A new potential oncogenic mutation in the FERM domain of JAK2 in *BCR-ABL1* negative and V617F negative chronic myeloproliferative neoplasms (CMPNs) revealed by a comprehensive screening of 17 tyrosine kinase coding genes

Paula Aranaz[1], Cristina Ormazábal[1], Cristina Hurtado, Ignacio Erquiaga, Maria J. Calasanz, Marina García-Delgado, Francisco J. Novo, José L. Vizmanos[*]

Department of Genetics, School of Sciences, University of Navarra, Pamplona, Spain

*These authors contributed equally to this work

[*]Corresponding author: José Luis Vizmanos, PhD. Department of Genetics, School of Sciences. Research Building. University of Navarra. C/ Irunlarrea 1. E-31008 Pamplona, Spain. E-mail address: jlvizmanos@unav.es.

Abstract

BCR-ABL1 negative chronic myeloproliferative neoplasms (CMPNs) are a heterogeneous group of clonal haematological malignancies. Over the last years, some genetic events in tyrosine kinase (TK) genes have been described as causal events of these diseases. In order to find new genetic aberrations underlying these diseases, we have used dHPLC and FISH to analyse 17 genes from two families of receptor-TKs (families III and IV) and three families of cytoplasmatic-TKs (Syk, Abl and Jak) on samples from 44 *BCR-ABL1* negative and *JAK2*V617F negative CMPN patients with different clinical phenotypes. While screening by FISH did not reveal novel chromosomal aberrations, we detected several sequence changes. Although none of them were frequent events, we identified a new potential activating mutation in the FERM domain of JAK2 (R340Q). None of the germline SNPs detected showed a different distribution between patients and controls. In summary, data presented here show that these genes are not frequently mutated or rearranged in CMPNs, suggesting that molecular events causing these disorders must be located in other genes.

Introduction

BCR-ABL1 negative chronic myeloproliferative neoplasms (CMPNs) are a heterogeneous group of clonal haematological malignancies characterised by abnormal proliferation and survival of one or more myeloid lineage cells. In some cases, they can evolve to acute myeloid leukemia (AML). These diseases include both "classic" CMPNs (essential thrombocythemia –ET-, polycythaemia vera –PV- and idiopathic myelofibrosis –IMF-), and "atypical" or "non-classic" CMPNs (like chronic eosinophilic leukaemia, chronic neutrophilic leukemia, hypereosinophilic syndrome, mast cell disease and myeloid neoplasms with eosinophilia) [1]. Unlike chronic myeloid leukemia (CML), in which *BCR-ABL1* fusion seems to be the primary cause of the disease, the pathogenesis of these CMPNs has remained elusive until recently. Over the last years, some genetic alterations (like mutations in *JAK2*) have been described, most of them activating tyrosine kinase (TK) genes that play a role similar to *ABL1* in CML.

TKs have an important role in cell growth and oncogenesis regulating multiple cellular processes. Gain-of-function abnormalities of TK genes can result in constitutive activation of the signalling pathways in which they are involved, contributing to tumour development and progression, and causing the abnormal proliferation and survival that characterize these diseases [2].

The RTKs (receptor tyrosine kinase) class III genes include *PDGFRA* (4q12), *PDGFRB* (5q33), *CSF1R* (5q33), *KIT* (4q12) and *FLT3* (13q12). RTKs class IV or FGFRs genes include *FGFR1* (8p11), *FGFR2* (10q26), *FGFR3* (4p16) and *FGFR4* (5q35). All of them code for receptors characterised by several immunoglobulin (Ig)-like domains in the extracellular ligand-binding region, a transmembrane domain, a

juxtamembrane domain and a splitted cytoplasmatic TK domain. Ligand binding to RTKs promotes receptor dimerization and subsequent activation of downstream signalling pathways.

Whereas RTKs are able to initiate signal transduction pathways, CTKs are mediators in the same or in other pathways triggered by receptors with and without TK activity. Genes from Syk family, *SYK* (9p22) and *ZAP70* (2q12), are both expressed in different cells from haematological tissue (mainly B and T cells). Both share the same structure, with two N-terminal SH2 protein-binding domains and a C-terminal TK domain. Genes from the Abl family (*ABL1* located in 9q34 and *ABL2* in 1q25) share an N-terminal SH3 domain, followed by a protein-binding SH2 domain, a central TK domain and a C-terminal DNA-binding domain. The Jak family is composed by four genes (*JAK1* in 1p31, *JAK2* in 9p24, *JAK3* in 19p13 and *TYK2* in19p13). All of them are important players of JAK-STAT signalling pathway triggered by several different cytokines which has been found aberrantly activated in CMPNs [3] mostly as consequence of *JAK2* mutations. Four JAKs share structural properties, with an N-terminal receptor-binding FERM domain, followed for an internal protein-binding SH2 domain, a unique autoinhibitory pseudokinase domain and a C-terminal TK domain.

To date, some fusions different to *BCR-ABL1* but involving also *ABL1* (*ETV6-ABL1*) and other TKs (*SYK*, *JAK2*, *FGFR1*, *PDGFRA*, *PDGFRB* and *FLT3*) have been described in isolated CMPN cases. Some of these genes are rearranged with multiple partners (for a review see [4]) and others, like *KIT* and *FLT3*, have been found also mutated in several neoplasms. However, so far only the *JAK2* V617F point mutation has been found recurrently in classic CMPNs and in a small number of non-classic CMPNs [5], so there is still a significant proportion of patients in which the disease-

causing molecular event remains to be discovered. In addition, recent reports have addressed the role of germline variation in MPN pathogenesis, defining putative disease susceptibility *loci* [6-9] that would explain, at least in part, the familial clustering observed in these diseases [10,11].

