



Published in final edited form as:

Leukemia. 2014 April ; 28(4): 935–938. doi:10.1038/leu.2014.7.

Whole-exome sequencing of polycythemia vera revealed novel driver genes and somatic mutation shared by T-cells and granulocytes

Linghua Wang¹, Sabina I. Swierczek², Jennifer Drummond¹, Kimberly Hickman², Kimberly Walker¹, Harshavardhan Doddapaneni¹, Donna M. Muzny¹, Richard A. Gibbs¹, David A. Wheeler^{1,*}, and Josef T. Prchal^{2,*}

¹Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030

²Division of Hematology, The University of Utah School of Medicine and VAH, Salt Lake City, UT 84132. Josef Prchal also represents investigators of MPD-RC (Myeloproliferative Disorders Research Consortium)

Abstract

To better understand the underlying molecular basis of polycythemia vera (PV), we performed whole-exome sequencing and DNA copy-number analysis of 31 *JAK2*^{V617F}-positive patients and further investigated the evolution of somatic mutations using longitudinal samples. In addition to *JAK2*^{V617F} and 9pUPD, we identified frequent recurrent somatic mutation in *ASXL1*, *TET2*, *DNMT3A*, *SF3B1* and *NFI*. Forty two percent of patients had a somatic mutation in at least one epigenetic modifier gene. In 4 of 31 patients, variant allele abundance suggested mutation of *JAK2*^{V617F} was preceded by other somatic mutations including *ASXL1*, *DNMT3A* and *SF3B1*. Strikingly, in 7 patients, apparent germline variants were detected at COSMIC codons in one or more PV-related genes in which we had also discovered somatic mutations across the cohort, suggesting that some pre-*JAK2*^{V617F} mutations contribute to substantial T-lymphocyte progeny. This study contributes to novel understanding of the complexity of PV pathogenesis.

Polycythemia vera (PV) is a myeloproliferative neoplasm (MPN) characterized by an elevated red cell mass caused by excessive myelopoiesis with propensity to transformation to myelofibrosis and acute leukemia¹. *JAK2*^{V617F} mutation² and acquired uniparental disomy on chromosome 9p (9pUPD)^{3, 4} are the most frequent somatic alterations. *JAK2* is mutated in over 90% of PV patients²; however, in spite of its prevalence it is regarded as necessary but not sufficient to cause the disease⁵. Yet, systematic mutation analysis of exome coding sequence has to be described. We here analyzed 31 *JAK2*^{V617F}-positive patients by

*To whom correspondence should be addressed: David A. Wheeler, Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Suite Alkt-N1619, MS: BCM226, Houston TX, 77030, USA, Phone: +1-713-798-7206, Fax: +1-713-798-5741, wheeler@bcm.edu and Josef T. Prchal, MD, Division of Hematology, 30N 1900E, 5C402 SOM, University of Utah, Salt Lake City, UT 84132, USA, Phone: +1-801-581-4220, Fax: +1-801-585-3432, josef.prchal@hsc.utah.edu.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Supplementary information is available at Nature Leukemia's website.

whole-exome sequencing (WES) and SNP array and further investigated the mutational evolution using longitudinal samples collected from 7 patients (Supplementary Methods).

WES targeting ~43 Mb of coding sequence yielded an average of 125-fold coverage (Supplementary Figure 1), with 95% of targeted bases covered to a depth of 20x (Supplementary Figure 2). In this study, we use the exomes of T cells as a reference genome and we defined sequence variants present in granulocytes but not in the T-cells from the same patient as somatic mutations. DNA copy-number analysis revealed multiple recurrent chromosomal arm-length changes including 9p UPD, trisomy of 9p and 1q, LOH at 18p, and single patients with 2p and 9q LOH, trisomy 2q and UPD at 22q. Through serial sampling, we observed a focal hemizygous deletion in PV28 (Supplementary Figure 3) at 9q22.31-32. No known tumor suppressors or oncogenes were found in this region and its functional importance is unclear at this time.

We validated 87 somatic non-silent mutations and a subset of important germline variants (Supplementary Table 1). The average non-silent somatic mutation rate was 0.08 per megabase (Mb), which was much lower than other adult cancers⁶. G->T transversion was the most frequent base change, followed by C->T transition (Supplementary Figure 4). The recurrent G1849T mutation leading to *JAK2*^{V617F} accounted for the majority of G->T transversions since all patients harbored *JAK2*^{V617F}. The most frequently mutated class of genes (Figure 1) was that involved in epigenetic modification. Forty-two percent of patients had at least one somatic mutation in an epigenetic modifier, indicating epigenetic regulation plays a key role in PV pathogenesis, as it does in other myeloid neoplasms. Many other known cancer-related genes were mutated in single patients as shown in Figure 1.

