

MYELOID NEOPLASIA

Germ line variants predispose to both *JAK2* V617F clonal hematopoiesis and myeloproliferative neoplasms

David A. Hinds,¹ Kimberly E. Barnholt,¹ Ruben A. Mesa,² Amy K. Kiefer,¹ Chuong B. Do,¹ Nicholas Eriksson,¹ Joanna L. Mountain,¹ Uta Francke,¹ Joyce Y. Tung,¹ Huong (Marie) Nguyen,³ Haiyu Zhang,⁴ Linda Gojenola,⁴ James L. Zehnder,^{4,5} and Jason Gotlib⁵

¹23andMe, Inc, Mountain View, CA; ²Division of Hematology & Medical Oncology, Mayo Clinic, Scottsdale, AZ; ³Division of Hematology/Oncology, University of Michigan Comprehensive Cancer Center, Ann Arbor, MI; ⁴Department of Pathology, Stanford University School of Medicine, Stanford, CA; and ⁵Division of Hematology, Department of Medicine, Stanford University School of Medicine/Stanford Cancer Institute, Stanford, CA

Key Points

- Germ line variants in *TERT*, *SH2B3*, *TET2*, *ATM*, *CHEK2*, *PINT*, and *GFI1B* are associated with *JAK2* V617F clonal hematopoiesis and MPNs.
- Age-related *JAK2* V617F clonal hematopoiesis is found in ~2 out of 1000 individuals in the general population.

We conducted a genome-wide association study (GWAS) to identify novel predisposition alleles associated with Philadelphia chromosome-negative myeloproliferative neoplasms (MPNs) and *JAK2* V617F clonal hematopoiesis in the general population. We recruited a web-based cohort of 726 individuals with polycythemia vera, essential thrombocythemia, and myelofibrosis and 252 637 population controls unselected for hematologic phenotypes. Using a single-nucleotide polymorphism (SNP) array platform with custom probes for the *JAK2* V617F mutation (V617F), we identified 497 individuals (0.2%) among the population controls who were V617F carriers. We performed a combined GWAS of the MPN cases plus V617F carriers in the control population (n = 1223) vs the remaining controls who were noncarriers for V617F (n = 252 140). For these MPN cases plus V617F carriers, we replicated the germ line *JAK2* 46/1 haplotype (rs59384377: odds ratio [OR] = 2.4, $P = 6.6 \times 10^{-89}$), previously associated with V617F-positive MPN. We also identified genome-wide significant associations in the *TERT* gene (rs7705526: OR = 1.8, $P = 1.1 \times 10^{-32}$), in *SH2B3* (rs7310615: OR = 1.4, $P = 3.1 \times 10^{-14}$),

and upstream of *TET2* (rs1548483: OR = 2.0, $P = 2.0 \times 10^{-9}$). These associations were confirmed in a separate replication cohort of 446 V617F carriers vs 169 021 noncarriers. In a joint analysis of the combined GWAS and replication results, we identified additional genome-wide significant predisposition alleles associated with *CHEK2*, *ATM*, *PINT*, and *GFI1B*. All SNP ORs were similar for MPN patients and controls who were V617F carriers. These data indicate that the same germ line variants endow individuals with a predisposition not only to MPN, but also to *JAK2* V617F clonal hematopoiesis, a more common phenomenon that may foreshadow the development of an overt neoplasm. (*Blood*. 2016;128(8):1121-1128)

Introduction

Although the *JAK2* V617F (V617F) mutation is identified in almost all patients with polycythemia vera (PV) and ~50% to 60% of patients with essential thrombocythemia (ET) and primary myelofibrosis (PMF),¹⁻⁴ these *BCR-ABL1*-negative myeloproliferative neoplasms (MPNs) exhibit a spectrum of phenotypes that cannot be explained by a single molecular abnormality. V617F gene dosage also contributes to phenotype; V617F allele burden tends to be higher in PV and PMF because of acquired uniparental disomy (UPD), whereas a lower allele burden is generally observed in ET patients.¹⁻⁵ Mutations in exon 12 of *JAK2* in PV,⁶ in the thrombopoietin (TPO) receptor *MPL* in ET and PMF,^{7,8} in *CBL*,⁹ or in *SH2B3* (*LNK*),¹⁰ also result in JAK-STAT pathway activation in MPNs. Somatic deletion and insertion mutations in the calreticulin (*CALR*) gene were identified in the majority of ET and PMF patients with nonmutated *JAK2* or *MPL*.^{11,12} MPN heterogeneity is also generated

by the acquisition of molecular abnormalities in addition to V617F (eg, *TET2*, *ASXL1*, *SRSF2*, *IDH 1/2*, *DNMT3A*, *EZH2*, and *TP53*),¹³ although no strict temporal order of these mutations defines the chronic or blast phase of MPNs.

Individual genetic background may influence the propensity for developing an MPN. A five- to sevenfold increased risk of MPN development was identified in first-degree relatives of patients with MPN.¹⁴ Three studies demonstrated that a germ line haplotype (GGCC, referred to as “46/1”) encompassing the 3’ region of the *JAK2* gene is associated with a three- to fourfold risk of developing a V617F-positive MPN.¹⁵⁻¹⁷ Patients who were heterozygous for this haplotype preferentially acquired the V617F mutation in *cis* with the predisposition allele (~75% of cases),^{15,16} suggesting that the haplotype may lead to hypermutability at the *JAK2* locus. However, the haplotype was also weakly associated with V617F-negative or *MPL*-mutant MPN,^{15,18-20}

Submitted June 20, 2015; accepted June 5, 2016. Prepublished online as *Blood* First Edition paper, June 30, 2016; DOI 10.1182/blood-2015-06-652941.

The online version of this article contains a data supplement.

There is an Inside *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

© 2016 by The American Society of Hematology

suggesting that it may also confer a more generalized propensity for MPN development independent of V617F. Although the 46/1 haplotype contributes to ~40% to 50% of the attributable risk for MPN development at a population level,¹⁶ additional inherited predisposition alleles likely contribute to this increased susceptibility.

In general population studies, the V617F mutation has also been identified in subjects without a diagnosed MPN or other overt hematologic disorder.²¹⁻²⁹ Although the clinical significance of V617F positivity (eg, *JAK2* V617F-mutant clonal hematopoiesis) in this context is unclear, some of these analyses have begun to characterize time-dependent rates of MPN development as well as the prevalence of non-MPN phenotypes in these individuals.

In this study, we performed genome-wide association analyses in order to identify additional germ line risk factors associated with MPNs, as well as *JAK2* V617F clonal hematopoiesis in the general population. This MPN research initiative employed a web-based strategy to recruit a cohort of MPN patients who submitted saliva samples for genotyping, as well as additional customers of the 23andMe Personal Genome Service (PGS) who served as population controls.

Methods

Subjects

Individuals with MPNs (including PV, ET, PMF, post-PV/ET myelofibrosis [MF], systemic mastocytosis, chronic myelogenous leukemia, chronic eosinophilic leukemia, and hypereosinophilic syndromes) were invited to participate in a free online MPN research initiative sponsored by 23andMe, Inc (Mountain View, CA) and were additionally provided with the 23andMe PGS at no cost. The 23andMe PGS is a direct-to-consumer genetic testing product of 23andMe, Inc (a more detailed description is provided in supplemental Appendix A, available on the *Blood* Web site). Customers submit a saliva DNA sample, are genotyped using a genome-wide single-nucleotide polymorphism (SNP) array, and receive a combination of ancestry analyses and reports of genetic contributions to health and other traits.