The activating abnormalities described in neoplasms and the impact of the targeted therapy have pointed to TK coding genes as obvious candidates for systematic screenings in cancer. In fact, this is the focus of a comprehensive study worldwide using different approaches on different tumours [12-17].

In this study, we have analysed all genes from the families III and IV of RTKs and Syk, Abl and Jak families of CTKs in a candidate-gene approach under the assumption that one or some of them could be altered in *BCR-ABL1* negative and *JAK2*V617F negative CMPNs. All of these genes have a functional role in haematopoietic cells and at least one gene of each family has been found fused or mutated in a hematological neoplasm. Two approaches have been used for this screening: FISH (fluorescence *in situ* hybridization) to detect putative cryptic rearrangements, and a sensitive, high-throughput and low-cost screening method (dHPLC, denaturing HPLC) to detect sequence variants.

Materials and methods

Material from 44 different caucasian CMPN patients (28 males and 16 females) with normal karyotype from several hospitals from the north of Spain was obtained with informed consent from individual patients and the study was approved by the internal Ethics Commitee. Four patients were diagnosed of PV, 15 ET, four IMF and 21 atypical CMPNs (aCMPN). Previously to the inclusion in our series, the *BCR-ABL1*

fusion was ruled out by conventional RT-PCR and FISH, and *JAK2*V617F mutation was negative by ARMS-PCR [18].

Cell suspensions in sufficient quantity and quality for FISH analysis were only available for 27 of the 44 patients for all the tests performed.

In addition, the analysis by dHPLC included 20 normal (no disease) samples used as controls in order to check the frequency of sequence changes observed (putative polymorphisms) in our population.

FISH

FISH analysis was performed with commercial probes for *PDGFRB* and *CSF1R* (*PDGFRB 5q33 Break Probe* from Kreatech Biotechnology BV, Amsterdam, The Netherlands), *ABL1* (*Vysis® LSI® BCR/ABL ES Dual Color Translocation Probe*, Vysis, Abbot Molecular, IL) and *ABL2* (*Vysis® LSI® 1p36/LSI® 1q25 and LSI® 19q13/19p13 Dual Color Probe*, Vysis, Abbot Molecular, IL). For the other genes, we designed two-colour FISH analyses with flanking probes (assay A) and one-colour analyses with a gene-spanning probe (assay B) using BAC clones according to the position showed by *UCSC Genome Browser* (http://genome.ucsc.edu). These clones were obtained from *BACPAC Resources Center* at *Children´s Hospital Oakland Research Institute* (Oakland, CA) and *Clone Resources* at *The Welcome Trust Sanger Institute* (Hixton, UK) (additional file 1). Assay B was only performed in cases in which assay A showed an ambiguous pattern.

Some of the TK genes selected for the analyses have been found fused with *ETV6* (12p13), so we also analyzed this gene with *Vysis® LSI® ETV6 Dual Color Breakapart Probe* (Abbott Molecular, Abbott Park, IL) in order to check if abnormalities in *ETV6* could unravel new TK fusions.

BACs were extracted from liquid LB cultures with *PhasePrep™ BAC DNA kit* (Sigma-Aldrich, St Louis, MO) and labelled by nick translation using the *CGH Nick Translation kit* (Abbott Molecular, Abbott Park, IL). Coprecipitation, hybridization and washing were performed using standard methods. Results were analyzed with a fluorescence microscope (*Nikon Eclipse 80i*, Nikon Instruments Inc, Japan) and *Isis FISH Imaging System* (MetaSystems, Germany) software. Every probe was previously tested on normal metaphases and nuclei to establish a cut-off normal value for each assay. Two hundred interphase nuclei, and metaphases when available, were analyzed for each assay and patient.

dHPLC analysis

Genomic DNA was obtained from all the samples and whole-genome amplification (WGA) was performed (*illustra™ GenomiPhi™ V2*, GE Healthcare, Piscataway, NJ) in order to obtain enough material for the mutational analyses. Several tests using DNA from original and WGA samples were performed to check if both results matched, and no discrepancies were observed.

For every gene screened, we selected exons coding the main functional protein domains such as the transmembrane (regulator) domain and the intracellular kinase domain. Due to its importance, the 25 coding exons of *JAK2* were analysed in 23 fragments. Primers located in flanking introns were designed using *Primer3* [19] (for a list for primers, size of fragments and dHPLC conditions see additional files 2). For each fragment, a mutant control fragment was obtained from normal DNA changing a nucleotide in the forward or reverse primer depending on the melting profile. This "mutant" fragment allowed us to validate the dHPLC assay. Melting profiles of PCR fragments were calculated in each case by *Navigator™ Software* v1.6.2 (Transgenomic Ltd., Omaha, NE), which helps to establish theoretical solvent

gradient and temperature/s of the analysis. Some fragments, due to their melting profiles, required more than one condition (temperature) to improve mutation detection. Conditions were then validated experimentally with appropriate controls. All the analyses were performed on a WAVE® 4500HT System (Transgenomic Ltd, Omaha, NE) with a DNASep® HT cartridge.

