Received: 7 October 2018 | Revised: 2 April 2019 | Accepted: 4 April 2019

DOI: 10.1002/mgg3.772

ORIGINAL ARTICLE

Molecular Genetics & Genomic Medicine Open Access WILEY

TET2 missense variants in human neoplasia. A proposal of structural and functional classification

Elena Bussaglia[1] | **Rosa Antón**[2] | **Josep F. Nomdedéu**[1] | **Pablo Fuentes-Prior**[2]

[1]Hematology Department and Diagnostic Hematology Group, Barcelona, Spain

[2]Molecular Bases of Disease, The Biomedical Research Institute Sant Pau (IIB Sant Pau), Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

Correspondence
Josep F. Nomdedéu and Pablo Fuentes-Prior, The Biomedical Research Institute Sant Pau (IIB Sant Pau), Hospital de la Santa Creu i Sant Pau, Sant Antoni Maria Claret 167, 08025 Barcelona, Spain.
Email: jnomdedeu@santpau.cat (J. F. N) and pfuentes@santpau.cat (P. F.-P.)

Funding information
Instituto de Salud Carlos III, Grant/Award Number: FIS PI 10/1739 (J.F.N.) and FIS PI 13/2729 (J.F.N.); Secretaría de Estado de Investigación, Desarrollo e Innovación, Grant/Award Number: SAF2014-57994-R (P.F.-P.); Agència de Gestió d'Ajuts Universitaris i de Recerca, Grant/Award Number: SGR383 (J.F.N.)

Abstract

Background: The human *TET2* gene plays a pivotal role in the epigenetic regulation of normal and malignant hematopoiesis. Somatic *TET2* mutations have been repeatedly identified in age-related clonal hematopoiesis and in myeloid neoplasms ranging from acute myeloid leukemia (AML) to myeloproliferative neoplasms. However, there have been no attempts to systematically explore the structural and functional consequences of the hundreds of *TET2* missense variants reported to date.

Methods: We have sequenced the *TET2* gene in 189 Spanish AML patients using Sanger sequencing and NGS protocols. Next, we performed a thorough bioinformatics analysis of TET2 protein and of the expected impact of all reported *TET2* missense variants on protein structure and function, exploiting available structure-and-function information as well as 3D structure prediction tools.

Results: We have identified 38 *TET2* allelic variants in the studied patients, including two frequent SNPs: p.G355D (10 cases) and p.I1762V (28 cases). Four of the detected mutations are reported here for the first time: c.122C>T (p.P41L), c.4535C>G (p.A1512G), c.4760A>G (p.D1587G), and c.5087A>T (p.Y1696F). We predict a complex multidomain architecture for the noncatalytic regions of TET2, and in particular the presence of well-conserved α+β globular domains immediately preceding and following the actual catalytic unit. Further, we provide a rigorous interpretation of over 430 missense SNVs that affect the TET2 catalytic domain, and we hypothesize explanations for ~700 additional variants found within the regulatory regions of the protein. Finally, we propose a systematic classification of all missense mutants and SNPs reported to date into three major categories (severe, moderate, and mild), based on their predicted structural and functional impact.

Conclusions: The proposed classification of missense *TET2* variants would help to assess their clinical impact on human neoplasia and may guide future structure-and-function investigations of TET family members.

KEYWORDS
5-methylcytosine, classification of mutations, epigenetic regulation, neoplasia, TET2

1 | INTRODUCTION

Methylation at the C5 position of cytosine bases to generate 5-methylcytosine (5mC) is one of the major epigenetic modifications of mammalian genomes, with a profound impact on chromatic structure and gene expression (Feng, Jacobsen, & Reik, 2010; Smith & Meissner, 2013). Members of the Ten-Eleven-Translocation (TET) subfamily of Fe(II)- and 2-oxoglutarate (2-OG)-dependent dioxygenases (EC 1.14.11. n2) help to revert this modification by iteratively oxidizing 5mC bases first to 5-hydroxymethylcytosine (5hmC) (Ito et al., 2010; Tahiliani et al., 2009), and then to 5-formylcytosine (5fC) and 5-carboxy-cytosine (5caC) (Ito et al., 2011). Since both 5fC and 5caC can be removed through the base excision-repair pathway (Cortellino et al., 2011; He et al., 2011; Ito et al., 2010,2011; Ko et al., 2010; Rampal et al., 2014), TET enzymes catalyze the first step in active DNA demethylation, and therefore play an important role in transcription regulation. In addition, 5hmC bases function as stable, independent epigenetic marks, which have been shown to accumulate for example at sites of DNA damage and to promote genome stability (Kafer et al., 2016). On the other hand, TET2 has been recently shown to oxidize specific mRNAs to promote infection-induced myelopoiesis (Shen et al., 2018).

TET2 has been shown to possess important functions that are independent of its enzymatic activity. For instance, TET2 involvement in the resolution of inflammation depends upon the ability of the enzyme to recruit the histone deacetylase, HDAC2, to specifically repress *IL-6* transcription (Zhang et al., 2015). Further, both catalytic and noncatalytic activities of TET2 are required for mast cell differentiation and proliferation (Montagner et al., 2016). Finally, it has been recently reported that the enzyme regulates age-related regenerative decline in the aging hippocampus, with important implications for neurogenic rejuvenation (Gontier et al., 2018). (For authoritative reviews on the function of TET proteins see e.g. refs. [Pastor, Aravind, & Rao, 2013; Rasmussen & Helin, 2016]).

TET2 is expressed at particularly high levels in hematopoietic cells, and is critically relevant for normal hematopoiesis. *Tet2* deletion in mice models causes an enlargement of the myeloid immature compartment raising the risk of acquiring additional transforming mutations (Li et al., 2011; Moran-Crusio et al., 2011; Quivoron et al., 2011). Not surprisingly, numerous *TET2* mutations have been described in a wide range of human myeloid malignancies ranging from acute myeloid leukemia (AML) to myeloproliferative neoplasms (MPN) (see e.g. [Abdel-Wahab et al., 2009; Delhommeau et al., 2009; Langemeijer et al., 2009]). *TET2* mutations are also common in some subgroups of mature T-cell lymphomas (Quivoron et al., 2011; Zang et al., 2017). In addition to hematological and lymphoid neoplasms, *TET2* variants have been more recently reported in nearly all cancer types, most notably colorectal, lung, and skin carcinomas. (For a complete list of articles describing *TET2* mutations see Table S2; see ref. (Ko et al., 2015) for a recent review of TET proteins and their role in hematological cancers). *TET2* mutation causes an expansion of immature hematopoietic precursors with a bias toward myeloid differentiation (Arends et al., 2018; Buscarlet et al., 2018).

TET2 alterations have also been described in age-related clonal hematopoiesis (ARCH; [Jaiswal et al., 2014]), as firstly demonstrated in elderly women (Busque et al., 2012). *TET2* mutations may also be detected in healthy women years before the development of leukemia (Desai et al., 2018). Importantly, somatic *TET2* mutations in normal elderly individuals with CH raise the mortality attributed to vascular events (Jaiswal et al., 2017), most likely by promoting an exacerbated atherosclerosis (Fuster et al., 2017). In this regard, TET2-deficiency results in an increased pro-inflammatory phenotype in murine macrophages, which may help to promote atherosclerosis (Abegunde, Buckstein, Wells, & Rauh, 2018; Cull, Snetsinger, Buckstein, Wells, & Rauh, 2017; Cull, Snetsinger, & Rauh, 2016). CH also alters cardiac repair by inflammatory pathways (Sano et al., 2018). Also in line with the inflammatory role of TET2, patients who underwent allogeneic bone marrow transplantation from donors with CH had an increased risk of chronic graft-versus-host disease (Frick et al., 2019). These findings are in line with TET2-mediated repression of inflammatory mediators such as IL-6, as mentioned above (Zhang et al., 2015). (For a recent review on CH and its relevance for hematopoietic malignancies, see [Bowman, Busque, & Levine, 2018]).