Participants provided informed consent and participated in the research online under a protocol approved by the external Association for the Accreditation of Human Research Protection Programs–accredited institutional review board, Ethical & Independent Review Services. Between August 2011 and December 2013, we enrolled a total of 1451 participants into this research initiative. We used data from additional customers of the 23andMe PGS who had agreed to participate in research between January 2011 and November 2013 as population controls for genome-wide association study (GWAS). For replication, we used customer data collected between January 2014 and February 2015.

Genotype data

DNA extraction and genotyping were performed on saliva samples by National Genetics Institute, a Clinical Laboratory Improvement Amendments–certified clinical laboratory and subsidiary of Laboratory Corporation of America. Individuals included in the GWAS were genotyped using a custom version of the Illumina HumanOmniExpress+ BeadChip. This platform has a base set of 730 000 SNPs, augmented with a custom set of ~250 000 SNPs selected by 23andMe. The custom content included 4 probes for the V617F somatic mutation. Samples that failed to reach a 98.5% call rate for SNPs on the standard portion of the array were reanalyzed. Individuals whose analyses failed repeatedly were recontacted by 23andMe customer service to provide replacement samples.

Participant genotype data were imputed against the August 2010 release of 1000 Genomes phase 1 reference haplotypes (refer to supplemental Appendix A for additional details regarding methods; see also <http://faculty.washington.edu/browning/beagle/beagle.html>, <http://genome.sph.umich.edu/wiki/Minimac>, and <http://www.1000genomes.org/>).³⁰⁻³² Replication samples were genotyped on a fully custom SNP array including 560 000 SNPs selected by 23andMe, with 8

probes for the V617F mutation. This data were imputed against the September 2013 release of 1000 Genomes phase 1 haplotypes, but methods and quality control were otherwise as described for the discovery GWAS.

Phenotype data

Phenotype data were collected by participant self-report using web-based surveys (supplemental Appendix A). Participants in the MPN research initiative filled out an online background survey that included questions about their specific diagnosis, treatment, and current disease status.

Somatic mutation analysis

To estimate V617F mutation burden from SNP array data, we computed a ratio of normalized intensity for the mutant allele to the sum of mutant and wild-type intensities, separately for each of the available probes. Individual observations with unusually low total intensity were removed, and then the median of the remaining observations was used as the mutation burden score.

We validated a subset of results using competitive allele-specific real-time polymerase chain reaction (RT-PCR) (CastPCR, assay *JAK2_12600_mu*; Life Technologies, Carlsbad, CA). Using serial dilutions of homozygous mutant DNA derived from the UKE-1 cell line (GM23245; Coriell Institute for Medical Research, Camden, NJ) into wild-type human DNA, we verified that the CastPCR assay was sensitive to the mutation at the 0.1% level in a wild-type background. Because of the potential variability of blood cells in saliva, which was used as the source of genotyping instead of peripheral blood or bone marrow, we evaluated the correlation between the allele fraction of *JAK2* V617F in saliva and blood in a cohort of MPN patients (supplemental Appendix B).

Statistical analysis

Individuals included in the genome-wide association analysis were selected for having >97% European ancestry, as determined through an analysis of local ancestry by comparison with the 3 HapMap 2 populations.³³ A maximal set of unrelated individuals was selected using a segmental identity-by-descent estimation algorithm.³⁴ Individuals were defined as close relatives if they shared >700 cM identity-by-descent, corresponding approximately to the minimal expected sharing between first cousins. Principal components analysis was used to model population substructure.

Genome-wide association tests were performed using imputed allele dosages. Each SNP was tested for association by logistic regression, assuming additive allelic effects. Age, sex, and 5 principal components³⁵ were included as covariates, and *P* values were computed using a likelihood ratio test. Results were adjusted for residual inflation using genomic control.³⁶

In regions with genome-wide significant associations, we performed stepwise conditional association tests to identify secondary signals. We added terms to the model for the most strongly associated SNP in a region and recomputed likelihood ratio test statistics for all other SNPs in an interval spanning 200 kb beyond any SNP in the region with $P < 1 \times 10^{-6}$. Results were genomic control–adjusted using the inflation factor from the original GWAS. We considered a secondary association to be significant if it had a Bonferroni-adjusted $P < .05$ accounting for the number of additional tests performed across the entire interval. Supplemental Appendix C lists the study populations that comprise the GWAS analyses included in this report, the genomic control inflation factor for each specified GWAS, and study design considerations. Supplemental Figure 1 displays the Q-Q plots for each GWAS.

Results

Detection of the somatic *JAK2* V617F mutation

We investigated whether we could detect the presence of V617F by comparing evidence for the mutation determined from our SNP arrays with self-reported V617F status. The array-based V617F mutation burden strongly correlated with self-reported mutation status (Figure 1A).

