y, Imura N. Expression and characterization of hypoxia-inducible factor (HIF)-3alpha in human kidney: suppression of HIF-mediated gene expression by HIF-3alpha. Biochem Biophys Res Commun. 2001;287:808-813.
137. Ang SO, Chen H, Hirota K, et al. Disruption of oxygen homeostasis underlies congenital Chuvash polycythemia. Nat Genet. 2002;32:614-621.
138. Percy MJ, Zhao Q, Flores A, et al. A family with erythrocytosis establishes a role for prolyl hydroxylase domain protein 2 in oxygen homeostasis. Proc Natl Acad Sci U S A. 2006;103:645-649.
139. Skoda R PJ. Lessons from familial myeloproliferative disorders. Semin Hematol. 2005;42:266-273.
140. Harrison CN, Gale RE, Wiestner AC, Skoda RC, Linch DC. The activating splice mutation in intron 3 of the thrombopoietin gene is not found in patients with non-familial essential thrombocythaemia. Br J Haematol. 1998;102:1341-1343.
141. Hess G, Rose P, Gamm H, Papadileris S, Huber C, Seliger B. Molecular analysis of the erythropoietin receptor system in patients with polycythaemia vera. British Journal of Haematology. 1994;88:794-802.
142. Le Couedic JP, Mitjavila MT, Villeval JL, et al. Missense mutation of the erythropoietin receptor is a rare event in human erythroid malignancies. Blood. 1996;87:1502-1511.
143. Kunishima S, Mizuno S, Naoe T, Saito H, Kamiya T. Genes for thrombopoietin and c-mpl are not responsible for familial thrombocythaemia: a case study. Br J Haematol. 1998;100:383-386.
144. Wiestner A, Padosch SA, Ghilardi N, et al. Hereditary thrombocythaemia is a genetically heterogeneous disorder: exclusion of TPO and MPL in two families with hereditary thrombocythaemia. Br J Haematol. 2000;110:104-109.
145. Kralovics R, Sokol L, Broxson EH, Jr., Prchal JT. The erythropoietin receptor gene is not linked with the polycythemia phenotype in a family with autosomal dominant primary polycythemia. Proceedings of the Association of American Physicians. 1997;109:580-585.
146. Collins A FJ, Teague J, Morton NE. A metric map of humans: 23,500 loci in 850 bands. Proc Natl Acad Sci U S A. 1996;93:14771-14775.
147. Broman KW MJ, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet. 1998;63:861-869.
148. Clayton DG WN, Smyth DJ, Pask R, Cooper JD, Maier LM, Smink LJ, Lam AC, Ovington NR, Stevens HE, Nutland S, Howson JM, Faham M, Moorhead M, Jones HB, Falkowski M, Hardenbol P, Willis TD, Todd JA. Population structure, differential bias and genomic control in a large-scale, case-control association study. Nat Genet. 2005;37:1243-1246.
149. Brodmann S, Passweg JR, Gratwohl A, Tichelli A, Skoda RC. Myeloproliferative disorders: complications, survival and causes of death. Ann Hematol. 2000;79:312-318.
150. Pearson TC. Evaluation of diagnostic criteria in polycythemia vera. Semin Hematol. 2001;38:21-24.