All the PCR reactions were performed with *AmpliTaq*™ *Gold* (Applied Biosystems, Foster City, CA) under standard conditions. After PCR cycling, samples were subjected to several cycles of heating and cooling in order to create heteroduplex molecules which improve mutation detection by dHPLC. For each fragment, we sequenced two samples of each different elution profile and results were analysed with *Mutation Surveyor* v3.10 (SoftGenetics LLC, State College, PA). The genomic reference sequences were obtained from *EnsEMBL.*

Association analysis

Although our main goal was not to genotype SNPs, our mutational analysis revealed some germline SNPs in exons and flanking intronic regions. Since the genetic background has been proposed as an important predisposing factor for the development of these diseases [11], we carried out statistical analyses to determine putative associations between allele and genotype frequencies of the SNPs detected and the development of the disease. Although our patient and control series were small and a preliminary association must be verified in larger series, an unequal distribution of alleles and/or genotypes could unravel a putative linkage disequilibrium with real predisposing variants.

For each sequence variant detected, we compared allele frequencies between patient and control groups using a two-tailed Fisher´s exact test. We used as control group

data obtained from 20 individuals without disease. In addition, when available, we also used genotype data from european individuals deposited in *dbSNP* (http://www.ncbi.nlm.nih.gov/projects/SNP/, caucasian populations). In addition, we compared genotype frequencies using a chi-square 2 x 3 contingency test with two degrees of freedom with two rows, one for patients and one for controls, and three columns, for the three different genotypes. As above, we compared patients *vs* our controls and patients *vs* controls plus european data from *dbSNP* when available. Hardy-Weinberg equilibrium test to ensure a correct genotyping were performed using *SNPStats* (http://bioinfo.iconcologia.net/index.php?module=Snpstats, [20]).

mRNA splicing prediction analysis

To check putative influence of intronic changes (both SNPs and mutations) on mRNA splicing, we used an mRNA splicing prediction web tool, *Human Splicing Finder* (http://www.umd.be/SSF/). This tool predicts the influence of nucleotide changes on mRNA splicing due to the effect on potential splice sites, splicing silencers and enhancers and branch points.

Results

FISH analysis

For *PDGFRA*, we only observed an abnormal FISH result with assay A (gene-flanking two colour labelled probes) in one aCMPN patient. This patient lost one red signal in 85% (170/200) of the nuclei and 8/10 metaphases analyzed, suggesting a deletion of the upstream 5´ part of *PDGFRA*. Subsequent analysis with a commercial FISH probe (*Poseidon*™ *Repeat Free*™ FIP-Chic-PDGFRA 4q12 Deletion/Break Dual-Color from Kreatech Biotechnology BV, Amsterdam, The Netherlands),

showed a pattern compatible with a *FIP1L1-PDGFRA* rearrangement not detected previously by conventional RT-PCR.

All other samples showed normal hybridization patterns for all the TK genes included in the analyses. In addition, no abnormal pattern with commercial assay for *ETV6* was found in any of the samples.

Mutation and association analyses

For each fragment, two samples from each different dHPLC elution profile were sequenced. We could detect 93 changes (see additional file 3) already described as polymorphisms in *dbSNP*. The most polymorphic gene was *SYK* with 11 SNPs detected, whereas the less polymorphic was *ABL1* with only one SNP detected. Only seven of the 93 changes detected were non-synonymous (rs3822214 in *KIT* -M541L-, rs17881656 in *FGFR3* -F384L-, rs351855 in *FGFR4* -G388R-, rs2230723 in *JAK2* –L393V-, rs3213409 in *JAK3* –V722I-, and rs12720356 –I684S- and rs55882956 –R703W- both in *TYK2*). None of them showed different distribution either in patients *vs* controls or in patients *vs* controls plus *dbSNP* data ($p>0.01$). For rs55882956 in *TYK2* no frequency data was available in *dbSNP*.

Thirty-five of the 93 SNPs were detected in exons but without effect on translation. Again, none of them showed different distribution either in patients *vs* controls or in patients *vs* controls plus *dbSNP* data ($p>0.01$). Remarkably, two of them (rs246395 in *PDGFRB* and rs3733542 in *KIT)* showed differences ($p=0.025$ and $p=0.005$) in the allelic frequencies when *dbSNP* data were considered but no differences were observed considering only our control data, so maybe there are some differences between european populations for these SNPs.

The remaining 51 SNPs were detected in flanking intronic regions. Some of them could have a potential effect on splicing according to *Human Splicing Finder* prediction (additional file 3) but again none of them showed a different distribution in patients *vs* controls or in patients *vs* controls plus *dbSNP* data (p>0.01).

Finally, we also found sequence changes not described in *dbSNP* (Table 1), all in a few number of individuals. Some of them were found both in patients and controls or only in control samples without effects on translation, except for 16283C>A (R925S) which was detected in *JAK3* in 1/43 patients (2,3%) and in 1/20 controls (5%). Only 57982G>A in *CSF1R* and 72438A>G in *FLT3* were detected in a significant proportion of patients and controls: 2/44 (4,5%) and 3/44 (6,8%) respectively in patients and 2/20 (10%) and 3/20 (15%) respectively in controls. Interestingly, we found a non-synonymous sequence change in exon 8 of *JAK2* (70507 G>A, R340Q) in a patient with an atypical CMPN (Figure 1A). The sequence pattern was compatible with a somatic event, although unfortunately we could not analyze another tissue of this patient to test this possibility. This change affects the FERM domain (Prosite #PS50057, InterPro #IPR000299). FERM domains are involved in protein-protein interactions mainly in the localization of proteins to the plasma membrane and they are found in cytoskeletal-associated proteins.