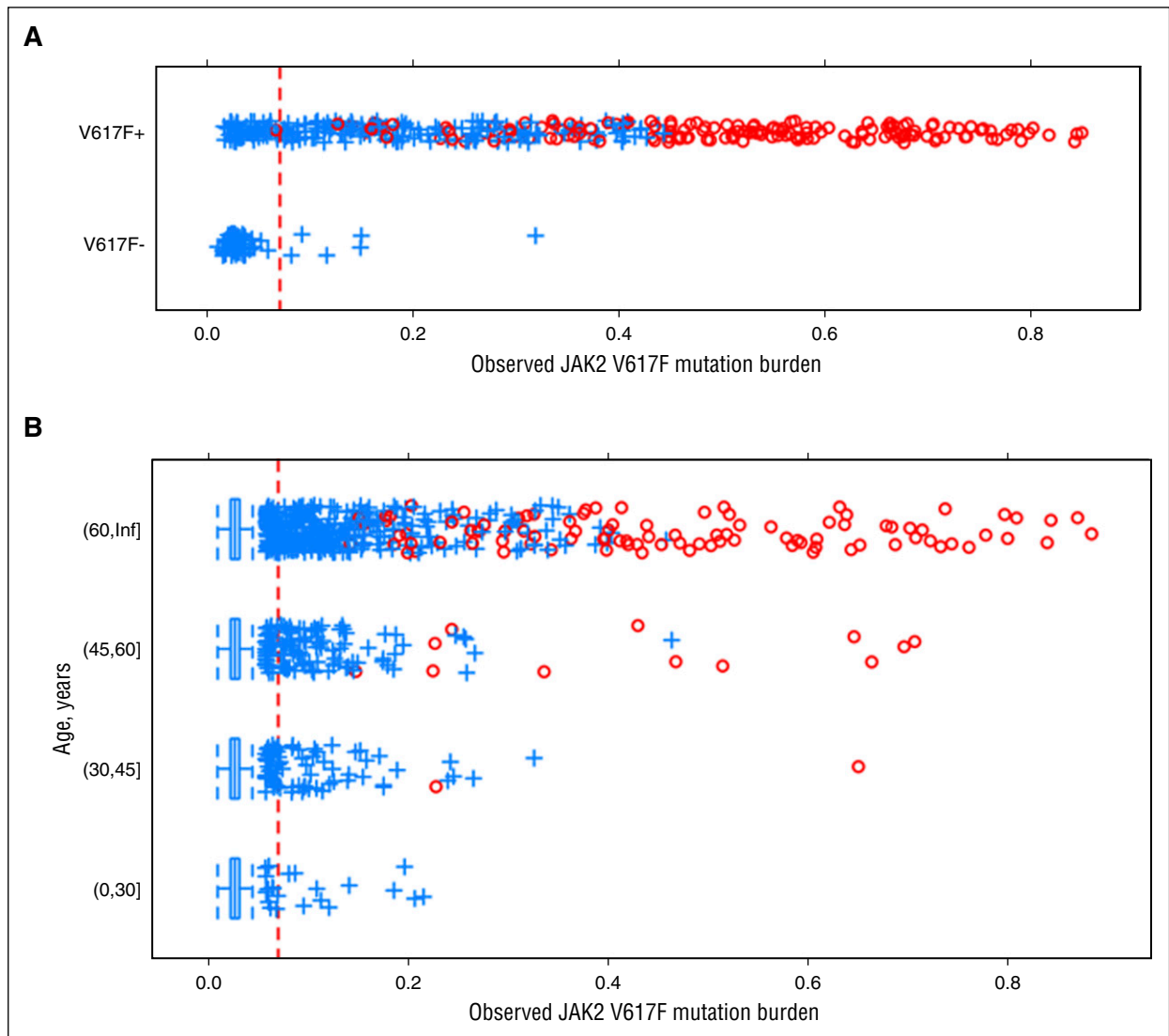


Figure 1. Distribution of JAK2 V617F array-based allele burden scores. Scores should be monotonically related to the proportion of the mutant allele in the DNA sample but were not calibrated. Individuals with evidence of UPD are shown as red circles. (A) Scores among MPN research initiative participants, stratified by self-reported mutation status (V617F-negative [-] or V617F-positive [+]). (B) Scores among population controls not enrolled in the MPN research initiative, stratified by age. The red dashed line represents the threshold used to classify V617F carriers in the analysis. Inf, infinity.

In population controls, we observed individuals who also had elevated V617F allele ratios, and both the prevalence and distribution of elevated allele ratios were strongly age dependent (Figure 1B). In both groups, higher allele burdens were accompanied by UPD (supplemental Appendix D).

We characterized classification performance of the array-based assay for V617F by using self-reported V617F-positive individuals to set a lower bound on the test sensitivity, assuming that any negative result in this group was a false negative. We used younger age population controls (age <40) to set a lower bound on the specificity of the test, on the assumption that any positive result for a younger age control was a false positive. We selected the detection threshold at which we expected a sensitivity of >85% and a specificity of >99.97%, giving an expected positive predictive value of >82% based on the observed distribution of positive results across all ages. Using this prespecified detection threshold, we selected 129 samples for validation using a sensitive RT-PCR assay. Of these, 126 samples tested positive for the V617F mutation, giving a specificity of 99.998% and positive predictive value of 97.7% (see supplemental Appendix E for additional

analysis of assay performance). In a separate cohort of 26 MPN patients, we found a strong positive correlation between the allele fraction of V617F in paired samples of saliva and blood (Pearson correlation coefficient, $R = 0.95$; supplemental Appendix B), indicating the feasibility of using saliva for genotyping of V617F status.

Cohort characteristics

Among the 1451 participants who enrolled in the MPN research initiative, we identified a cohort of 726 unrelated individuals (Table 1) with a self-reported diagnosis of ET (N = 246), PV (N = 258), PMF (N = 87), or related/overlapping MPN diagnoses (designated M-R) including ET/PV, PV + MF, or ET + MF (total N = 135) (supplemental Table 1). This M-R group primarily represented patients with dynamically evolving or overlapping diagnoses of ET and PV, and PV or ET that transformed to MF. Among the 252 637 unrelated population controls, we identified 497 individuals (0.2%) who had evidence for the V617F mutation (V617F carriers). Prevalence of V617F varied from 0.03% under the age of 30, to

Table 1. GWAS cohort demographics

	MPN cases	Population controls	V617F carriers
N	726 (100.0%)	252 637 (100.0%)	497 (100.0%)
Sex			
Male	223 (30.7%)	138 608 (54.9%)	295 (59.4%)
Female	503 (69.3%)	114 029 (45.1%)	202 (40.6%)
Age			
1-30	17 (2.3%)	35 149 (13.9%)	11 (2.2%)
31-45	120 (16.5%)	74 166 (29.4%)	45 (9.1%)
46-60	290 (39.9%)	66 390 (26.3%)	91 (18.3%)
61-112	299 (41.2%)	76 932 (30.5%)	350 (70.4%)
V617F self-reported status			
N/A	273 (37.6%)		
V617F-	92 (12.7%)		
V617F+	361 (49.7%)		

0.46% over the age of 60. This group tended to be older than participants in the MPN research initiative (70% over the age of 60 vs 31% of MPN participants). Although MPN initiative participants were predominantly female (69.3%), individuals in the population controls identified by their somatic mutation status were predominantly male (59.4%).

We tested 342 other self-reported phenotypes for association with V617F carrier status in the population controls, adjusting for age, sex, and genetic principal components (supplemental Tables 2 and 3). A few of these control individuals, although not enrolled in the MPN research initiative, independently reported having an MPN, and this was strongly associated with V617F carrier status (odds ratio [OR] = 238, $P = 5.6 \times 10^{-27}$). Nonspecific reports of blood disorders were also more common in V617F carriers (OR = 9.78, $P = 9.5 \times 10^{-21}$). V617F carriers were less likely to report having high cholesterol (OR = 0.53, $P = 1.4 \times 10^{-8}$) and had lower body mass index ($\beta = -1.25 \text{ kg/m}^2$, $P = 5.3 \times 10^{-6}$). Among other hematologic traits, V617F carriers reported more blood clots (OR = 1.8, $P = .014$), less anemia (OR = 0.45, $P = .019$), and higher prevalence of stroke (OR = 1.9, $P = .03$), but these results were not significant after adjusting for multiple tests. An analysis of reasons V617F+ controls gave for enrolling in 23andMe is provided in supplemental Appendix F.

Genome-wide association analysis

We performed separate genome wide association analyses comparing the 726 MPN cases with all 252 637 population controls (“MPN GWAS”; supplemental Appendix C) and compared the 497 V617F carriers from the control population with the remaining 252 140 V617F negative controls (“V617F status GWAS”; supplemental Appendix C). In both analyses, we observed strong associations with variants in the 9p24.1 region around *JAK2* and the 5p22 region in *TERT* (supplemental Figure 2; supplemental Table 4). An association in the *SH2B3* gene was suggestive ($P < 1 \times 10^{-6}$) in the MPN analysis and was genome-wide significant in the analysis of V617F carrier status among population controls.