The *TET2* gene maps to chromosome 4q24 contains 11 exons, and has an open reading frame of 6009 nt. The encoded protein contains 2002 amino acid residues (isoform A, NM 001127208), with the actual catalytic unit located in the C-terminal part of the protein (Figure 1a). The three-dimensional (3D) structure of this TET2 region and the mechanism of 5mC oxidation have been recently characterized by X-ray crystallography (Hu et al., 2013, 2015). The structural work revealed that the catalytic unit is comprised of two tandem cysteine-rich modules, followed by the actual catalytic domain (Figure 1b,c). This domain, termed double-stranded β-helix (DSBH) for its major structural feature, contains a large, poorly conserved insertion between approximate positions Cys1464 and Glu1841. (All numbers correspond to the human enzyme). Thus, the structurally characterized fragment is an artificial construct in which residues Ser1481-Asn1843 are replaced by a 15-residues-long Gly/Ser linker, and which therefore corresponds to only ≈22% of the full-length protein. On the other hand, there is no structural information available for the long N-terminal region of the protein, Met1-Pro1131. Interestingly, a second transcript of the *TET2* gene encodes for a protein that is essentially comprised of

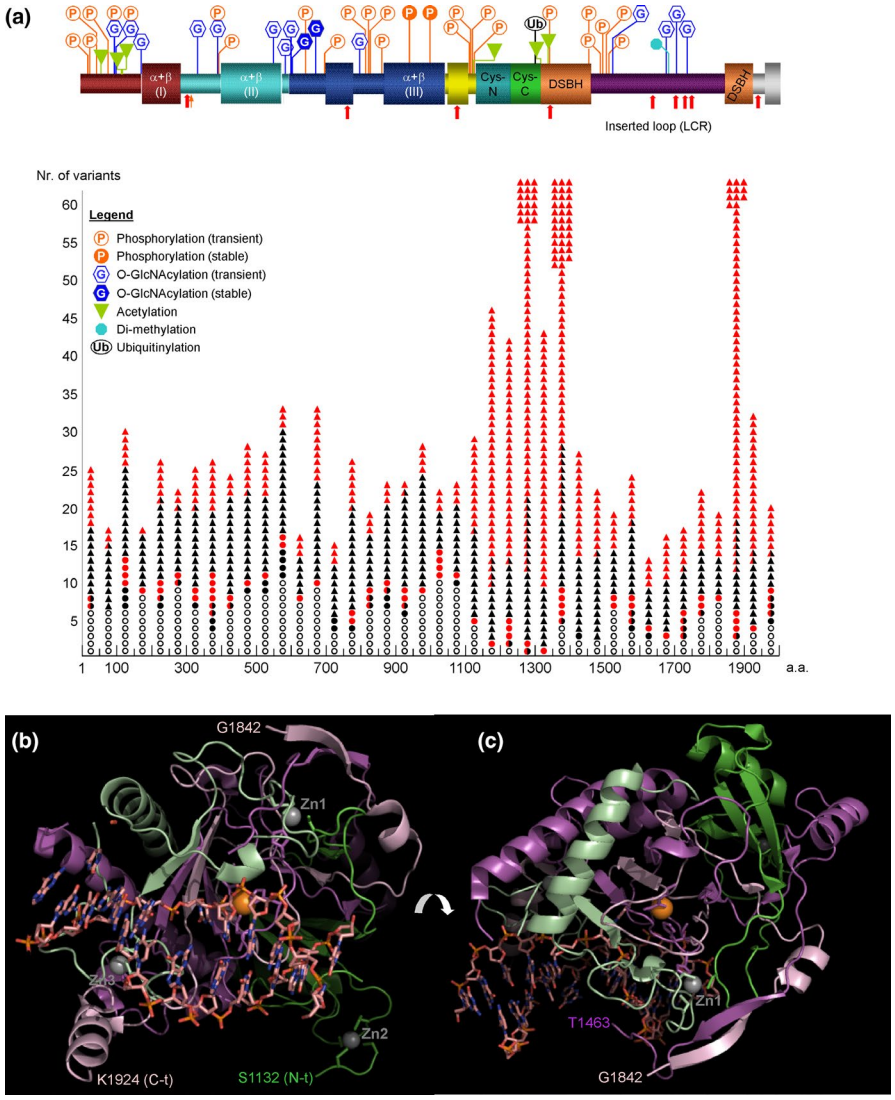


FIGURE 1 Domain organization and distribution of missense variants in human TET2. (a) The domain architecture of the TET2 protein is schematically represented on top of the figure. Posttranslational modification (PTM) sites were taken from public databases, or predicted, either by similarity with the mouse enzyme (Bauer et al., 2015) or using software given in Methods. Notice that the long polypeptide stretch preceding the actual catalytic unit is predicted to contain several putative, relatively well-conserved globular modules, separated by more variable PTM-rich linkers. Single-nucleotide variations reported in the human *TET2* gene are given below the scheme, with the following code: missense mutations are represented by triangles, colored red if identified in hematopoietic and lymphoid neoplasms or black, if reported in solid tumors. Validated SNPs are given as circles, which are colored red and black if they have been in addition reported in cancer patients, or left empty otherwise. (b and c) Two approximately perpendicular views of the three-dimensional structure of TET2 catalytic unit. Major structural domains are shown as ribbons, colored green (Cys-N module), light green (Cys-C), magenta and light pink (N- and C-terminal halves of the DSBH domain, respectively). Zn^{2+} and Fe^{2+} ions are represented as gray and orange spheres, respectively; the side chains of their coordinating residues are also shown. The bound DNA oligonucleotide is shown with all its non-hydrogen atoms, color-coded (carbon, light pink; nitrogen, blue; oxygen, red; and phosphor, orange). Note the complex interdomain contacts, not only between the N- and C-terminal halves of the catalytic domain, but also between the Zn^{2+}-binding domains with each other and with the DSHB module. In panel c, note the long insertion between residues Thr1463 and Gly1842 (C-terminal end of N-terminal half and N-terminal start of the C-terminal half, respectively). DSBH, domain, termed double-stranded β-helix

these N-terminal residues, suggesting that it plays an important functional role(s) (isoform B, NM 017628).

As mentioned above, somatic *TET2* mutations have been repeatedly found in myeloid neoplasia. The pathogenic attribution of *TET2* variants to hematologic neoplasms relies on the earlier description in other cases, functional studies and/or the absence of the newly described variant in the DNA from non-neoplastic samples. However, the identification of *TET2* mutations in ARCH and the presence of allele variants of *TET2* in nonhematologic neoplasms may complicate interpretation of new variants. In silico prediction of the functional consequences of *TET2* variants could be clinically useful in case

these criteria are not met. However, this requires a deep knowledge of both the wild-type protein and representative *bona fide* neoplastic mutants. Rigorous pathogenicity prediction of SNVs is limited to the structurally characterized TET2 catalytic unit, but has not been systematically attempted to date.