Discussion

Over the last years, the knowledge about the pathogenesis of the *BCR-ABL1* negative CMPNs has grown significantly, mainly due to the description of the *JAK2*V617F mutation in an important proportion of classic CMPNs, but also thanks to the characterization of multiple gene fusions involving several TK genes in a few number of CMPN patients. Taken together, all of these genetic lesions lead to TK activation,

resulting in aberrant signalling. Nevertheless, there are still a significant proportion of CMPN patients in which the molecular aberration which drives the disease remains unclear. In addition, during the past few years, some studies have shown association of different polymorphisms with the development of the neoplastic process, some of them in myeloproliferative neoplasms [6,14,21-24].

We considered the hypothesis that some TK genes, from families III and IV of RTKs and Syk, Jak and Abl families of CTKs, could be altered by cryptic genomic aberrations or point mutations in an important proportion of CMPN patients, explaining, at least in part, the development of these diseases. Some genes from these families have already been found mutated or rearranged in these and other tumours.

Our FISH analyses did not show any new cryptic chromosomal rearrangement in the patients analyzed. This suggests that gene-fusions of these TK genes remain as very infrequent events in these disorders. A *FIP1L1-PDGFRA* fusion that had been missed by conventional RT-PCR was detected by FISH in one patient. Our results confirm that, as previously reported by other groups, FISH analysis is a good technique to detect the *FIP1L1-PDGFRA* gene fusion when conventional RT-PCR fails due to the variability of the breakpoints in *FIP1L1*. In fact, a multiplex PCR from genomic material has been recently reported to improve the detection of this rearrangement [25].

We chose denaturing-HPLC as a powerful method for somatic mutation discovery in neoplastic diseases. With serial dilutions of mutant controls, we have been able to detect a mutation present in 3-5% molecules of the sample (data not shown) whereas the sensitivity of direct sequencing is around 10-20%.

Applying a dHPLC comprehensive screening approach on candidate genes, we have detected, as expected, many sequence variants previously described as SNPs in *dbSNP*. Only seven of these changes were non-synonymous, 35 were synonymous and 51 were located in intronic regions close to exons. Some of them could have a potential effect on splicing according to *in silico* analyses using *Human Splicing Finder*. However, none of them showed differences in genotype and allelic frequencies between patients and controls, although two, rs246395 in *PDGFRB* and rs3733542 in *KIT*, showed differences between our population data and data deposited in *dbSNP* for european populations. This fact can be due to population substructure or differences in allelic frequencies between subpopulations.

Recently, it has been reported that somatic mutations in *JAK2* are preferentially found within a particular inherited haplotype that could explain almost half of the CMPN risk attributable to inherited factors. The risk haplotype is tagged by a SNP located close to V617F mutation [7-9]. Two hypotheses could explain this association [26]. First, (hypermutability hypothesis) the risk haplotype could tag a hypermutable *locus*. A second alternative explanation ("fertile ground" hypothesis) would be that although the somatic mutation rate is the same on different *JAK2* haplotypes, only one of them would have a stronger selective advantage. Although the experimental design allowed us to detect sequence variants on intronic regions flanking exons, we could not directly test these hypotheses because our focus was on detection of sequence (somatic) variants in exons, so that we did not genotype the SNPs that have been found associated with the V617F mutation. In addition, the risk factor haplotype has been found in *JAK2*V617F positive cases, whereas our analyses were focused on *JAK2*V617F negative cases. In this regard, there is only one report showing a very

low association between *JAK2* risk haplotype and the development of the disease in *JAK2*V617F negative cases, which was not confirmed in additional series [7].

We also found sequence changes not described previously as SNPs (table 1). Only one, 72438 A>G in *FLT3* intron 16 with no potential effect on splicing, was observed in more than two patients and controls. Some of them were observed in one or two patients but also in control samples. Most of these sequence variants showed dHPLC and sequencing patterns compatible with changes present in the germ line, suggesting that they could be *bona fide* rare SNPs not yet described. However, no sample from another tissue was available to confirm this. Only 72454delTG and 84872C>A in *FLT3* and 15525 T>C in *JAK3* could have a putative effect on splicing according to *in silico* analyses performed with *Human Splicing Finder.*

Finally, only two non-synonymous changes were observed, R340Q (70507 G>A) in *JAK2* exon 8 and R925S (16286 C>A) in *JAK3* exon 20. The latter was found in one out of 43 patients (a patient with ET) but also in an individual from the control group. This fact points to a rare SNP. However, this change affects the kinase domain and recently a non-synonymous SNP (rs3213409) in exon 15 (pseudokinase domain) of *JAK3* (V722I) has been demonstrated as an activating point mutation in a patient with acute megakaryoblastic leukemia [27], although other genetic or environmental factors could promote the oncogenic effect of this variant.