Motivated by the consistency of results from these 2 analyses, we performed a “combined GWAS” analysis of the 1223 MPN cases plus V617F carriers vs 252 140 V617F noncarriers (Figure 2; supplemental Appendix C). In the combined analysis, 4 loci were genome-wide significant at the $P < 5 \times 10^{-8}$ level (Table 2; supplemental Figure 3). The strongest signal, in the 9p24.1 region, was in the *JAK2* gene (rs5938477: OR = 2.46, $P = 6.6 \times 10^{-89}$), consistent with prior reports of association at this locus. Strength of association of the *JAK2* variant rs59384377 varied with MPN diagnoses and generally followed the pattern of V617F mutation prevalence across diagnoses, with the strongest effect and highest prevalence in PV and the smallest effect and lowest prevalence in ET (Figure 3; supplemental Figure 4). We observed strong evidence for association extending beyond the end of the *JAK2* gene; however, the SNPs with the strongest effects were near the 5' end of the gene, in the second intron. After adjusting for the effect of rs59384377 (supplemental Figure 5; supplemental Table 5), we observed a secondary association in the same region (rs10974900, $P = 6.2 \times 10^{-6}$), 17 kb from rs59384377.

The second significant locus was in *TERT*, the gene encoding the protein component of telomerase reverse transcriptase. *TERT* rs7705526 was associated with an OR of 1.8 per G allele, $P = 2.9 \times 10^{-32}$. *TERT* variant rs7705526 was associated with ET, PV, and MF with similar effect sizes. We identified a secondary association at *TERT* rs2853677 (supplemental Figure 6; supplemental Table 5) that remained genome-wide significant after conditioning on rs7705526 ($P = 1.1 \times 10^{-12}$). We identified a third association with lead SNP rs7310615 in the *SH2B3* gene (OR = 1.5, $P = 3.1 \times 10^{-14}$) in strong linkage disequilibrium ($r^2 = 0.94$) with nonsynonymous variant

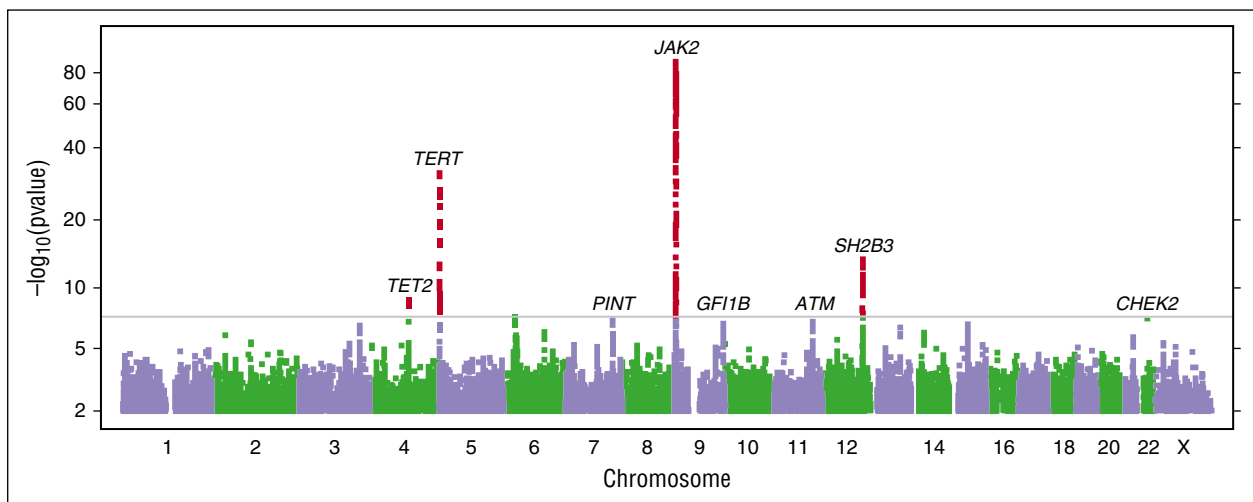


Figure 2. Manhattan plot of the combined GWAS of MPN cases and V617F carriers. Results with $P < 5 \times 10^{-8}$ (conventional threshold for genome-wide significance) are shown in red. We have also labeled suggestive results ($P < 1 \times 10^{-6}$) discussed in the text. Gene labels are provided for cross referencing with other results and are not intended to suggest a causal basis for the observed associations.

Table 2. Association statistics for lead SNPs

Gene context	SNP	Band	Position	Alleles	Dose.b	MPN		V617F+		Combined	
						P	OR	P	OR	P	OR
[<i>JAK2</i>]	rs59384377	9p24.1	5005034	A/T	0.265	4.3e-57	2.46	3.8e-35	2.33	6.6e-89	2.41
[<i>TERT</i>]	rs7705526	5p15.33	1285974	C/A	0.436	2.5e-16	1.69	1.3e-18	1.96	2.9e-32	1.80
[<i>TERT</i>]	rs2853677	5p15.33	1287194	A/G	0.355	5.2e-16	1.56	9.3e-19	1.80	5.2e-32	1.65
[<i>SH2B3</i>]	rs7310615	12q24.12	111865049	G/C	0.509	5.3e-07	1.32	2.5e-09	1.50	3.1e-14	1.39
[<i>CXXC4</i> —] <i>TET2</i>	rs1548483	4q24	105749895	C/T	0.037	1.6e-07	2.12	.0019	1.75	2.0e-09	1.97
[<i>CHEK2</i>]	rs555607708	22q12.1	29091857	I/D	0.002	1.1e-05	4.49	.0015	4.07	7.5e-08	4.35
[<i>ATM</i>]	rs1800056	11q22.3	108138003	T/C	0.014	1.2e-05	2.18	.011	1.78	6.5e-07	2.02
[<i>PINT</i>]	rs58270997	7q32.3	130729394	C/T	0.267	1.8e-06	1.43	.0079	1.27	1.1e-07	1.36
[<i>GFI1B</i> —] <i>GTF3C5</i>	rs621940	9q34.13	135870130	C/G	0.160	1.2e-05	1.35	.0031	1.28	1.9e-07	1.32

Band, cytogenetic band; position, chromosomal position in National Center for Biotechnology Information Build 37 coordinates; dose.b, average imputed dosage for the second listed allele; gene context, a representation of the gene(s) flanking the SNP position; I/D, insertion/deletion; P, the SNP association test P value.

rs3184504, W262R. The fourth identified association was with the lead SNP rs1548483 (OR = 2.0, $P = 2.0 \times 10^{-9}$), which falls 300 kb upstream of the *TET2* gene. In a secondary analysis of other MPN diagnoses not included in the GWAS (supplemental Table 6), rs1548483 was nominally associated with chronic myelogenous leukemia ($N = 130$, $P = .034$, OR = 2.1) and systemic mastocytosis ($N = 210$, $P = .013$, OR = 2.0), other diagnoses enlisted within the MPN research initiative.

We observed suggestive association signals with relatively large effect sizes at low-frequency nonsynonymous variants in *CHEK2* (1100delC, rs555607708: OR = 4.4, $P = 7.5 \times 10^{-8}$) and *ATM* (F858L, rs1800056: OR = 2.2, $P = 6.5 \times 10^{-7}$), with consistent effects in both MPN cases and V617F carriers. Two additional suggestive associations with smaller effect sizes from this combined GWAS included rs58270997 in the long intergenic noncoding RNA *PINT* ($P = 1.1 \times 10^{-7}$), and rs621940 downstream of the gene for growth factor independent 1B (*GFI1B*) ($P = 1.9 \times 10^{-7}$) (Table 2). We tested all pairs of index SNPs for interactions, but none were significant after adjusting for multiple testing (supplemental Appendix G).