Here we report the identification and analysis of *TET2* SNVs identified in 189 Spanish AML patients. Spurred by the lack of structural information on the noncatalytic regions of the enzyme, we have combined bioinformatics results and available functional information into a unified model of TET2 overall architecture. We present a rigorous interpretation of over 430 missense variants that affect the TET2 catalytic domain, and provide likely explanations for ~700 additional variants found in the regulatory regions of the protein. Finally, we present a proposal of classification of missense *TET2* SNVs that takes into account these structural and functional data.

2 | METHODS

2.1 | Patients

Samples from patients with AML, diagnosed according to standard criteria (Beer et al., 2010), were analyzed for *TET2* mutations using Sanger sequencing and NGS protocols as reported elsewhere (Nomdedéu et al., 2012, 2017).

2.2 | DNA extraction and sequencing

Genomic DNA was extracted from bone marrow aspirates or from peripheral blood of healthy donors and patients using standard methods, and analyzed for mutations in *NPM1*, *FLT3*, *CEBPA*, and *MLL* genes using well-established protocols (Munoz et al., 2003, 2001, 2003; Nomdedéu et al., 2012). For mutational analysis of the *TET2* gene, DNA was amplified using primers spanning the entire coding region and PCR conditions reported elsewhere (Delhommeau et al., 2009). PCR products were purified with ExoSAP-IT (Mouradov et al., 2014), and sequenced using the BigDye™ Terminator Cycle Sequencing kit. Sequence analysis was performed on an ABI PRISM-3100 Genetic Analyzer. In 20 cases, a targeted NGS approach that allowed a mean coverage of 88% of the coding *TET2* sequence was also used, and variations were confirmed by Sanger sequencing (Nomdedéu et al., 2017).

2.3 | Bioinformatics analysis

Missense mutations were retrieved from the COSMIC database (http://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=TET2) or extracted from published work. Validated SNPs were taken from the NCBI database (http://www.ncbi.nlm.nih.gov/snp). For the identification of disordered regions, protein sequences were submitted to the DisMeta meta server (http://www-nmr.cabm.rutgers.edu/bioinformatics/disorder/; [Huang, Acton, & Montelione, 2014]) and in addition to servers PrDOS (http://prdos.hgc.jp/; [Ishida & Kinoshita, 2007]) and DICHOT (http://idp1.force.cs.is.nagoya-u.ac.jp/dichot/; [Fukuchi, Hosoda, Homma, Gojobori, & Nishikawa, 2011]). Local sequence motifs were predicted with ELM (http://elm.eu.org/; [Dinkel et al., 2016]) PROSITE (http://prosite.expasy.org/; [Sigrist et al., 2012]) and DLocalMotif (http://bioinf.scmb.uq.edu.au:8080/dlocalmotif/; [Mehdi, Sehgal, Kobe, Bailey, & Bodén, 2013]). Caspase cleavage sites were predicted with CaspDB (http://caspdb.sanfordburnham.org/; [Kumar, van Raam, Salvesen, & Cieplak, 2014]) and calpain cleavage sites with CaMPDB (http://calpain.org/predict.rb?cls=substrate; [duVerle, Ono, Sorimachi, & Mamitsuka, 2011]). Experimentally determined posttranslational modifications were taken from PhosphoSitePlus (http://www.phosphosite.org/; [Hornbeck et al., 2014]) or from cited references (Bauer et al., 2015; Nakagawa et al., 2015).

Secondary structure was predicted with JPred4 (http://www.compbio.dundee.ac.uk/jpred/; [Drozdetskiy, Cole, Procter, & Barton, 2015]) and PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/; [Buchan, Minneci, Nugent, Bryson, & Jones, 2013]) assisted also by the results of YASPIN (http://www.ibi.vu.nl/programs/yaspinwww/; [Lin, Simossis, Taylor, & Heringa, 2005]) and of β-hairpins predicted with βhairPred (http://www.imtech.res.in/raghava/bhairpred/; [Kaur & Raghava, 2003]). The impact of selected SNVs was assessed through metaservers PredictSNP (http://loschmidt.chemi.muni.cz/predictsnp/; [Bendl et al., 2014]) and Meta-SNP (http://snps.biofold.org/meta-snp; [Capriotti, Altman, & Bromberg, 2013]). Threading and modeling were performed with RaptorX (http://raptorx.uchicago.edu/; [Källberg et al., 2012]) or Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2/; [Kelley, Mezulis, Yates, Wass, & Sternberg, 2015]), and models were finally refined with ModRefiner (https://zhanglab.ccmb.med.umich.edu/ModRefiner/; [Xu & Zhang, 2011]). The impact of point variants on protein structure and function was assessed with PROVEAN (http://provean.jcvi.org/, ref. [Choi, Sims, Murphy, Miller, & Chan, 2012]), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/index.shtml, ref. [Adzhubei et al., 2010]), and CUPSAT (http://cupsat.tu-bs.de/; ref. [Parthiban, Gromiha, & Schomburg, 2006]). Structure figures were prepared with PyMol (www.pymol.org).

3 | RESULTS AND DISCUSSION

3.1 | Novel *TET2* mutations in Spanish neoplasia patients

Thirty-eight *TET2* allelic variants were identified in 189 studied AML patients, including two frequent SNPs: p.G355D (10 cases) and p.I1762V (28 cases) (Table S1). Four of the detected mutations are reported here for the first time: c.122C>T (p.P41L), c.4535C>G (p.A1512G), c.4760A>G (p.D1587G), and c.5087A>T (p.Y1696F) (Figure S1). Further, in 100 patients, we excluded the presence of additional variations

affecting residues that bind Zn^{2+} ions in TET2 catalytic unit. Finally, none of the following missense mutations or SNPs was detected in 100 healthy controls: p.Y867H, p.T1270A, p.R1359P, p.A1512G, and p.P1723S.

3.2 | Predicted domain organization of TET2 protein

A growing number of missense mutations and validated SNPs in *TET2* (summarized in Table S2 and Table S3, respectively) affect the large N-terminal, uncharacterized region of the protein (Met1-Pro1131) or the long loop inserted within the catalytic unit (Cys1464-Asn1843). Although both stretches are commonly described as intrinsically disordered regions (IDRs), it seems likely that such large, relatively well-conserved polypeptides are relevant for TET2 structure and function. This prompted us to integrate the results of several bioinformatics tools into a working model of full-length TET2 architecture (Figure 1a).