Recurrent somatic mutations were observed in: *ASXL1*, *DNMT3A*, *TET2*, *SF3B1* and *PDE4C* (Figure 1). *ASXL1* belongs to the Enhancer of trithorax and Polycomb gene families, which are responsible for maintaining activation and silencing of PcG and trxG proteins⁷. Disruption of *ASXL1*, through inactivating mutations in exon 12 is prevalent in post-PV myelofibrosis, but is identified only rarely in cases of essential thrombocytosis or PV⁸. Stein and coworkers sequenced exon 12 of *ASXL1* in 42 PV cases but found only one patient (2%) with a nonsense mutation⁸. We identified 4 inactivating somatic mutations in *ASXL1* (12.9%), 2 frame-shift and 2 nonsense. All 4 loss-of-function mutations were identified in exon 12. This was a 6-fold higher mutation rate than previously reported¹⁰ and is similar to other MPN.

Somatic *DNMT3A* mutations were reported at low frequency in PV (2.7%)⁹. The reported mutations were identified in the terminal exon, at position M880 and R882. In this study, we identified 3 somatic *DNMT3A* mutations (9.7%), one was identified at the known hotspot, R882, and the other two were novel frame-shift mutations at codon K456. *SF3B1* encodes subunit 1 of the splicing factor 3b, which is important for anchoring the spliceosome to precursor mRNA. Mutation of *SF3B1* is frequent in most MPN, having been reported in myelodysplasia with ring sideroblasts (65%)¹⁰, myelodysplastic syndrome (20%)¹⁰, primary myelofibrosis (7%)¹¹ and essential thrombocythaemia (3%), but it has not been reported in PV¹⁰. In this study, we identified 3 *SF3B1* mutations in 2 patients (9.7%), patient PV5 carried two mutations and both were reported by COSMIC (<http://cancer.sanger.ac.uk/>

cancergenome/projects/cosmic/). Interestingly, phosphodiesterase 4C, *PDE4C*, is mutated in two patients. *PDE4C* hydrolyzes the second messenger cAMP and therefore mediates a variety of responses to extracellular signals. Although mutation in this gene is rarely observed in cancer, one of the mutations we discovered was reported in COSMIC, suggesting it may be functionally relevant.

The fraction of reads with a given mutation, the variant allele fraction, is proportional to the number of nuclei in the tumor sample harboring the mutation. Since the granulocytes in PV patients are clonal, by X-inactivation in females¹², the variant allele fraction of the mutations reported in Figure 1, should correspond to the order in which they appeared in the patient. Three patients, PV5, PV8 and PV24, exhibited tumor variant allele fraction in the key epigenetic modifier genes, *ASXL1*, *DNMT3A*, or *SF3B1* that were higher than *JAK2*^{V617F} (Figure 2A). Interestingly mutation in a gene associated with immunosuppression in solid tumors, *IDO1*, is also found at higher variant allele burden than that of *JAK2*^{V617F}. These results suggest that mutation of these genes might have preceded the acquisition of *JAK2*^{V617F}.

In 7 patients we could determine the order of appearance of mutations directly by longitudinal sampling (Figure 2B). Patients PV1, PV3, PV8, PV10, PV23 and PV29, harbored only *JAK2*^{V617F} mutation in 2011. Upon follow-up in 2013, four of them had acquired additional mutations, particularly in key epigenetic modifier genes *DNMT3A*, *SMARCA2* and *TET2*, but also in *SF3B1*, *STK11*, *CDH3* and *TP53*. One patient harbored *IDH2* and *JAK2* in 2011, but did not acquire any additional mutations. Results from a parallel study of these patients showed that 10% of them acquired 9p UPD prior to *JAK2*^{V617F} mutations (Figure 1, Group III patients), whereas 42% had acquired *JAK2*^{V617F} mutation first (Figure 1, Group II, AMBER13-LEU-1191).