We tested whether each index SNP could discriminate between V617F-positive or V617F-negative MPNs, using a case-only approach. Although *JAK2* rs1327494 was more strongly associated with V617F-positive disease (OR = 2.7, $P = 4.5 \times 10^{-18}$), *TERT* rs7705526 was only nominally associated with V617F-positive disease (OR = 1.3, $P = .04$). *TERT* rs2853677, *SH2B3* rs7310615, and *TET2* rs1548483 were not significantly associated with V617F status in MPN cases (Table 3; all $P > .05$). However, we observed a trend toward larger effect sizes for V617F-positive cases than for V617F-negative cases vs controls (supplemental Figure 7).

Replication analysis

Using the same array-based detection method, we identified 446 *JAK2* V617F carriers in an additional set of 169 467 unrelated individuals (supplemental Table 7; supplemental Appendix C). In this set, we confirmed associations between V617F status and lower risk of high cholesterol (OR = 0.59, $P = 3.9 \times 10^{-5}$), lower body mass index ($\beta = -1.4 \text{ kg/m}^2$, $P = 4.5 \times 10^{-6}$), and higher risk of blood clots (OR = 2.3, $P = 9.3 \times 10^{-6}$). This replication analysis of 446 *JAK2* V617F carriers vs 169 021 noncarriers confirmed associations with similar effect sizes that were observed from the combined GWAS analysis (Table 4).

We combined results from the combined GWAS and replication cohorts, using Fisher's method to compute joint P values. The 4 suggestive associations described for the combined GWAS analysis (*ATM*, *CHEK2*, *PINT*, and downstream of *GFI1B*) reached genome-wide significance in this joint analysis (Table 4, Joint P column). In the combined cohorts, we tested for associations between clinical

phenotypes and either V617F allele burden or UPD status, and found that the associations seen for V617F presence tended to be stronger at higher allele burdens or with UPD. Among our GWAS loci, the risk allele for *TERT* rs59384377 was associated with lower allele burden and risk of UPD, and the risk allele for *GFI1B* rs621940 was associated with higher burdens and risk of UPD (supplemental Appendix H).

Discussion

Our study extends the understanding of the genetic predispositions underlying *BCR-ABL1*-negative MPNs and *JAK2* V617F clonal hematopoiesis in the general population. In addition to the 46/1 *JAK2* haplotype, we identify predisposition alleles associated with *TERT*, *SH2B3*, *TET2*, *ATM*, *CHEK2*, *PINT*, and *GFI1B* in both MPN patients and in population controls who are V617F carriers. These genes impact diverse biologic pathways such as cellular aging (*TERT*), JAK-STAT signaling (*JAK2*, *SH2B3*), epigenetic regulation (*TET2*), DNA damage repair and/or tumor suppressor function (*ATM*, *CHEK2*, *PINT*), and erythroid/ megakaryocyte development (*GFI1B*).

Age-related clonal hematopoiesis in the general population was recently found to be associated with adverse outcomes, including an increased risk of hematologic cancer and all-cause mortality.^{27-29,37} *JAK2* was among the most commonly mutated genes in these studies, which also included *DNMT3A*, *TET2*, *ASXL1*, and *TP53*. The predisposition alleles we have identified in our study have similar ORs for V617F-positive and V617F-negative MPNs, as well as *JAK2* V617F clonal hematopoiesis in unselected population controls.

Besides the replicated *JAK2* 46/1 haplotype, the second most statistically significant SNPs in our analysis were 2 *TERT* loci, rs7705526 and rs2853677, previously identified in several cancer GWAS studies.³⁸⁻⁴¹ Since our initial identification of *TERT* (eg, SNP rs2853677) as a predisposition allele for MPNs,⁴² another germ line variant in *TERT*, rs2736100, was identified in Icelanders diagnosed with MPN.⁴³ Our lead SNP, rs2853677, is in moderate linkage disequilibrium with rs2736100 ($r^2 = 0.54$), and we see similar evidence for association at rs2736100 (OR = 1.6, $P = 7.5 \times 10^{-13}$). SNP rs2736100 has previously been associated with increased red blood cell and platelet counts and white blood cells of the myeloid lineage.⁴³ Recently, a third study corroborated *TERT* as a predisposition gene for MPN.⁴⁴

Our association with rs7310615 in the *SH2B3* gene, in strong linkage disequilibrium with nonsynonymous variant rs3184504, or *SH2B3* W262R, corroborates prior data showing an association between W262R and MPN patients.⁴⁵ Relevant to MPNs, SH2B3 binds to the

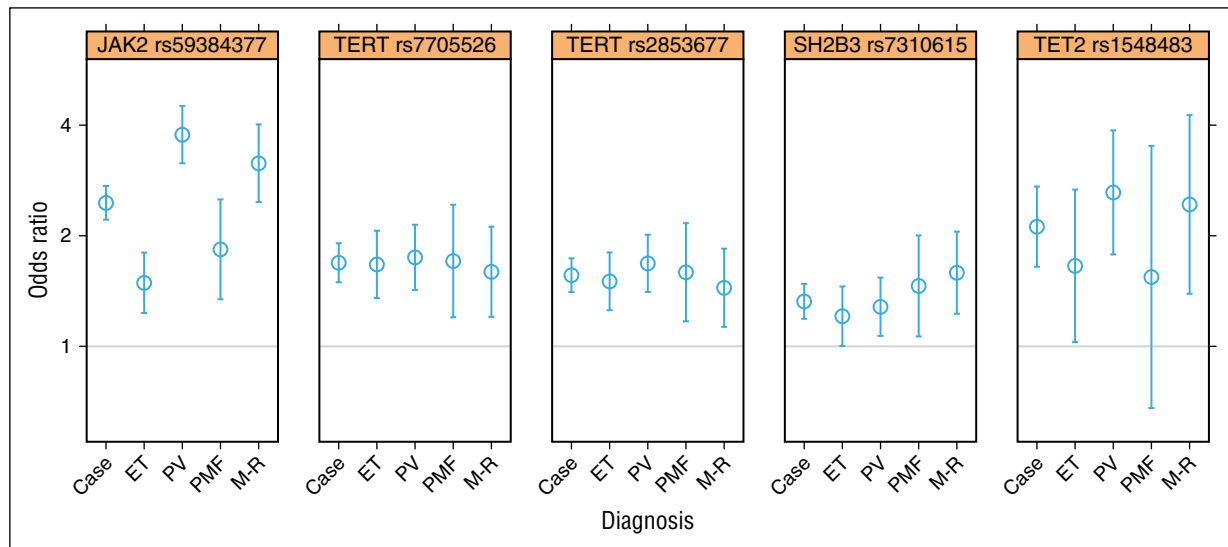


Figure 3. SNP associations with specific MPN diagnoses vs population controls. Here, the "case" result represents all cases used in the GWAS.