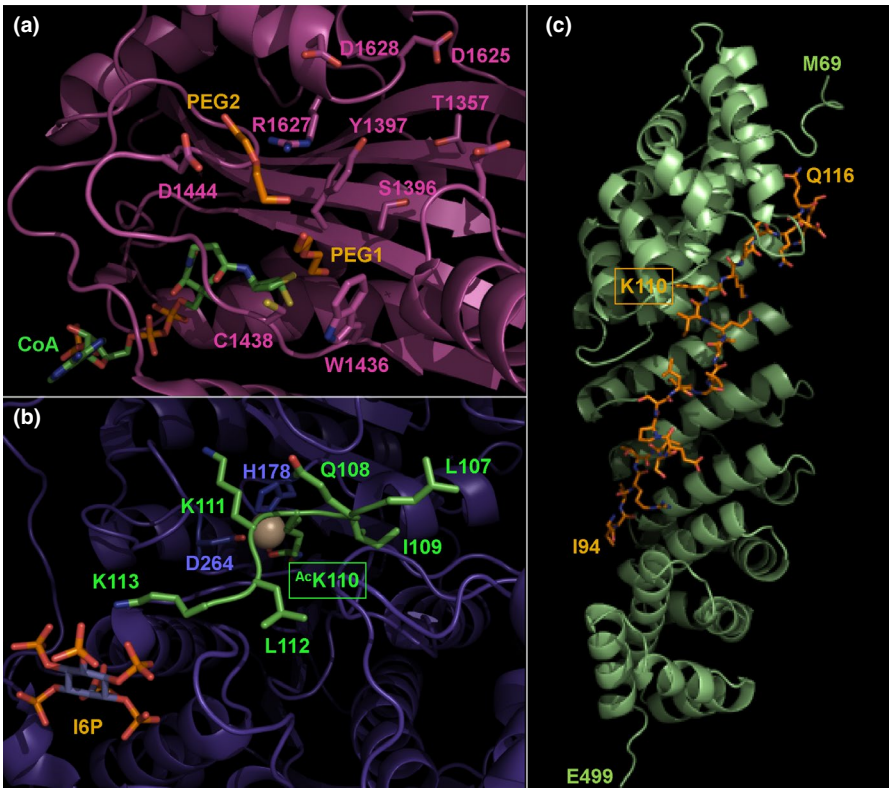


FIGURE 2 Relevance of the non-catalytic regions of TET2 protein for enzyme function. Focus is on experimentally verified or predicted roles of residues within the Gly93-Gln116 TET2 stretch. (a) Close-up of human p300 catalytic region, highlighting pockets that could accommodate the substrate lysine residue (Lys100; ref. [Zhang et al., 2017]) in TET2, as well as surrounding residues. p300 secondary structure elements are colored pink, and the CoA cofactor is shown with all its nonhydrogen atoms to the bottom-left, color-coded (carbon, green; oxygen, red; nitrogen, blue; phosphor, orange and sulfur, yellow). Well-ordered PEG molecules bound close to the active site in the reported crystal structure (PDB code 4PZR, [Maksimoska, Segura-Peña, Cole, & Marmorstein, 2014]) are also shown color-coded (carbon, orange; oxygen, red). Molecule PEG1 overlays with the incoming lysine residue to be acetylated, while PEG2 was proposed to mimic a substrate residue with a small side chain, ideally glycine, flanking the substrate lysine. In TET2, however, Lys110 is flanked by residues with bulkier side chains, and the path of the TET2 peptide in the actual p300·TET2 complex might differ from that with canonical substrates, and/or it would require some rearrangements of residues surrounding the p300 active site for binding in a productive conformation. On the other hand, the shallow negatively charged pocket on p300 termed "P2 site" (top right in this panel) might accommodate Lys113; this interaction is critical for effective substrate binding (Liu et al., 2008). (b) 3D model of TET2 peptide Leu107-Lys113 with acetylated Lys110 bound to HDAC1. The model has been generated according to the crystal structure of human HDAC1 bound to a substrate mimic based on the sequence of the histone H4 tail (5ICN, [Watson et al., 2016]). HDAC1 secondary structure elements are colored purple. Only side chains of HDAC1 residues that coordinate the catalytic Zn^{2+} ion are shown. TET2 peptide is green, with all side chains shown color-coded. Note that the side chain of Lys113 is in the position to engage in electrostatic interactions with the bound inositol hexaphosphate (I6P) molecule, a physiologically relevant activator of HDACs (Watson et al., 2016). (c) 3D model of the Gly93-Gln116 polypeptide of human TET2 bound to the nuclear importer, importin-α (IMP-α). The model has been developed based on the structure of human IMP-α bound to CPB80 (Dias, Wilson, Rojas, Ambrosio, & Cerione, 2009). The IMP-α helices are colored lime green, and the TET2 fragment in shown with all is non-hydrogen atoms, color-coded (carbon, orange; oxygen, red and nitrogen, blue). Note that the TET2 polypeptide runs in an extended, approximately antiparallel direction to the IMP-α superhelix

Most notably, the results of our analyses strongly suggest that TET2 contains several globular domains in addition to the catalytic unit, which might adopt definite 3D structures upon binding for example to cognate modules in specific coregulators. These domains are linked by IDRs that are targets of various posttranslational modifications (PTMs; Figure 1a), and which appear to be functionally relevant. For instance, residue Lys110 has been recently reported to be acetylated by p300/CBP, which in turn results in DNMT1 binding, enhanced protein stability, and higher catalytic activity in vivo (Zhang et al., 2017). (See Figure 2a for a close-up of p300 catalytic region, highlighting pockets that accommodate the substrate lysine residue, as well as surrounding residues). These findings explain the negative impact of mutant p.Lys110Arg, identified in a patient with myelodysplastic syndrome (Papaemmanuil et al., 2013). Alternatively, mutations that interfere with the deacetylation of ^{Ac}Lys100 by HDAC1/2 might result in a hyperactive TET2 enzyme. (For a model of the Leu107-Lys113 TET2 peptide bound to HDAC1, see Figure 2b). Further, a well-conserved bipartite nuclear localization signal is predicted between residues Gly93 and Asp115, and a 3D model of the putative complex between this TET2 stretch and the nuclear transport receptor, importin-α, is given in Figure 2c. A sequence alignment around this NLS is given in Figure S2a and details of the putative TET2—importin-α interaction interface are presented in Figures S2b–e. Alternatively, the Gly93-Asp115 peptide might interact with the TPR-repeat domain of another known TET2 interactor, *O*-linked GlcNAc transferase (Chen, Chen, Bian, Fujiki, & Yu, 2013; Deplus et al., 2013; Vella et al., 2013), which shares structural similarities and a common binding mode with importin-α (Jinek et al., 2004). These interactions might be regulated by phosphorylation of residue Ser99 by AMP-activated kinase, which has been recently shown to stabilize TET2 protein. This PTM thus links hyperglycemia and diabetes to cancer through modification of the 5hmC landscape (Wu et al., 2018). We also note that the highly variable Asp297-Thr395 linker is a likely major target of cysteine proteinases that regulate TET2 activity: caspases (Ko et al., 2013) and calpains (Wang & Zhang, 2014). Removal of the N-terminal region and concomitant loss of binding sites for the important interactors mentioned above might explain, at least partly, the regulatory role of these proteinases.

Furthermore, threading algorithms could identify accurate templates for some of the previously unforeseen globular domains, in particular for the Thr396-Gln574 and Gln866-Lys1044 stretches. (The low, previously unappreciated sequence similarity between these two stretches is shown in Figure S3, 3D models for these regions are given in Figure S4a,b, and model quality is illustrated in Figure S4c,d, respectively). Perhaps more relevant for TET2 enzymatic activity, the sequences immediately preceding and following the catalytic unit (residues Ala1045-Pro1131 and Lys1924-Ile2002 in human TET2, respectively) are highly conserved from reptiles to humans and have previously unrecognized counterparts in TET3 (Figure S5a,b), pointing to important conserved functions. The results of secondary structure prediction algorithms suggest that these regions might also fold into small α+β domains. Of note, we have been able to recombinantly express most of these regions in *E. coli* at high yields. The recombinant proteins are soluble in physiological conditions and could be straightforwardly purified without signs of degradation, indicating the presence of well-folded globular domains (manuscript in preparation).