Novel sequence variants found in both granulocytes and T-cells from the same patients are putative germline mutations. However, in 7 patients, these putative germline variants were in genes that were somatically mutated in other patients in the cohort (Figure 1, blue tiles). Moreover, a high proportion of these so-called germline mutations were likely to be functionally relevant either because they were truncating frameshift or nonsense mutations or the same mutations could be found in COSMIC. For example, the tumor suppressor *NF1* was mutated in 4 patients' T-cells and granulocytes. Among them, 3 variants are presented in COSMIC. Germline mutation of *NF1* is linked to neurofibromatosis type 1, a debilitating dominant genetic disorder characterized by a higher risk for juvenile myelomonocytic leukemia with a potential progression to acute myeloid leukemia (AML)¹³. Symptoms of neurofibromatosis type 1 were not observed in our PV patients, thus it is highly unlikely these patients have true germline mutation in this gene. Similar variants were also found in two patients in *ASXL1*, and one patient each in *TET2*, *DNMT3A* and *IDH2*. The *IDH2* variant R140Q is a hotspot for somatic mutation in AML and other cancers. Based on these results, *NF1* emerges as a frequently mutated gene (16%) in PV; *ASXL1* mutated in 19% of patients, 5-fold higher than previously reported ($P = 0.02$, Fisher's exact test); and *DNMT3A* and *TET2* each mutated in 13% of our cohort.

These mutations could not be explained by contamination of the T-cells by granulocytes because the T-cells harbored little or no *JAK2*^{V617F}, whereas these mutations are present in similar variant allele fractions in the T-cells and granulocytes (Supplementary Figure 5). Therefore we conclude these were somatic mutations in long-lived T cells. For the mutations to have populated both the T-cell and granulocyte lineages in equal proportions, they must have occurred in pluripotent stem cells early in the lives of these patients, well before the *JAK2*^{V617F} mutation that favors myeloid differentiation.

The involvement of T cells by some of the somatic mutations was unexpected. Studies based on the distribution of X-chromosome allelic usage have been used to assess the relative contribution of pluripotent stem cells to B and T cell lymphopoiesis versus myelopoiesis. Indeed, in normal healthy females the X-chromosome allelic usage of myeloid and lymphoid progeny was identical suggesting that these are all derived from relative small number of hematopoietic stem cells present in female embryo at the time of X-chromosome inactivation¹⁴. However, the majority of T cells, unlike myeloid progeny in PV are polyclonal¹⁵ suggesting that the mutations were observed were derived from stem cells prior to the clonal event necessary for discernible PV phenotype. However, a small proportion of T cells may be part of the PV clone³, and *JAK2*^{V617F} mutant cells have preferentially myeloid potential but not exclusively as some *JAK2*^{V617F} lymphoid progeny are also formed¹⁶. In this report we present a novel observation that some pre- *JAK2*^{V617F} somatic mutations contribute to substantial T-lymphocyte progeny.

In summary, we sequenced 31 PV pairs of granulocytes and T cells presumed to have germline genome and identified multiple novel driver genes and pre-*JAK2* mutations and described the signatures of clonal evolution during PV progression in some patients. This study contributes to our understanding of the pathogenesis of PV and underscores the complexity of genetic lesions in the typical PV genome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by research funding from the National Human Genome Research Institute (NHGRI, grant number: 5U54HG003273) to DW and from the National Institutes of Health (NIH, grant number: NIH-P01CA108671) to JP. We thank Christian Buhay, Donna Morton, Huyen Dinh, Ritika Raj, Lora Lewis, Christie Kovar, Sandra Lee, Michelle Bellair and Zhu Yiming for their excellent technical support.

References

1. Tefferi A, Spivak JL. Polycythemia vera: scientific advances and current practice. *Seminars in hematology*. 2005; 42(4):206–20. [PubMed: 16210034]
2. James C, Ugo V, Le Couedic JP, Staerk J, Delhommeau F, Lacout C, et al. A unique clonal *JAK2* mutation leading to constitutive signalling causes polycythaemia vera. *Nature*. 2005; 434(7037): 1144–8. [PubMed: 15793561]
3. Kralovics R, Guan Y, Prchal JT. Acquired uniparental disomy of chromosome 9p is a frequent stem cell defect in polycythemia vera. *Experimental hematology*. 2002; 30(3):229–36. [PubMed: 11882360]