TPO receptor MPL; upon cytokine stimulation with TPO, SH2B3 also binds strongly to JAK2 and inhibits downstream STAT activation, resulting in negative feedback.^{46,47} *SH2B3*^{-/-} mice exhibit an MPN phenotype, including an expanded hematopoietic stem cell compartment with increased self-renewal, megakaryocytic hyperplasia, splenomegaly, leukocytosis, and thrombocytosis.⁴⁸ Somatic mutations in *SH2B3* have been identified at low frequency in MPN^{10,12} and result in aberrant JAK-STAT activation.¹⁰

This study's identification of a predisposition allele associated with *TET2* is intriguing given the common finding of somatic *TET2* mutations in a spectrum of myeloid malignancies, including MPN.^{49,50} In MPN, *TET2* mutations can be identified in hematopoietic stem cells and can either precede or follow the acquisition of V617F.^{13,49} The mutational order of these 2 genes can influence the clinical and biologic behavior of these neoplasms.⁵¹ Mutations in *TET2* constitute one of the early molecular lesions in myeloid neoplasms, including MPN; similarly, it is one of the more commonly mutated genes in individuals with clonal hematopoiesis.^{27-29,52}

In the joint analysis results of MPN cases plus V617F carriers vs noncarriers of V617F from the 2 unrelated control populations, genome-wide significance was observed for the variants *CHEK2* 1100delC and *ATM* F858L, as well as *PINT* and *GFI1B*. *ATM* and the G2 checkpoint kinase *CHEK2* are closely related components of the DNA-damage response pathway, which plays a critical role in maintaining genomic integrity. Mutations in genes of this pathway compromise DNA

repair and are linked to several types of malignancies and heritable cancer syndromes.⁵³⁻⁵⁶ Although the germ line variants *CHEK2* 1100delC and *ATM* F858L have been associated with chronic lymphocytic leukemia,⁵⁷ sparse data exist for their relationship to myeloid malignancies, including MPN.^{58,59} Somatic mutations in *CHEK2* were recently identified in 3 of 151 MPN patients,¹² highlighting a potential role of this gene in MPN pathogenesis. The long intergenic noncoding RNA *PINT* is regulated by p53 and interacts with polycomb repressive complex 2, a key regulator of hematopoietic stem cell differentiation and maintenance frequently inactivated in myeloid malignancies.⁶⁰ *GFI1B* is a zinc-finger transcriptional repressor that is essential for erythropoiesis and megakaryopoiesis.⁶¹ A nonsense mutation in *GFI1B* was identified in a family with autosomal dominant gray platelet syndrome.⁶² The truncated *GFI1B* mutant protein acts in a dominant-negative manner to block the normal development of megakaryocytes and platelets.

A recent study has reported additional associations with MPN at 3q26.2 downstream of *MECOM*, and at 6q23.3 between *HBSIL* and *MYB*.⁴⁴ At 3q26.2, we find weak support for the reported association at rs2201862 (GWAS $P = .02$, OR = 0.91 for T allele) but more evidence at an intronic site in *MECOM* (rs3851379: GWAS $P = 2.7 \times 10^{-7}$, OR = 1.24 for G allele; replication $P = .044$, OR = 1.15; joint $P = 2.3 \times 10^{-7}$). At the 6q23.3 locus, we replicate the reported association (GWAS $P = .0040$, OR = 1.14 for A allele; replication $P = .27$, OR = 1.09; joint $P = .0085$).

Studies have reported the detection of chromosomal abnormalities in blood, saliva, or buccal samples from SNP array data.^{63,64} However, these investigations did not report detection of somatic point mutations, largely because these variants are not present on standard genome-wide genotyping arrays designed to assay germ line variation. We show that it is feasible to detect the V617F mutation from SNP array data with high sensitivity and specificity. Although we do not expect this test to match the sensitivity of targeted blood-based assays, it may remain attractive as a screen followed by secondary testing. Although the variable composition of DNA in saliva from blood may affect the sensitivity of this assay, we demonstrated a very strong correlation between the V617F allele fraction in the peripheral blood and saliva of individual MPN patients. In addition to its application to blood and bone marrow, RT-PCR quantification of V617F mutant allele burden in saliva merits further exploration.

Table 3. Association of SNPs with *JAK2* V617F positivity in MPN cases

SNP	OR	95% CI	P
<i>JAK2</i> rs59384377	2.71	[2.13,3.44]	4.5e-18
<i>TERT</i> rs7705526	1.33	[1.02,1.73]	.035
<i>TERT</i> rs2853677	1.06	[0.84,1.32]	.64
<i>SH2B3</i> rs7310615	1.10	[0.87,1.39]	.41
<i>TET2</i> rs1548483	1.28	[0.74,2.20]	.38
<i>CHEK2</i> rs555607708	3.11	[0.83,11.65]	.068
<i>ATM</i> rs1800056	1.24	[0.64,2.43]	.52
<i>PINT</i> rs58270997	1.11	[0.83,1.48]	.50
<i>GFI1B</i> rs621940	0.86	[0.66,1.13]	.30

95% CI, confidence interval for OR; OR, odds ratio for *JAK2*⁺ vs *JAK2*⁻ status in MPN cases; P, association test P value.

Table 4. Replication cohort results for lead SNPs

SNP	OR	95% CI	P	Joint P*
JAK2 rs59384377	2.04	[1.78,2.32]	2.3e-24	3.9e-110
TERT rs7705526	1.70	[1.47,1.97]	2.1e-12	6.2e-42
TERT rs2853677	1.70	[1.49,1.95]	6.2e-15	3.4e-44
SH2B3 rs7310615	1.30	[1.14,1.49]	.00011	1.4e-16
TET2 rs1548483	1.89	[1.41,2.54]	8.7e-05	5.3e-12
CHEK2 rs555607708	4.89	[2.41,9.91]	.00042	7.8e-10
ATM rs1800056	1.96	[1.30,2.94]	.0035	4.7e-8
PINT rs58270997	1.40	[1.20,1.65]	4.2e-5	1.2e-10
GFI1B rs621940	1.34	[1.14,1.59]	.00070	3.2e-9

OR, odds ratio for the GWAS risk allele in the replication cohort; P, association test P value in the replication cohort.

*Joint P: overall test of combined GWAS results (Table 2) and replication cohort results using Fisher's method.

The 0.2% rate of V617F positivity in our control population mirrors the rates of 0.14% and 0.17% found in 2 Copenhagen general population studies,^{22,26} and the rates of 0.19%²⁷ and 0.18%²⁸ in 2 studies of age-related clonal hematopoiesis. In our study, the median age of V617F carriers was 68 years, with the incidence rising from 0.03% for ages 21 to 30 to 1.1% for ages 81 to 90. In our analysis and the Copenhagen General Population Study,²² V617F carrier status was more frequently associated with male sex, increasing age, and a strikingly similar multifactorial-adjusted OR for MPN (OR = 221 for the Copenhagen study; OR = 238 for the current GWAS). In addition, ORs for venous thromboembolism and deep venous thrombosis were 3.1 and 4.6 in the Copenhagen study; in our analysis, V617F carriers were similarly more likely to report a history of blood clots (OR = 1.8, P = .014). The 0.2% prevalence rate (age-adjusted: 0.16%) of JAK2 V617F in the control cohort is roughly double the estimated age-adjusted prevalence rate of combined JAK2 V617F-positive MPN (PV + ET + PMF + post-PV MF + post-ET MF) in the United States based on data from 2 large health care plans.⁶⁵ Therefore, the excess prevalence of JAK2 V617F clonal hematopoiesis among unselected controls likely represents a combination of (1) persons with an undiagnosed MPN, (2) individuals who may develop a future MPN, and (3) and those who will never develop a hematologic disorder.