151. Murphy S. Diagnostic criteria and prognosis in polycythemia vera and essential thrombocythemia. Seminars in Hematology. 1999;36:9-13.
152. Jaffe ES, Harris NL, Stein H. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. Lyon: IARC Press; 2001.
153. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. Blood. 2002;100:2292-2302.
154. Thiele J, Kvasnicka HM, Orazi A. Bone marrow histopathology in myeloproliferative disorders--current diagnostic approach. Semin Hematol. 2005;42:184-195.
155. Kovacsovics TJ, Bachelot C, Toker A, et al. Phosphoinositide 3-kinase inhibition spares actin assembly in activating platelets but reverses platelet aggregation. J Biol Chem. 1995;270:11358-11366.
156. Bauermeister DE. Quantitation of bone marrow reticulin--a normal range. Am J Clin Pathol. 1971;56:24-31.
157. Rudzki Z, Kawa R, Okoń K, Szczygieł E, Stachura J. Objective, planimetry-based assessment of megakaryocyte histological pictures in Philadelphia chromosome-negative chronic myeloproliferative disorders: a perspective for a valuable adjunct diagnostic tool. Virchows Archiv. 2005;12:1-9.
158. Kralovics R, Buser AS, Teo SS, et al. Comparison of molecular markers in a cohort of patients with chronic myeloproliferative disorders. Blood. 2003;102:1869-1871.
159. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25:402-408.
160. Kralovics R, Teo SS, Buser AS, et al. Altered gene expression in myeloproliferative disorders correlates with activation of signaling by the V617F mutation of Jak2. Blood. 2005;106:3374-3376.
161. Weber JL, Wong C. Mutation of human short tandem repeats. Hum Mol Genet. 1993;2:1123-1128.
162. Li W. Molecular evolution. Sunderland, MA: Sinauer; 1997.
163. Thiele J, Kvasnicka HM, Fischer R. Histochemistry and morphometry on bone marrow biopsies in chronic myeloproliferative disorders - aids to diagnosis and classification. Ann Hematol. 1999;78:495-506.
164. Bartl R, Frisch B. Clinical significance of bone marrow biopsy and plasma cell morphology in MM and MGUS. Pathol Biol (Paris). 1999;47:158-168.
165. Li J, Xia Y, Kuter DJ. Interaction of thrombopoietin with the platelet c-mpl receptor in plasma: binding, internalization, stability and pharmacokinetics. Br J Haematol. 1999;106:345-356.
166. Eyster ME, Saletan SL, Rabellino EM, et al. Familial essential thrombocythemia. Am J Med. 1986;80:497-502.
167. Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, C. L. dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics. 2004;20:1233-1240.
168. Towbin H, Staehelin T, J G. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci U S A. 1979;76:4350-4354.
169. Witke W SA, Hartwig JH, Azuma T, Stossel TP, Kwiatkowski DJ. Hemostatic, inflammatory, and fibroblast responses are blunted in mice lacking gelsolin. Cell. 1995;81:41-51.
170. Maury CP BM. Isolation and characterization of cardiac amyloid in familial amyloid polyneuropathy type IV (Finnish): relation of the amyloid protein to variant gelsolin. biochem Biophys Acta. 1990;1096:84-86.
171. Ghiso J HM, Prelli F, Novello J, Frangione B. Gelsolin variant (Asn-187) in familial amyloidosis, Finnish type. Biochem J. 1990;272:827-830.
172. de la Chapelle A TR, Boysen G, Santavy J, Bleeker-Wagemakers L, Maury CP, Kere J. Gelsolin-derived familial amyloidosis caused by asparagine or tyrosine substitution for aspartic acid at residue 187. Nat Genet. 1992;2:157-160.
173. Geng YJ AT, Tang JX, Hartwig JH, Muszynski M, Wu Q, Libby P, Kwiatkowski DJ. Caspase-3-induced gelsolin fragmentation contributes to actin cytoskeletal collapse, nucleolysis,

and apoptosis of vascular smooth muscle cells exposed to proinflammatory cytokines. Eur J Cell Biol. 1998;77:294-302.
174. Sagawa N FH, Banno Y, Nozawa Y, Katoh H, Kuzumaki N. Gelsolin suppresses tumorigenicity through inhibiting PKC activation in a human lung cancer cell line, PC10. Br J Cancer. 2003;88:606-612.
175. De Corte V BE, Boucherie C, Mareel M, Vandekerckhove J, Gettemans J. Gelsolin-induced epithelial cell invasion is dependent on Ras-Rac signaling. Embo J. 2002;21:6781-6790.
176. Rojnuckarin P KK. Actin reorganization and proplatelet formation in murine megakaryocytes: the role of protein kinase calpha. Blood. 2001;97:154-161.
177. Prchal JT. Molecular biology of polycythemias
Pathogenetic mechanisms of polycythemia vera and congenital polycythemic disorders. Intern Med. 2001;40:681-687.
178. Ang SO, Chen H, Gordeuk VR, et al. Endemic polycythemia in Russia: mutation in the VHL gene. Blood Cells, Molecules & Diseases. 2002;28:57-62.
179. GL S. Hypoxia-inducible factor 1: control of oxygen homeostasis in health and disease. Pediatr Res. 2001.
180. Wang GL, Jiang BH, GL. S. Effect of protein kinase and phosphatase inhibitors on expression of hypoxia-inducible factor 1. Biochem Biophys Res Commun. 1995;216:669-675.
181. Sutter CH, Laughner E, GL. S. Hypoxia-inducible factor 1alpha protein expression is controlled by oxygen-regulated ubiquitination that is disrupted by deletions and missense mutations. Proc Natl Acad Sci U S A. 2000;97:4748-4753.
182. Livak KJ, TD S. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25:402-408.