4. Wang K, Swierczek S, Hickman K, Hakonarson H, Prchal JT. Convergent mechanisms of somatic mutations in polycythemia vera. *Discovery medicine*. 2011; 12(62):25–32. [PubMed: 21794206]
5. Nussenzweig RH, Swierczek SI, Jelinek J, Gaikwad A, Liu E, Verstovsek S, et al. Polycythemia vera is not initiated by JAK2V617F mutation. *Experimental hematology*. 2007; 35(1):32–8. [PubMed: 17198871]
6. Wheeler DA, Wang L. From human genome to cancer genome: The first decade. *Genome Res*. 2013; 23(7):1054–62. [PubMed: 23817046]
7. Fisher CL, Pineault N, Brookes C, Helgason CD, Ohta H, Bodner C, et al. Loss-of-function Additional sex combs like 1 mutations disrupt hematopoiesis but do not cause severe myelodysplasia or leukemia. *Blood*. 2010; 115(1):38–46. [PubMed: 19861679]
8. Stein BL, Williams DM, O'Keefe C, Rogers O, Ingersoll RG, Spivak JL, et al. Disruption of the ASXL1 gene is frequent in primary, post-essential thrombocytosis and post-polycythemia vera myelofibrosis, but not essential thrombocytosis or polycythemia vera: analysis of molecular genetics and clinical phenotypes. *Haematologica*. 2011; 96(10):1462–9. [PubMed: 21712540]
9. Rao N, Butcher CM, Lewis ID, Ross DM, Melo JV, Scott HS, et al. Clonal and lineage analysis of somatic DNMT3A and JAK2 mutations in a chronic phase polycythemia vera patient. *British journal of haematology*. 2012; 156(2):268–70. [PubMed: 21859430]
10. Papaemmanuil E, Cazzola M, Boulwood J, Malcovati L, Vyas P, Bowen D, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *The New England journal of medicine*. 2011; 365(15):1384–95. [PubMed: 21995386]
11. Lasho TL, Finke CM, Hanson CA, Jimma T, Knudson RA, Ketterling RP, et al. SF3B1 mutations in primary myelofibrosis: clinical, histopathology and genetic correlates among 155 patients. *Leukemia: official journal of the Leukemia Society of America, Leukemia Research Fund, UK*. 2012; 26(5):1135–7.
12. Swierczek SI, Agarwal N, Nussenzweig RH, Rothstein G, Wilson A, Artz A, et al. Hematopoiesis is not clonal in healthy elderly women. *Blood*. 2008; 112(8):3186–93. [PubMed: 18641369]
13. Side L, Taylor B, Cayouette M, Conner E, Thompson P, Luce M, et al. Homozygous inactivation of the NF1 gene in bone marrow cells from children with neurofibromatosis type 1 and malignant myeloid disorders. *The New England journal of medicine*. 1997; 336(24):1713–20. [PubMed: 9180088]
14. Prchal JT, Prchal JF, Belickova M, Chen S, Guan Y, Gartland GL, et al. Clonal stability of blood cell lineages indicated by X-chromosomal transcriptional polymorphism. *The Journal of experimental medicine*. 1996; 183(2):561–7. [PubMed: 8627167]
15. Prchal JT, Guan YL, Prchal JF, Barany F. Transcriptional analysis of the active X-chromosome in normal and clonal hematopoiesis. *Blood*. 1993; 81(1):269–71. [PubMed: 8093348]
16. Jamieson CH, Gotlib J, Durocher JA, Chao MP, Mariappan MR, Lay M, et al. The JAK2 V617F mutation occurs in hematopoietic stem cells in polycythemia vera and predisposes toward erythroid differentiation. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(16):6224–9. [PubMed: 16603627]