In summary, we have identified inherited loci in or near *TERT*, *SH2B3*, *TET2*, *ATM*, *CHEK2*, *PINT*, and *GFI1B* that predispose to both age-related JAK2 V617F clonal hematopoiesis in the general population, as well as MPN independent of V617F status. In particular,

JAK2, *SH2B3*, *TET2*, and *CHEK2* represent 4 genes wherein both inherited variation and somatic mutation have been found to contribute to V617F clonal hematopoiesis and/or MPN development. The use of SNP arrays with custom probes for specific somatic mutations can be applied to analyses of other genes implicated in age-related clonal hematopoiesis and hematologic cancers.

Acknowledgments

The authors thank the participants in the 23andMe MPN Research Initiative and customers of 23andMe who participate in research, as well as the employees of 23andMe, who together made this research possible.

J.G. is supported by the Charles and Ann Johnson Foundation.

Authorship

Contribution: D.A.H. conceived the study design, conducted statistical analysis and interpretation of the SNP array data, cowrote the primary manuscript draft, and revised the final report; K.E.B. conceived the study, organized the case and control cohort demographic data, and critically reviewed the manuscript; R.A.M., A.K.K., C.B.D., N.E., J.L.M., U.F., J.Y.T., H.(M.)N., H.Z., L.G., and J.L.Z. contributed to the intellectual content and design of the GWAS study and supporting analyses and critically reviewed and edited the manuscript; J.G. conceived the study design, analyzed the data, cowrote the primary draft of the manuscript, and revised the final report; and all authors gave final approval of the manuscript.

Conflict-of-interest disclosure: D.A.H., K.E.B., A.K.K., C.B.D., N.E., J.L.M., U.F., and J.Y.T. are current or former employees of, and own stock or stock options in, 23andMe, Inc R.A.M., J.L.Z., and J.G. are unpaid advisors to, and collaborators with, 23andMe, Inc The remaining authors declare no competing financial interests.

Correspondence: David A. Hinds, 23andMe, Inc, 1390 Shorebird Way, Mountain View, CA 94043; e-mail: dhinds@23andme.com; and Jason Gotlib, Stanford Cancer Institute, 875 Blake Wilbur Dr, Room 2324, Stanford, CA 94305-5821; e-mail: jason.gotlib@stanford.edu.