The R340Q sequence variant could be a driver mutation. In fact, sequence chromatogram shows a pattern compatible with a somatic mutation (wild type allele in a very low proportion and mutant allele in a proportion >50%, figure 1A) and this change affects to a JAK2 interspecies conserved aminoacid present also in JAK3 (figure 1B). However, definite demonstration of its oncogenic potential should be made with functional analyses. This change affects the FERM domain, which is

involved in JAK2-receptor binding with a putative role on the regulation of the catalytic activity of the kinase [28]. Several different activating mutations apart from V617F have been described in the SH2 domain, the pseudokinase and the kinase coding regions of *JAK2* (Figure 1A), but, to the best of our knowledge, mutations in the FERM domain of JAK2 have not been described to date. Recently, it has been reported that the function of V617F is dependent on the JAK2 FERM domain. This domain seems to be inhibitory in wild-type JAK2 to keep JAK2 in an inactive basal state [29]. In addition, some mutations with transforming activity have been described in the FERM domain in JAK1 and JAK3. For instance, *JAK3*P132T was described in an acute megakaryoblastic leukemia (AMKL) patient [27] and I87T was found in a patient with Down Syndrome and AMKL with a previous history of transient myeloproliferative disorder [30]. The *JAK1*K204M mutation has been recently described in an adult with a B acute lymphoblastic leukemia [28], also with transforming activity. Although *JAK*2R340Q could be a very rare mutation, it will be interesting to explore its recurrence. Unfortunately most studies on *JAK2* are focused only in pseudokinase and kinase coding regions, so other changes such as this could be missed.

In summary, both FISH and dHPLC mutation screening have shown that genes coding for RTKs class III and IV and the CTKs families Abl, Syk and Jak are not frequently mutated in *BCR-ABL1* negative and V617F*JAK2* negative CMPN patients. Moreover there are still some patients with PV in which the molecular event is not known. Our results also show that dHPLC is a reliable and efficient method for mutation detection, although most of the changes detected are SNPs or passenger variants that do not contribute to the neoplastic process [31]. Only the non-synonymous mutation 70507 G>A (R340G) found in *JAK2* exon 8, coding for FERM

domain, could be a novel oncogenic driver change, although this requires confirmation by functional analyses.

Conclusions

In conclusion, molecular events or cryptic rearrangements causing *BCR-ABL1* negative and V617F *JAK2* negative CMPNs, if they exist, must be located in other TK coding genes or in genes coding for other molecules participating in the same signal transduction pathways. Despite the number of genes screened only a putative driver mutation, R340Q in *JAK2*, has been identified in a patient with atypical CMPN. Our results agree with other TK screening analyses performed mainly by resequencing in several cancers [12-17], in which the frequency of mutations in TK genes was found to be lower than initially expected.

Acknowledgements

This work has been funded with the help of the Institute of Health Carlos III (FIS PI040037), Spanish Ministry of Science and Innovation (SAF 2007-62473), the PIUNA Program of the University of Navarra and the Caja Navarra Foundation through the Program "You choose, you decide" (Project 10.830). FJN is a recipient of a "Jerónimo de Ayanz" award from the Government of Navarra.

References

1. Tefferi A, Vardiman JW. Classification and diagnosis of myeloproliferative neoplasms: the 2008 World Health Organization criteria and point-of-care diagnostic algorithms. Leukemia 2008;22:14-22.
2. De Keersmaecker K, Cools J. Chronic myeloproliferative disorders: a tyrosine kinase tale. Leukemia 2006;20:200-5.

3. Levine RL, Gilliland DG. Myeloproliferative disorders. Blood 2008;112:2190-8.

4. Macdonald D, Cross NC. Chronic myeloproliferative disorders: the role of tyrosine kinases in pathogenesis, diagnosis and therapy. Pathobiology 2007;74:81-8.

5. Tefferi A. Classification, diagnosis and management of myeloproliferative disorders in the JAK2V617F era. Hematology Am Soc Hematol Educ Program 2006;240-5.

6. Pardanani A, Fridley BL, Lasho TL, Gilliland DG, Tefferi A. Host genetic variation contributes to phenotypic diversity in myeloproliferative disorders. Blood 2008;111:2785-9.

7. Jones AV, Chase A, Silver RT, Oscier D, Zoi K, Wang YL, Cario H, Pahl HL, Collins A, Reiter A, Grand F, Cross NC. JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. Nat Genet 2009;41:446-9.

8. Kilpivaara O, Mukherjee S, Schram AM, Wadleigh M, Mullally A, Ebert BL, Bass A, Marubayashi S, Heguy A, Garcia-Manero G, Kantarjian H, Offit K, Stone RM, Gilliland DG, Klein RJ, Levine RL. A germline JAK2 SNP is associated with predisposition to the development of JAK2(V617F)-positive myeloproliferative neoplasms. Nat Genet 2009;41:455-9.

9. Olcaydu D, Harutyunyan A, Jager R, Berg T, Gisslinger B, Pabinger I, Gisslinger H, Kralovics R. A common JAK2 haplotype confers susceptibility to myeloproliferative neoplasms. Nat Genet 2009;41:450-4.

10. Bellanne-Chantelot C, Chaumarel I, Labopin M, Bellanger F, Barbu V, De Toma C, Delhommeau F, Casadevall N, Vainchenker W, Thomas G, Najman A. Genetic and clinical implications of the Val617Phe JAK2 mutation in 72 families with myeloproliferative disorders. Blood 2006;108:346-52.