CURRICULUM VITAE

Family Name: LIU
First Name: Kun
Date of Birth: May 3rd, 1976
Nationality: China
Marriage status: Single
Mailing Address: Ryffstrasse 58, Basel 4056, Switzerland
Telephone: 061 265 3282 (Office); 0764455592 (Mobile)
Email: kun.liu@unibas.ch

Education

PhD (Biomedical Research)	University of Basel, Switzerland	2007
Master of Science	National University of Singapore	2003
Bachelor of Medicine (M.B.B.S)	China Medical University	2001

Research Experience

10/2003 — present PhD student
Experimental Hematology laboratory, Department of Research, University Hospital Basel

Thesis: *Genetic studies of hereditary myeloproliferative diseases*

07/2001 — 10/2003 Research Scholar (Master sponsored by research scholarship)
Hepato-Oncogenetics Lab, Department of Physiology, Faculty of Medicine, National University of Singapore

Thesis: *Exploration of the functional significance of mig-2 in human cancer cell susceptibility to cytotoxic agents*

Professional Skills

General Research skills
- Manage and develop research projects independently during graduate studies
- Write manuscripts; Contribute to research grant proposal writing
- Give public presentations; Instruct new student in lab

Molecular & Cell Biology
- PCR; real-time PCR; Molecular Cloning; Western blotting
- Cell culture; retrovirus productionand infection of cell lines and primary cells; Cell proliferation assays (XTT/MTT)
- FACS analysis; Apoptosis assay (TUNEL)

Genetics
- Affymetrix SNP chip arrays; dCHIP analysis

- DNA sequencing; Mutation screening
- Microsatellite mapping; Fragment analysis

Animal experiments
- Holding animal working permit in Switzerland
- Experience in mouse blood sampling, organ removal/dissection, and bone marrow transplantation

Hematology
- Blood cell separation; EEC culture; transformation of B lymphocytes from patients with EBV

Other Experience

01/2000 — 07/2001 Medical Intern
The 2nd Affiliated Hospital of China Medical University, Shenyang, China
Tasks: Practice clinical medicine in Department of Internal Medicine, Surgical Department, Department of Gynaecology and Obstetrics, and Department of Paediatrics

09/1995 — 01/2000 Part-time Tutor
Tasks:Instructing high school students in English, Chinese and Mathematics

Extra-curricular activities

Chairman	Association of Chinese Students and Scholars in Basel	(2006-2007)
Vice Chairman	Association of Chinese Students and Scholars in Basel	(2004-2006)
Representative	Student Union of China Medical University	(1995-2001)

Scholarship & Awards

Research scholarship by National University of Singapore	(2001—2003)
Yearly Prize University Scholarship by China Medical University	(1995--- 2000)
Third Class Prize in China National College Students English Contest	(1999)
Excellent Student Leader by China Medical University	(1997)

Publications

• A *de novo* splice donor mutation in the thrombopoietin gene causes hereditary thrombocythemia in a Polish family (in revision)

• Mutation analysis in families with hereditary thrombocythemia and identification of a founder effect for a MPL mutation (in press)

• Genetic association analysis of a family with congenital secondary polycythemia (manuscript in preparation)