References

- Baxter EJ, Scott LM, Campbell PJ, et al; Cancer Genome Project. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet*. 2005;365(9464):1054-1061.
- James C, Ugo V, Le Couédic JP, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythemia vera. *Nature*. 2005;434(7037):1144-1148.
- 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(4):387-397.
- Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med*. 2005;352(17):1779-1790.
- Vannucchi AM, Antonioli E, Guglielmelli P, Pardanani A, Tefferi A. Clinical correlates of JAK2V617F presence or allele burden in myeloproliferative neoplasms: a critical reappraisal. *Leukemia*. 2008;22(7):1299-1307.
- Scott LM, Tong W, Levine RL, et al. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *N Engl J Med*. 2007;356(5):459-468.
- Pikman Y, Lee BH, Mercher T, et al. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med*. 2006;3(7):e270.
- Pardanani AD, Levine RL, Lasho T, et al. MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood*. 2006;108(10):3472-3476.
- Grand FH, Hidalgo-Curtis CE, Ernst T, et al. Frequent CBL mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms. *Blood*. 2009;113(24):6182-6192.
- Oh ST, Simonds EF, Jones C, et al. Novel mutations in the inhibitory adaptor protein LNK drive JAK-STAT signaling in patients with myeloproliferative neoplasms. *Blood*. 2010;116(6):988-992.
- Klampff T, Gisslinger H, Harutyunyan AS, et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *N Engl J Med*. 2013;369(25):2379-2390.
- Nangalia J, Massie CE, Baxter EJ, et al. Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *N Engl J Med*. 2013;369(25):2391-2405.
- Kim E, Abdel-Wahab O. Focus on the epigenome in the myeloproliferative neoplasms. *Hematology Am Soc Hematol Educ Program*. 2013;2013(1):538-544.
- Landgren O, Goldin LR, Kristinsson SY, Helgadóttir EA, Samuelsson J, Björkholm 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(6):2199-2204.
15. Jones AV, Chase A, Silver RT, et al. JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. *Nat Genet*. 2009;41(4):446-449.
 16. Kilpivaara O, Mukherjee S, Schram AM, et al. A germline JAK2 SNP is associated with predisposition to the development of JAK2 (V617F)-positive myeloproliferative neoplasms. *Nat Genet*. 2009;41(4):455-459.
 17. Olcaydu D, Harutyunyan A, Jäger R, et al. A common JAK2 haplotype confers susceptibility to myeloproliferative neoplasms. *Nat Genet*. 2009;41(4):450-454.
 18. Jones AV, Campbell PJ, Beer PA, et al. The JAK2 46/1 haplotype predisposes to MPL-mutated myeloproliferative neoplasms. *Blood*. 2010;115(22):4517-4523.
 19. Tefferi A, Lasho TL, Patnaik MM, et al. JAK2 germline genetic variation confers disease susceptibility in primary myelofibrosis regardless of V617F mutational status: nullizygosity for the JAK2 46/1 haplotype is associated with inferior survival. *Leukemia*. 2010;24(1):105-109.
 20. Pardanani A, Lasho TL, Finke CM, et al. The JAK2 46/1 haplotype confers susceptibility to essential thrombocythemia regardless of JAK2V617F mutational status-clinical correlates in a study of 226 consecutive patients. *Leukemia*. 2010;24(1):110-114.
 21. Nielsen C, Bojesen SE, Nordestgaard BG, Kofoed KF, Birgens HS. JAK2V617F somatic mutation in the general population: myeloproliferative neoplasm development and progression rate. *Haematologica*. 2014;99(9):1448-1455.
 22. Nielsen C, Birgens HS, Nordestgaard BG, Bojesen SE. Diagnostic value of JAK2 V617F somatic mutation for myeloproliferative cancer in 49 488 individuals from the general population. *Br J Haematol*. 2013;160(1):70-79.
 23. Xu X, Zhang Q, Luo J, et al. JAK2(V617F): Prevalence in a large Chinese hospital population. *Blood*. 2007;109(1):339-342.
 24. Passamonti F, Rumi E, Pietra D, Lazzarino M, Cazzola M. JAK2 (V617F) mutation in healthy individuals. *Br J Haematol*. 2007;136(4):678-679.
 25. Sidon P, El Housni H, Dessars B, Heimann P. The JAK2V617F mutation is detectable at very low level in peripheral blood of healthy donors. *Leukemia*. 2006;20(9):1622.
 26. Nielsen C, Birgens HS, Nordestgaard BG, Kjaer L, Bojesen SE. The JAK2 V617F somatic mutation, mortality and cancer risk in the general population. *Haematologica*. 2011;96(3):450-453.
 27. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 2014;371(26):2488-2498.
 28. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 2014;371(26):2477-2487.
 29. Xie M, Lu C, Wang J, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med*. 2014;20(12):1472-1478.
 30. Abecasis GR, Altshuler D, Auton A, et al; 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing [published correction appears in *Nature*. 2011;473(7348):544]. *Nature*. 2010;467(7319):1061-1073.
 31. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*. 2007;81(5):1084-1097.
 32. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet*. 2012;44(8):955-959.
 33. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*. 2003;164(4):1567-1587.
 34. Henn BM, Hon L, Macpherson JM, et al. Cryptic distant relatives are common in both isolated and cosmopolitan genetic samples. *PLoS One*. 2012;7(4):e34267.
 35. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909.
 36. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55(4):997-1004.
 37. Steensma DP, Bejar R, Jaiswal S, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015;126(1):9-16.
 38. Van Dyke AL, Cote ML, Wenzlaff AS, et al. Chromosome 5p region SNPs are associated with risk of NSCLC among women. *J Cancer Epidemiol*. 2009;2009:242151.
 39. Shiraiishi K, Kunitho H, Daigo Y, et al. A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population. *Nat Genet*. 2012;44(8):900-903.
 40. Bojesen SE, Pooley KA, Johnatty SE, et al; Australian Cancer Study; Australian Ovarian Cancer Study; Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer (kConFab); Gene Environment Interaction and Breast Cancer (GENICA); Swedish Breast Cancer Study (SWE-BRCA); Hereditary Breast and Ovarian Cancer Research Group Netherlands (HEBON); Epidemiological study of BRCA1 & BRCA2 Mutation Carriers (EMBRACE); Genetic Modifiers of Cancer Risk in BRCA1/2 Mutation Carriers (GEMO). Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat Genet*. 2013;45(4):371-384.
 41. Mocellin S, Verdi D, Pooley KA, et al. Telomerase reverse transcriptase locus polymorphisms and cancer risk: a field synopsis and meta-analysis. *J Natl Cancer Inst*. 2012;104(11):840-854.
 42. Hinds DA, Barnholt KE, Zehnder JL, et al. A germline variant in the TERT gene is a novel predisposition allele associated with myeloproliferative neoplasms [abstract]. *Blood*. 2012;120(21). Abstract 707.
 43. Oddsson A, Kristinsson SY, Helgason H, et al. The germline sequence variant rs2736100_C in TERT associates with myeloproliferative neoplasms. *Leukemia*. 2014;28(6):1371-1374.
 44. Tapper W, Jones AV, Kralovics R, et al. Genetic variation at MECOM, TERT, JAK2 and HBS1L-MYB predisposes to myeloproliferative neoplasms. *Nat Commun*. 2015;6:6691.
 45. Lesteven E, Picque M, Conejero Tonetti C, et al. Association of a single-nucleotide polymorphism in the SH2B3 gene with JAK2V617F-positive myeloproliferative neoplasms. *Blood*. 2014;123(5):794-796.
 46. Tong W, Lodish HF. Lnk inhibits Tpo-*mpl* signaling and Tpo-mediated megakaryocytopoiesis. *J Exp Med*. 2004;200(5):569-580.
 47. Bersenev A, Wu C, Balcerak J, Tong W. Lnk controls mouse hematopoietic stem cell self-renewal and quiescence through direct interactions with JAK2. *J Clin Invest*. 2008;118(8):2832-2844.
 48. Velazquez L, Cheng AM, Fleming HE, et al. Cytokine signaling and hematopoietic homeostasis are disrupted in Lnk-deficient mice. *J Exp Med*. 2002;195(12):1599-1611.
 49. Delhommeau F, Dupont S, Della Valle V, et al. Mutation in TET2 in myeloid cancers. *N Engl J Med*. 2009;360(22):2289-2301.
 50. Abdel-Wahab O, Mullally A, Hedvat C, et al. Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood*. 2009;114(1):144-147.
 51. Ortmann CA, Kent DG, Nangalia J, et al. Effect of mutation order on myeloproliferative neoplasms. *N Engl J Med*. 2015;372(7):601-612.
 52. Busque L, Patel JP, Figueroa ME, et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet*. 2012;44(11):1179-1181.
 53. Bell DW, Varley JM, Szydo TE, et al. Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. *Science*. 1999;286(5449):2528-2531.
 54. Meijers-Heijboer H, van den Ouweland A, Klijn J, et al; CHEK2-Breast Cancer Consortium. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet*. 2002;31(1):55-59.
 55. Webb EL, Rudd MF, Sellick GS, et al. Search for low penetrance alleles for colorectal cancer through a scan of 1467 non-synonymous SNPs in 2575 cases and 2707 controls with validation by kin-cohort analysis of 14 704 first-degree relatives. *Hum Mol Genet*. 2006;15(21):3263-3271.
 56. Fletcher O, Johnson N, dos Santos Silva I, et al; kConFab Investigators; AOCs Group; GENICA Consortium; Breast Cancer Association Consortium. Missense variants in ATM in 26,101 breast cancer cases and 29,842 controls. *Cancer Epidemiol Biomarkers Prev*. 2010;19(9):2143-2151.
 57. Rudd MF, Sellick GS, Webb EL, Catovsky D, Houlston RS. Variants in the ATM-BRCA2-CHEK2 axis predispose to chronic lymphocytic leukemia. *Blood*. 2006;108(2):638-644.
 58. Collado M, Landt O, Barragan E, et al. Locked nucleic acid-enhanced detection of 1100delC-CHEK2 germ-line mutation in Spanish patients with hematologic malignancies. *Clin Chem*. 2004;50(11):2201-2204.
 59. Janiszewska H, Bąk A, Pilarska M, et al. A risk of essential thrombocythemia in carriers of constitutional CHEK2 gene mutations. *Haematologica*. 2012;97(3):366-370.
 60. Shiekhattar R. PINTing for p53. *Genome Biol*. 2013;14(9):132-134.
 61. Randrianarison-Huetz V, Laurent B, Bardet V, Blobe GC, Huetz F, Duménil D. Gfi-1B controls human erythroid and megakaryocytic differentiation by regulating TGF-beta signaling at the bipotent erythro-megakaryocytic progenitor stage. *Blood*. 2010;115(14):2784-2795.
 62. Monteferrario D, Bolar NA, Marneth AE, et al. A dominant-negative GFI1B mutation in the gray platelet syndrome. *N Engl J Med*. 2014;370(3):245-253.
 63. Jacobs KB, Yeager M, Zhou W, et al. Detectable clonal mosaicism and its relationship to aging and cancer. *Nat Genet*. 2012;44(6):651-658.
 64. Laurie CC, Laurie CA, Rice K, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet*. 2012;44(6):642-650.
 65. Mehta J, Wang H, Iqbal SU, Mesa R. Epidemiology of myeloproliferative neoplasms in the United States. *Leuk Lymphoma*. 2014;55(3):595-600.