11. Landgren O, Goldin LR, Kristinsson SY, Helgadottir EA, Samuelsson J, Bjorkholm M: Increased risks of polycythemia vera, essential thrombocythemia, and myelofibrosis among 24,577 first-degree relatives of

11,039 patients with myeloproliferative neoplasms in Sweden. Blood 2008;112:2199-204.

12. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew YE, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan MH, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA, Stratton MR. Patterns of somatic mutation in human cancer genomes. Nature 2007;446:153-8.

13. Ruhe JE, Streit S, Hart S, Wong CH, Specht K, Knyazev P, Knyazeva T, Tay LS, Loo HL, Foo P, Wong W, Pok S, Lim SJ, Ong H, Luo M, Ho HK, Peng K, Lee TC, Bezler M, Mann C, Gaertner S, Hoefler H, Iacobelli S, Peter S, Tay A, Brenner S, Venkatesh B, Ullrich A. Genetic alterations in the tyrosine kinase transcriptome of human cancer cell lines. Cancer Res 2007;67:11368-76.

14. Tomasson MH, Xiang Z, Walgren R, Zhao Y, Kasai Y, Miner T, Ries RE, Lubman O, Fremont DH, McLellan MD, Payton JE, Westervelt P, Dipersio JF, Link DC, Walter MJ, Graubert TA, Watson M, Baty J, Heath S, Shannon WD, Nagarajan R, Bloomfield CD, Mardis ER, Wilson RK, Ley TJ. Somatic mutations and germline sequence variants in the expressed tyrosine kinase genes of patients with de novo acute myeloid leukemia. Blood 2008;111:4797-808.

15. Loriaux MM, Levine RL, Tyner JW, Frohling S, Scholl C, Stoffregen EP, Wernig G, Erickson H, Eide CA, Berger R, Bernard OA, Griffin JD, Stone RM, Lee B, Meyerson M, Heinrich MC, Deininger MW, Gilliland DG, Druker BJ. High-throughput sequence analysis of the tyrosine kinome in acute myeloid leukemia. Blood 2008;111:4788-96.

16. Brown JR, Levine RL, Thompson C, Basile G, Gilliland DG, Freedman AS. Systematic genomic screen for tyrosine kinase mutations in CLL. Leukemia 2008;22:1966-9.

17. Tyner JW, Loriaux MM, Erickson H, Eide CA, Deininger J, Macpartlin M, Willis SG, Lange T, Druker BJ, Kovacsovics T, Maziarz R, Gattermann N, Deininger MW. High-throughput mutational screen of the tyrosine kinome in chronic myelomonocytic leukemia. Leukemia 2008;23:406-9.

18. Jones AV, Kreil S, Zoi K, Waghorn K, Curtis C, Zhang L, Score J, Seear R, Chase AJ, Grand FH, White H, Zoi C, Loukopoulos D, Terpos E, Vervessou EC, Schultheis B, Emig M, Ernst T, Lengfelder E, Hehlmann R, Hochhaus A, Oscier D, Silver RT, Reiter A, Cross NC. Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. Blood 2005;106:2162-8.

19. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365-86.

20. Sole X, Guino E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. Bioinformatics 2006;22:1928-9.

21. Schnittger S, Kohl TM, Leopold N, Schoch C, Wichmann HE, Kern W, Lohse P, Hiddemann W, Haferlach T, Spiekermann K. D324N single-nucleotide polymorphism in the FLT3 gene is associated with higher risk of myeloid leukemias. Genes Chromosomes Cancer 2006;45:332-7.

22. Pastinen T, Ge B, Hudson TJ. Influence of human genome polymorphism on gene expression. Hum Mol Genet 2006;15 (Spec No 1):R9-16.

23. Gondek LP, Dunbar AJ, Szpurka H, McDevitt MA, Maciejewski JP. SNP Array Karyotyping Allows for the Detection of Uniparental Disomy and Cryptic Chromosomal Abnormalities in MDS/MPD-U and MPD. PLoS ONE 2007;2:e1225.

24. El-Omar EM, Ng MT, Hold GL. Polymorphisms in Toll-like receptor genes and risk of cancer. Oncogene 2008;27:244-52.

25. Score J, Walz C, Jovanovic JV, Jones AV, Waghorn K, Hidalgo-Curtis C, Lin F, Grimwade D, Grand F, Reiter A, Cross NC. Detection and molecular

monitoring of FIP1L1-PDGFRA-positive disease by analysis of patient-specific genomic DNA fusion junctions. Leukemia 2009;23:332-9.

26. Campbell PJ. Somatic and germline genetics at the JAK2 locus. Nat Genet 2009;41:385-6.

27. Walters DK, Mercher T, Gu TL, O'Hare T, Tyner JW, Loriaux M, Goss VL, Lee KA, Eide CA, Wong MJ, Stoffregen EP, McGreevey L, Nardone J, Moore SA, Crispino J, Boggon TJ, Heinrich MC, Deininger MW, Polakiewicz RD, Gilliland DG, Druker BJ. Activating alleles of JAK3 in acute megakaryoblastic leukemia. Cancer Cell 2006;10:65-75.