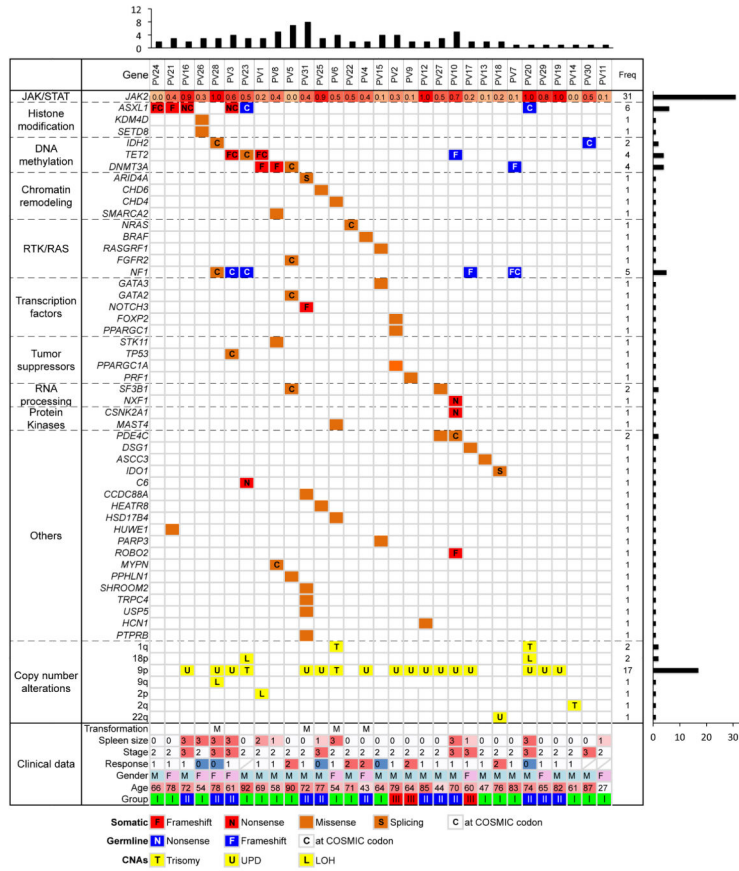


Figure 1. Summary of somatic mutations, germline variants and somatic DNA copy-number alterations identified in 31 PV cases
 Somatic mutations were shown in orange and red based on their mutation type and germline variants were shown in blue. For germline variants, only those frame-shift variants, nonsense variants or those variants reported in COSMIC database (<http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/>) were shown and only those genes with somatic mutations identified in the cohort were listed. The total number of somatic mutations identified per patient and gene was shown at the top and right, respectively. The allelic fraction of *JAK2*^{V617F} was shown at the top row and color scaled. The group classification was based on the relationship of *JAK2*^{V617F} and 9p aUPD. Subgroup I, patients harbored *JAK2*^{V617F} in a heterozygous state without detectable 9p aUPD; Subgroup II, patients had *JAK2*^{V617F} with an allelic fraction in direct proportion to the level of 9p aUPD; Subgroup III, patients harbored 9p aUPD at approximately twice the level of the *JAK2*^{V617F} allelic burden. The clinical data were summarized at the bottom. CNAs, somatic DNA copy-number alterations: T, trisomy; L, hemizygous deletion; U, uniparental disomy. Transformation: M, myelofibrosis. Spleen size: 0, the size of spleen is not palpable; 1, < 3cm; 2, < 6cm; 3, > 6cm. Stage: 1, low; 2, intermediate; 3, high. Response to therapy: 0, no response; 1, partial response; 2, complete response. Gender: M, male; F, female.

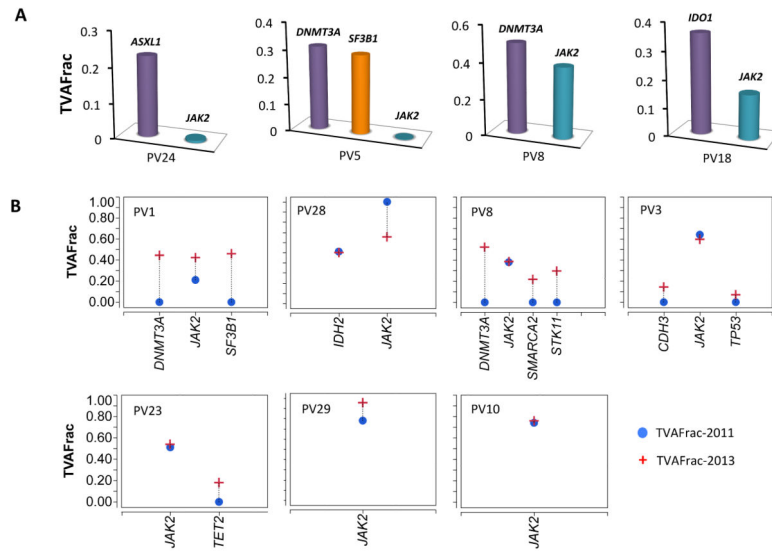


Figure 2. The pre-*JAK2* mutations and signatures of mutational evolution
(A) Patients with pre-*JAK2* somatic mutations. **(B)** Changes of somatic mutations during PV progression. The longitudinal samples were collected from 7 PV patients during a 2-year period from 2011 to 2013. TVAFrac, the variant allelic fraction in granulocytes.