3.3 | Genotype–phenotype correlation: a proposed classification of *TET2* missense mutations

In Table S2 and Table S3, we present a brief analysis of all missense mutants and validated SNPs of human *TET2* reported to date, respectively. In each case, sequence conservation, impact on experimentally confirmed or predicted PTMs and likely structural implications are considered. Considering the available results of structure and function studies (Hu et al., 2013, 2015) and bioinformatics data summarized in these tables and Figure 1a, we propose a hypothetical but testable classification of the pathogenicity of missense SNVs identified in the human *TET2* gene into three major categories (Table 1).

3.3.1 | Severe mutations

Among them, we distinguish between (a) exchanges that impair binding of the essential cofactors, 2-OG and/or Fe^{2+} (Figure 3a) and (b) mutations that would have an important impact on the global 3D structure of the TET2 catalytic unit (e.g., by disrupting major internal H-bonded networks, Figure 3b, or by introducing polar/charged residues in the densely packed hydrophobic core, Figure 3c). It is conceivable that type Ib mutations would have a more severe impact on TET2 protein function.

3.3.2 | Intermediate mutations

This category comprises (a) replacements of core residues of the catalytic unit with a limited impact on its structure and/or stability (Figure 3d), as well as (b) mutations that would disrupt the 3D structure of regulatory domains. Also included in this class are nonconservative exchanges of strictly conserved exposed residues that might weaken TET2 interactions with either (c) substrate DNA (Figure 3e), (d) with other protein domains or with other proteins (Figure 3f), or that (e) would interfere with conserved PTMs.

TABLE 1 A proposed classification of *TET2* missense mutations

Mutant class	Mutation impact	Examples
Ia	Impairment of enzyme activity, but overall protein structure essentially conserved. (Replacement of Fe^{2+}/2-OG ligands, or major steric clashes with these side chains)	p.R1261C/H/L/G, p.S1284F, p.H1881R/Q/N, p.R1896M/S/G, p.S1898F
Ib	Severely compromised 3D structure and/or stability of the catalytic unit. (Introduction of polar or charged residues in the hydrophobic core, disruption of major internal H-bonded networks, replacement of Zn^{2+} ligands)	p.R1176G/T, p.L1229R, p.C1289Y/F, p.L1322Q, p.L1329Q/R, p.R1359C/H/P, p.H1380Y/R/N/P, p.V1864E
IIa	Creation of small cavities in protein core, destabilization of major secondary structure elements and/or of partially buried H-bonded networks	p.E1207Q/D, p.L1231P, p.W1233G, p.Y1245S/C, p.E1879Q, p.I1897S, p.M1907K
IIb	Impairment of the 3D structure of noncatalytic domains	p.A215V, p.V238F, p.A241V, p.P463L, p.L1065F, p.K1094I, p.T1985I
IIc	Impaired recognition and processing of substrate DNA	p.W1291R, p.K1299E/N, p.R1302G, p.S1303R/N, p.H1904R/Q, p.K1905E
IId	Impaired docking of other TET2 domains and/or of interacting proteins (e.g., IDAX, importin-α, OGT)	p.R96C, p.E1206K, p.E12434G, p.D1314G, p.E1320A
IIe	Altered posttranslational modification of strictly conserved residues with functional implications	p.K110R, p.K1339R
IIIa	Conservative mutations of buried residues within the catalytic unit, with only minimal effect on the protein stability	p.I1175V, p.I1195V, p.V1199I, p.L1229I, p.F1287L, p.L1332F, p.L1360M, p.A1919V
IIIb	Conservative mutations of exposed, not strictly conserved residues within the catalytic unit	p.S1204C, p.K1243R, p.L1312V, p.E1357Q, p.E1405Q
IIIc	Mutations of nonconserved residues within the noncatalytic regions	p.P23S, p.K53E, p.S334F, p.S585L, p.A727T, p.S1497P

3.3.3 | Mutations with mild to negligible effects

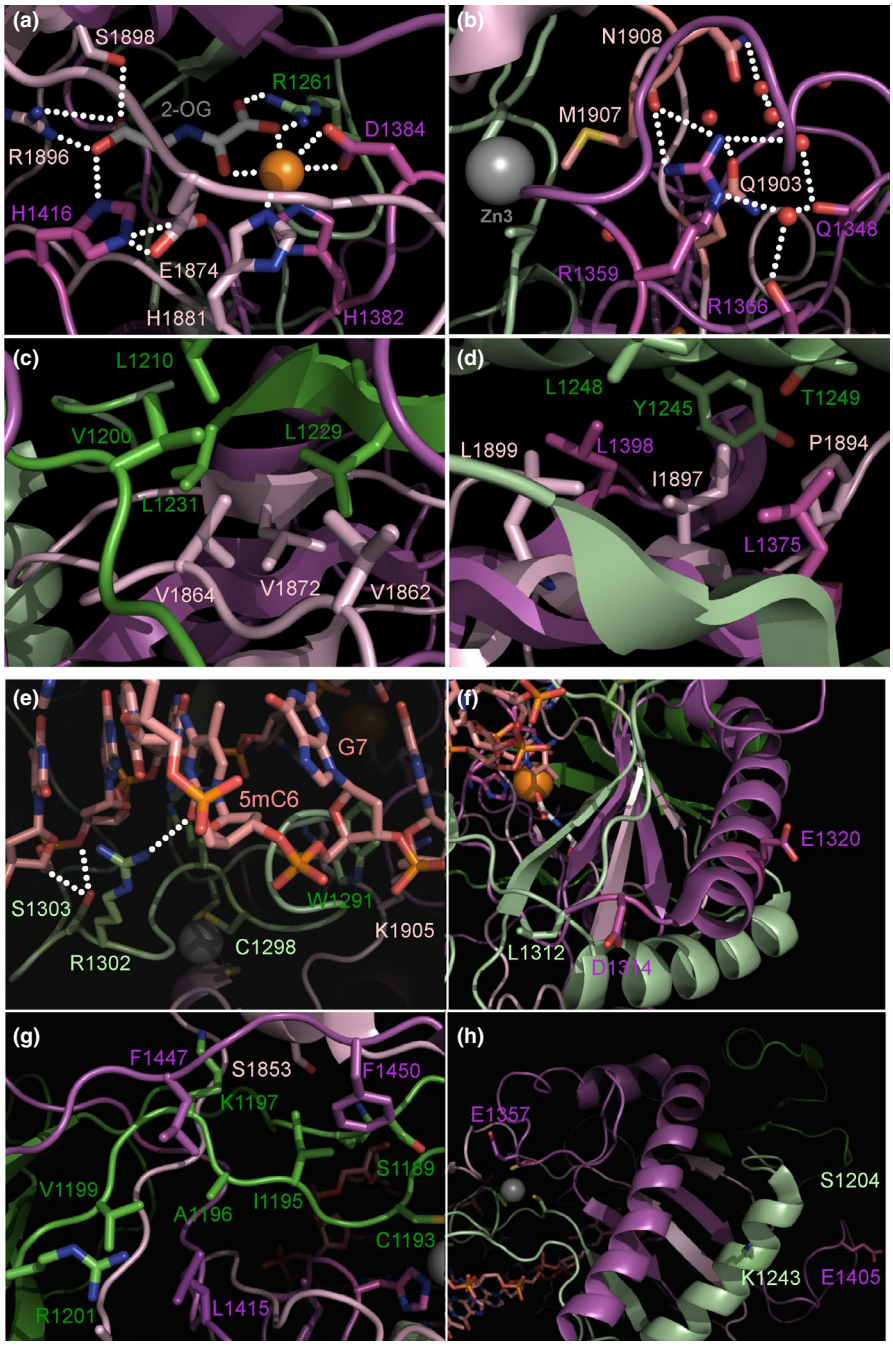
This category includes (a) conservative exchanges of buried residues within the catalytic unit, which would thus be fully tolerated (Figure 3g), (b) mutations that either directly or indirectly affect nonconserved PTMs (Figure 3h), and (c) other replacements of nonconserved residues within the noncatalytic modules of the enzyme.