28. Flex E, Petrangeli V, Stella L, Chiaretti S, Hornakova T, Knoops L, Ariola C, Fodale V, Clappier E, Paoloni F, Martinelli S, Fragale A, Sanchez M, Tavolaro S, Messina M, Cazzaniga G, Camera A, Pizzolo G, Tornesello A, Vignetti M, Battistini A, Cave H, Gelb BD, Renauld JC, Biondi A, Constantinescu SN, Foa R, Tartaglia M. Somatically acquired JAK1 mutations in adult acute lymphoblastic leukemia. J Exp Med 2008;205:751-8.

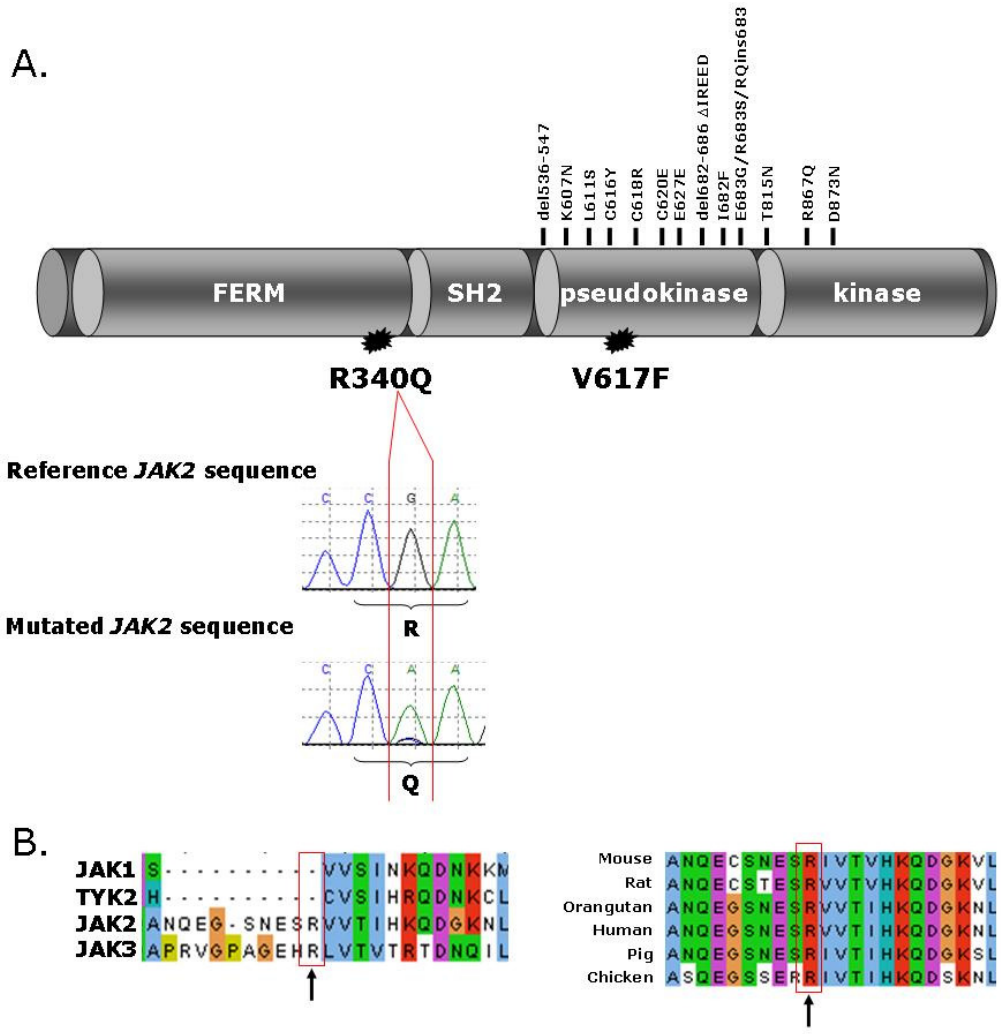
29. Zhao L, Ma Y, Seemann J, Huang LJ. A regulating role of JAK2 FERM domain in hyperactivation of JAK2(V617F). Biochem J 2009 Nov 24. [Epub ahead of print]

30. Kiyoi H, Yamaji S, Kojima S, Naoe T. JAK3 mutations occur in acute megakaryoblastic leukemia both in Down syndrome children and non-Down syndrome adults. Leukemia 2007;21:574-6.

31. Frohling S, Scholl C, Levine RL, Loriaux M, Boggon TJ, Bernard OA, Berger R, Dohner H, Dohner K, Ebert BL, Teckie S, Golub TR, Jiang J, Schittenhelm MM, Lee BH, Griffin JD, Stone RM, Heinrich MC, Deininger MW, Druker BJ, Gilliland DG. Identification of driver and passenger mutations of FLT3 by high-throughput DNA sequence analysis and functional assessment of candidate alleles. Cancer Cell 2007;12:501-13.

Figures

Figure 1. (A) Location of the R340Q mutation found in exon 8 of *JAK2* affecting FERM domain. The position of activating mutations found to date is also indicated. (B) Clustal-W alignment (http://www.ebi.ac.uk/Tools/clustalw2) of human Jak proteins (left) and JAK2 proteins in several species (right). Position 340 is conserved between human JAK2 and JAK3 and highly conserved in different JAK2 proteins.



Tables

Table 1 - Changes (not previously described as SNPs) observed in patients and controls. NA: not applicable.