We compared our structure- and function based classification with tools routinely used to predict the impact of missense variants, PROVEAN and PolyPhen, and also with an algorithm that predicts changes in structural stability upon point mutations, CUPSAT. The results of these analyses for all SNVs that map to the TET2 catalytic unit are presented in the Table S4. As can be seen, type I mutations, for instance, are usually found among the most deleterious by PROVEAN (values much lower than the threshold between "neutral" and "deleterious" replacements, −2.50). Type Ia/Ib variants are also commonly predicted as "probably damaging" by PolyPhen. There are, however, a large number of variants that are also predicted as deleterious/probably damaging, for which careful inspection of the deposited 3D structures suggests at most a limited impact on the structure and/or function of the enzyme. On the other hand, some replacements of residues that are part of internal, complex H-bonded networks (e.g., p.Arg1161Ser, p.Arg1176Gly/Thr) or that coordinate Zn^{2+} ions (e.g., Cys1221Arg, Cys1289Ser/Phe) are predicted as "stabilizing" by CUPSAT, even though loss of these interactions in the mutant proteins are quite likely to compromise overall folding and stability. In conclusion, neither simple analysis of residue conservation nor "blind" consideration of structural features allow for a detailed classification of TET2 variants, and we propose that our scheme should be used instead for predicting the likely impact of point mutations in the *TET2* gene.

Of note, the mutation pattern along the TET2 sequence is strikingly different between solid and hematologic neoplasms (Figure 1a). However, it remains to be investigated whether *TET2* variants that belong to different categories are unevenly distributed in myeloid neoplasia, solid tumors and tumor-infiltrating lymphocytes, and whether they would have different effects on atherogenesis (Bowman et al., 2018).

3.4 | Clinical significance of *TET2* mutations

Previous attempts to use the *TET2* gene as a predictor of clinical variables including disease-free survival (DFS) and overall survival (OS) have led to contradictory results. For instance, the presence of *TET2* mutated was reported as a neutral (Bejar



et al., 2011) or even favorable prognostic biomarker in MDS (Kosmider, Gelsi-Boyer, Cheok, et al., 2009), and in chronic myelomonocytic leukemia (CMML) (Grossmann et al., 2011; Kohlmann et al., 2010). Further, *TET2* mutation status did not influence OS or DFS in a cohort of cytogenetically normal AML patients younger than 60 years (Damm et al., 2014), and *TET2* gene status was not significantly correlated with DFS or OS in de novo AML (Nibourel et al., 2010). In addition, in a large cohort of younger adult patients with AML, *TET2* variants did not impact the response to induction therapy and clinical outcome (Gaidzik et al., 2012). Finally, in a more recent study of an even larger cohort of 1,540 AML patients, *TET2* mutations were much more frequently associated to "favorable risk" or "intermediate risk" than to "adverse risk" categories according to the European LeukemiaNet (ELN) classification. Of note, *TET2* mutations were quite rarely found in isolation in AML patients, indicating that they are not sufficient for overt leukemia (Papaemmanuil et al., 2016). Also along these lines, the presence of mutant *TET2* did not affect survival or leukemic transformation in patients with polycythemia vera or primary myelofibrosis (Tefferi, 2010; Tefferi et al., 2009), and *TET2*-mutated MPN patients were not cytogenetically different from their *TET2*-wild-type counterparts (Hussein et al., 2010). Finally, the results of a recent meta-analysis suggest that *TET2* mutations may not impact prognosis on OS of patients with MDS (Guo, Zhang, Zou, Fan, & Lyu, 2017).

FIGURE 3 Representative examples of mutations affecting the structure and function of human TET2 catalytic unit. (a) Close-up of TET2 catalytic site. Shown in the picture are only cofactors Fe^{2+} (orange sphere) and 2-OG (color-coded: carbon, pink; nitrogen, red; oxygen, blue), as well as the side chains of residues directly involved in their coordination (color-coded as the 2-OG cofactor, but with carbon atoms green). Note that several *TET2* missense mutations affect these residues (e.g., p.Arg1261Cys/His/Leu/Gly, p.His1881Arg/Glu/Asn; class Ia mutants). (b and c), representative examples of proposed type Ib variants. (b) Mutations likely to disrupt the overall 3D structure of the catalytic unit. Shown is a close-up around residue Arg1359. Note the intricate network of strong H-bonds centered on its side chain, which directly donates H-bonds to the carbonyl oxygen from Met1907 (C-terminal half of the DSBH domain), but is also connected through water molecules to the carbonyls of Gln1348 and Arg1366 (N-terminal half). Note also that the residue immediately preceding Arg1359 is a Zn^{2+} ligand. Therefore, mutants p.Arg1359Cys/His/Pro would most likely result in the collapse of the whole 3D structure. (c) Representative example of a mutation introducing a polar or charged residue in the densely packed hydrophobic core. Replacement of Val1864 by a glutamate, as in p.Val1864Glu, would result in major clashes of the mutant Glu1864 carboxylate with the side chains of aliphatic residues shown in the picture and/or with main chain carbonyl oxygen atoms of these and/or other residues. (d) Close-up around residue Ile1897. Only side chains of residues that make at least one vdW contact with one of the atoms of the Ile1897 side chain are shown, color-coded according to the domain they belong to. Note that replacement of the aliphatic side chain by a shorter, polar serine, although disfavored, would not be expected to cause major structural rearrangements, as the mutant Ser1897 side chain would not clash with any of the surrounding core residues (class IIa mutation). Note in addition the somehow polar environment created by the nearby Thr1249 side chain. (e) Close-up of the substrate DNA-binding site; the side chains of some of the residues mutated in cancer patients are shown with all of their nonhydrogen atoms. Replacement of single residues, such as in mutants p.Arg1302Gly or p.Lys1905Glu, is likely to affect the rate of DNA oxidation, but would not be expected to completely abolish processing of the 5mC residue (class IIc variants). (f) Close-up showing exposed, well-conserved residues that are mutated in some cancer patients (e.g., p.Leu1312Val, p.Asp1314Gly, p.Glu1320Ala). Their relative proximity to the bound DNA oligonucleotide suggests that they might form a binding site for a TET2 cofactor such as IDAX/CXXC4, which recruits TET2 to DNA (Ko et al., 2013) (type IId mutants). (g) Close-up around Ser1189-Arg1201 TET2 stretch. Only a few side chains of interacting residues are shown for simplicity. Some TET2 missense variants that affect residues within this sequence are likely to be fully tolerated without any important rearrangements of TET2 protein structure (e.g., p.Ile1195Val, p.Val1199Ile; class IIIa). (h) Close-up showing not conserved, exposed TET2 residues. Missense variants of these residues that introduce physicochemically related residues, commonly found in TET2 from other species (e.g., p.Ser1204Cys, p.Lys1243Arg, p.Glu1405Gln) are unlikely to have any impact on TET2 function (class III variants). DSBH, domain, termed double-stranded β-helix

In striking contrast, *TET2* mutations have been associated with decreased OS in de novo AML patients with intermediate-risk cytogenetics compared to *TET2*-wild-type patients (Abdel-Wahab et al., 2009), and other authors have corroborated this association between *TET2* mutations and poor prognosis in AML (Aslanyan et al., 2014; Metzeler et al., 2011; Ohgami et al., 2015). Further, *TET2* gene mutation has been reported to negatively impact patient survival in CMML (Kosmider, Gelsi-Boyer, Ciudad, et al., 2009) as well as in the most common peripheral T-cell lymphomas (Lemonnier et al., 2012). Furthermore, the nonepigenetic effects of *TET2* variants may also be of some relevance in hematologic neoplasms, as it has been recently shown for other TET family members (Khoueiry et al., 2017).

Notwithstanding other differences in the molecular pathologies of various hematological neoplasms, we consider that the apparently contradictory results listed above may be due, in part, to studying patient cohorts in which *TET2* mutations have largely different structural/functional consequences. Therefore, establishment of a scoring system and patient stratification according to our proposed classification might be helpful in future studies aimed at establishing the clinical relevance of *TET2* variants. In particular, we might expect a significant correlation between the presence of severe *TET2* mutations (classes Ia/Ib) and a poor disease outcome. Furthermore, these patients could benefit the most from treatment with hypomethylating drugs, as suggested in recent studies (Bejar et al., 2014). In this regard, a much better response to the DNA hypomethylating agent, azacitidine has been reported in AML and MDS patients carrying *TET2* mutations than in wild-type patients (82% vs. 45% response rate, including hematological improvement; [Itzykson et al., 2011]).

The understanding of the structural and functional consequences of *TET2* mutations may also be useful for the optimal design of advanced immunotherapeutic approaches. It has been recently reported the cure of a patient with chronic lymphocytic leukemia carrying the class IIa p.Glu1879Gln variant in one *TET2* allele, after disruption of the second *TET2* allele in its T cells through lentiviral integration (Fraietta et al., 2018). Altogether, consideration of the structure and function impact of *TET2* SNVs might have important implications for predicting the prognosis of leukemia patients and their stratification in future clinical trials.

ACKNOWLEDGMENTS

We thank a large number of colleagues for answering our questions regarding *TET2* mutants identified in their laboratories. This work was supported by grants SAF2014-57994-R, AGAUR 2014-SGR-01214 (P.F.-P.), FIS PI 10/1739, 13/2729, and AGAUR SGR383 (J.N.).

CONFLICT OF INTEREST

None of the authors declare any conflict of interest.

AUTHORS CONTRIBUTIONS

E.B. performed experiments and analyzed results. R.A. and P.F.-P. performed bioinformatics analyses. J.N. and P.F.-P. planned experiments, discussed the results, and wrote the paper.

ORCID

Pablo Fuentes-Prior https://orcid.org/0000-0002-6618-3204

REFERENCES

Abdel-Wahab, O., Mullally, A., Hedvat, C., Garcia-Manero, G., Patel, J., Wadleigh, M., … Levine, R. L. (2009). Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood*, *114*(1), 144–147. https://doi.org/10.1182/blood-2009-03-210039

Abegunde, S. O., Buckstein, R., Wells, R. A., & Rauh, M. J. (2018). An inflammatory environment containing TNFα favors Tet2-mutant clonal hematopoiesis. *Experimental Hematology*, *59*, 60–65. https://doi.org/10.1016/j.exphem.2017.11.002

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., … Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Arends, C. M., Galan-Sousa, J., Hoyer, K., Chan, W., Jäger, M., Yoshida, K., … Damm, F. (2018). Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. *Leukemia*, *32*(9), 1908–1919. https://doi.org/10.1038/s41375-018-0047-7

Aslanyan, M. G., Kroeze, L. I., Langemeijer, S. M. C., Koorenhof-Scheele, T. N., Massop, M., van Hoogen, P., … Jansen, J. H. (2014). Clinical and biological impact of *TET2* mutations and expression in younger adult AML patients treated within the EORTC/GIMEMA AML-12 clinical trial. *Annals of Hematology*, *93*(8), 1401–1412. https://doi.org/10.1007/s00277-014-2055-7

Bauer, C., Göbel, K., Nagaraj, N., Colantuoni, C., Wang, M., & Müller, U. (2015). Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT). *Journal of Biological Chemistry*, *290*(8), 4801–4812. https://doi.org/10.1074/jbc.M114.605881

Beer, P. A., Delhommeau, F., LeCouedic, J. P., Dawson, M. A., Chen, E., Bareford, D., … Green, A. R. (2010). Two routes to leukemic transformation after a *JAK2* mutation-positive myeloproliferative neoplasm. *Blood*, *115*(14), 2891–2900. https://doi.org/10.1182/blood-2009-08-236596

Bejar, R., Lord, A., Stevenson, K., Bar-Natan, M., Perez-Ladaga, A., Zaneveld, J., … Ebert, B. L. (2014). TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients. *Blood*, *124*(17), 2705–2712. https://doi.org/10.1182/blood-2014-06-582809

Bejar, R., Stevenson, K., Abdel-Wahab, O., Galili, N., Nilsson, B., Garcia-Manero, G., … Ebert, B. L. (2011). Clinical effect of point mutations in myelodysplastic syndromes. *New England Journal of Medicine*, *364*(26), 2496–2506. https://doi.org/10.1056/NEJMoa1013343

Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E. D., Zendulka, J., … Damborsky, J. (2014). PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. *PLOS Computational Biology*, *10*(1), e1003440. https://doi.org/10.1371/journal.pcbi.1003440

Bowman, R. L., Busque, L., & Levine, R. L. (2018). Clonal hematopoiesis and evolution to hematopoietic malignancies. *Cell Stem Cell*, *22*(2), 157–170. https://doi.org/10.1016/j.stem.2018.01.011

Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, *41*(W1), W349–W357. https://doi.org/10.1093/nar/gkt381

Buscarlet, M., Provost, S., Zada, Y. F., Bourgoin, V., Mollica, L., & Dubé, M.-P. (2018). Lineage restriction analyses in CHIP indicate myeloid bias for *TET2* and multipotent stem cell origin for *DNMT3A*. *Blood*, *132*(3), 277–280. https://doi.org/10.1182/blood-2018-01-829937

Busque, L., Patel, J. P., Figueroa, M. E., Vasanthakumar, A., Provost, S., Hamilou, Z., … Levine, R. L. (2012). Recurrent somatic *TET2* mutations in normal elderly individuals with clonal hematopoiesis. *Nature Genetics*, *44*(11), 1179–1181. https://doi.org/10.1038/ng.2413

Capriotti, E., Altman, R. B., & Bromberg, Y. (2013). Collective judgment predicts disease-associated single nucleotide variants. *BMC Genomics*, *14*(3), S2. https://doi.org/10.1186/1471-2164-14-s3-s2

Chen, Q., Chen, Y., Bian, C., Fujiki, R., & Yu, X. (2013). TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature*, *493*(7433), 561–564.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE*, *7*(10), e46688. https://doi.org/10.1371/journal.pone.0046688

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., … Bellacosa, A. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*, *146*(1), 67–79. https://doi.org/10.1016/j.cell.2011.06.020

Cull, A. H., Snetsinger, B., Buckstein, R., Wells, R. A., & Rauh, M. J. (2017). Tet2 restrains inflammatory gene expression in macrophages. *Experimental Hematology*, *55*, 56–70.e13. https://doi.org/10.1016/j.exphem.2017.08.001

Cull, A., Snetsinger, B., & Rauh, M. J. (2016). Tet2 deficiency leads to an increased inflammatory phenotype in murine macrophages. *Blood*, *128*(22), 708–708.