Gene	Location	Base[1]	Patient	AA change	Splicing effect[2]	Frequency		Patients
						Patients	Controls	
PDGFRA	Intron 20	59678 G>A	14752, 18148	NA	No effect	2/44 (4,5%)	0/20 (0%)	2 aCMPN
CSF1R	Intron 19	57982 G>A	16737, 18708	NA	No effect	2/44 (4,5%)	2/20 (10%)	2 ET
KIT	Intron 17	78549 T>C		NA	No effect	0/44 (0%)	1/20 (5%)	
FLT3	Intron 16	72438 A>G	15703, 15081, 11615	NA	No effect	3/44 (6,8%)	3/20 (15%)	1 ET, 2 aCMPN
	Intron 16	72454delTG	9606	NA	Breaks potential splicing site	2/44 (4,5%)	0/20 (0%)	1 IMF
	Intron 20	84872 C>A	16659	NA	Breaks potential splicing site	1/44 (2,3%)	0/20 (0%)	1 ET
FGFR3	Intron 12	11635insGCC	16547	NA	No effect	1/44 (2,3%)	0/20 (0%)	1 ET
FGFR4	Exon 10	6546 C>T	18708	NO (G435G)	NA	1/44 (2,3%)	0/20 (0%)	1 ET
	Exon 11	6846 C>T	14886	NO (A501A)	NA	1/44 (2,3%)	0/20 (0%)	1 ET
ABL1	Intron 4	149290 C>T	17826		No effect	1/41 (2,5%)	0/20 (0%)	1 ET
	Intron 9	166395 C>T			No effect	0/44 (0%)	1/20 (5%)	
ZAP70	Intron 2	10431delC	16659		No effect	1/44 (2,3%)	0/20 (0%)	1 ET
JAK1	Intron 13	118975 C>T	15703		No effect	1/43 (2,3%)	0/20 (0%)	1 ET
	Exon 20	126901 C>T	16547	NO (D947D)	NA	1/43 (2,3%)	0/20 (0%)	1 ET
JAK2	Exon 8	70507G>A	16572	**R340Q**	NA	1/44 (2,3%)	0/20 (0%)	1 aCMPN
	Exon 19	96506 C>T	18148	NO (D820D)	NA	1/44 (2,3%)	0/20 (0%)	1 aCMPN
	Exon 25	141607 C>T	18630	Non coding region	NA	1/44 (2,3%)	0/20 (0%)	1 PV
JAK3	Exón 15	12791G>A	14837	NO (L643L)	NA	1/43 (2,3%)	0/20 (0%)	1 ET
	Intron 19	15525 T>C			Breaks potential branch site	0/43 (0%)	1/20 (5%)	
	Exon 20	16286 C>A	14837	R925S	NA	1/43 (2,3%)	1/20 (5%)	1 ET
	Intron 20	16389 C>T			No effect	0/43 (0%)	2/20 (10%)	
TYK2	Intron 12	16955 C>T	19537, 16547		No effect	2/43 (4,7%)	0/20 (0%)	1 PV, 1 ET
	Exón 18	24345 C>T		NO (R952R)	NA	0/43 (0%)	1/20 (5%)	
	Exón 23	27725 C>T	21762	Non coding region	NA	1/43 (2,3%)	0/20 (0%)	1 PV

[1] numbering according EnsEMBL (http://www.ensembl.org) data: *PDGFRA* (ENSG00000134853), *PDGFRB* (ENSG00000113721), *CSF1R* (ENSG00000182578), *KIT* (ENSG00000157404), *FLT3* (ENSG00000122025), *FGFR1* (ENSG00000077782), *FGFR2* (ENSG00000066468), *FGFR3* (ENSG00000068078), *FGFR4* (ENSG00000160867), *JAK1* (ENSG00000162434), *JAK2* (ENSG00000096968), *JAK3* (ENSG00000105639), *TYK2* (ENSG00000105397), *ABL1* (ENSG00000097007), *ABL2* (ENSG00000143322), *SYK* (ENSG00000165025), *ZAP70* (ENSG00000115085).

[2] data from *Human Splicing Finder* (http://www.umd.be/SSF/)

Additional files

Additional file 1 – BAC clones selected for FISH assays type A (two-colour FISH with flanking probes) and type B (one-colour FISH with a gene-spanning probe).

For the assay A, two clones on each side of the gene were selected and labelled in order to obtain a good and strong signal. For *FGFR3* only one clone on each side was sufficient. These clones were labelled with *Spectrum Orange dUTP** (SO, 5' clone) and *Spectrum green dUTP** (SG, 3' clone) both from Abbott Molecular (Abbott Park, IL). The normal hybridization pattern for assay A would be two green-red fusion signals. Clones selected for assay B were labelled with *Spectrum Orange dUTP** (SO). For this assay the normal hybridization pattern would be two separated red signals..

Additional file 2 – Primers and dHPLC conditions used for mutation scan analysis.

Three primers were used for each fragment. Patient samples were amplified with forward (Fw) and reverse (Rv) primer. A control DNA was amplified with a mutant primer (Fw + Rv mut or Fw mut + Rv) to generate mutant control DNA used to validate the assay. Changes in the sequence of the primers to generate mutant control fragment are showed underlined. In addition, the tables show the size of the PCR fragment analyzed and the temperature of the dHPLC assay. TM: transmembrane domain; TK: Tyrosine Kinase domain.

Additional file 3 – SNPs detected by mutation scanning of candidate genes.