Damm, F., Markus, B., Thol, F., Morgan, M., Göhring, G., & Schlegelberger, B. (2014). *TET2* mutations in cytogenetically normal acute myeloid leukemia: Clinical implications and evolutionary patterns. *Genes, Chromosomes & Cancer*, *53*(10), 824–832. https://doi.org/10.1002/gcc.22191

Delhommeau, F., Dupont, S., Valle, V. D., James, C., Trannoy, S., & Massé, A. (2009). Mutation in *TET2* in myeloid cancers. *New England Journal of Medicine*, *360*(22), 2289–2301. https://doi.org/10.1056/NEJMoa0810069

Deplus, R., Delatte, B., Schwinn, M. K., Defrance, M., Méndez, J., Murphy, N., … Fuks, F. (2013). TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/

COMPASS. *The EMBO Journal*, *32*(5), 645–655. https://doi.org/10.1038/emboj.2012.357

Desai, P., Mencia-Trinchant, N., Savenkov, O., Simon, M. S., Cheang, G., Lee, S., … Hassane, D. C. (2018). Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nature Medicine*, *24*(7), 1015–1023. https://doi.org/10.1038/s41591-018-0081-z

Dias, S. M. G., Wilson, K. F., Rojas, K. S., Ambrosio, A. L. B., & Cerione, R. A. (2009). The molecular basis for the regulation of the cap-binding complex by the importins. *Nature Structural & Molecular Biology*, *16*(9), 930–937. https://doi.org/10.1038/nsmb.1649

Dinkel, H., Van Roey, K., Michael, S., Kumar, M., Uyar, B., Altenberg, B., … Gibson, T. J. (2016). ELM 2016—data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Research*, *44*(D1), D294–D300. https://doi.org/10.1093/nar/gkv1291

Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). JPred4: A protein secondary structure prediction server. *Nucleic Acids Research*, *43*(W1), W389–W394. https://doi.org/10.1093/nar/gkv332

duVerle, D. A., Ono, Y., Sorimachi, H., & Mamitsuka, H. (2011). Calpain cleavage prediction using multiple kernel learning. *PLoS ONE*, *6*(5), e19035. https://doi.org/10.1371/journal.pone.0019035

Feng, S., Jacobsen, S. E., & Reik, W. (2010). Epigenetic reprogramming in plant and animal development. *Science*, *330*(6004), 622–627. https://doi.org/10.1126/science.1190614

Fraietta, J. A., Nobles, C. L., Sammons, M. A., Lundh, S., Carty, S. A., & Reich, T. J. (2018). Disruption of *TET2* promotes the therapeutic efficacy of CD19-targeted T cells. *Nature*, *558*(7709), 307–312.

Frick, M., Chan, W., Arends, C. M., Hablesreiter, R., Halik, A., Heuser, M., … Damm, F. (2019). Role of donor clonal hematopoiesis in allogeneic hematopoietic stem-cell transplantation. *Journal of Clinical Oncology*, *37*(5), 375–385. https://doi.org/10.1200/JCO.2018.79.2184

Fukuchi, S., Hosoda, K., Homma, K., Gojobori, T., & Nishikawa, K. (2011). Binary classification of protein molecules into intrinsically disordered and ordered segments. *BMC Structural Biology*, *11*(1), 29. https://doi.org/10.1186/1472-6807-11-29

Fuster, J. J., MacLauchlan, S., Zuriaga, M. A., Polackal, M. N., Ostriker, A. C., Chakraborty, R., … Walsh, K. (2017). Clonal hematopoiesis associated with Tet2 deficiency accelerates atherosclerosis development in mice. *Science*, *355*(6327), 842–847. https://doi.org/10.1126/science.aag1381

Gaidzik, V. I., Paschka, P., Späth, D., Habdank, M., Köhne, C.-H., & Germing, U. (2012). *TET2* mutations in acute myeloid leukemia (AML): Results from a comprehensive genetic and clinical analysis of the AML Study Group. *Journal of Clinical Oncology*, *30*(12), 1350–1357. https://doi.org/10.1200/jco.2011.39.2886

Gontier, G., Iyer, M., Shea, J. M., Bieri, G., Wheatley, E. G., Ramalho-Santos, M., & Villeda, S. A. (2018). Tet2 rescues age-related regenerative decline and enhances cognitive function in the adult mouse brain. *Cell Reports*, *22*(8), 1974–1981. https://doi.org/10.1016/j.celrep.2018.02.001

Grossmann, V., Kohlmann, A., Eder, C., Haferlach, C., Kern, W., & Cross, N. C. P. (2011). Molecular profiling of chronic myelomonocytic leukemia reveals diverse mutations in >80% of patients with *TET2* and *EZH2* being of high prognostic relevance. *Leukemia*, *25*, 877–879. https://doi.org/10.1038/leu.2011.10

Guo, Z., Zhang, S.-K., Zou, Z., Fan, R.-H., & Lyu, X.-D. (2017). Prognostic significance of *TET2* mutations in myelodysplastic syndromes: A meta-analysis. *Leukemia Research*, *58*, 102–107. https://doi.org/10.1016/j.leukres.2017.03.013

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., … Xu, G.-L. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, *333*(6047), 1303–1307. https://doi.org/10.1126/science.1210944

Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Skrzypek, E. (2014). PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Research*, *43*(D1), D512–D520. https://doi.org/10.1093/nar/gku1267

Hu, L., Li, Z. E., Cheng, J., Rao, Q., Gong, W., Liu, M., … Xu, Y. (2013). Crystal structure of TET2-DNA complex: Insight into TET-mediated 5mC oxidation. *Cell*, *155*(7), 1545–1555. https://doi.org/10.1016/j.cell.2013.11.020

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., & Hou, H. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature*, *527*(7576), 118–122.

Huang, Y. J., Acton, T. B., & Montelione, G. T. (2014). DisMeta: A meta server for construct design and optimization. *Methods in Molecular Biology*, *1091*, 3–16. https://doi.org/10.1007/978-1-62703-691-7_1

Hussein, K., Abdel-Wahab, O., Lasho, T. L., Van Dyke, D. L., Levine, R. L., Hanson, C. A., … Tefferi, A. (2010). Cytogenetic correlates of TET2 mutations in 199 patients with myeloproliferative neoplasms. *American Journal of Hematology*, *85*(1), 81–83. https://doi.org/10.1002/ajh.21562

Ishida, T., & Kinoshita, K. (2007). PrDOS: Prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Research*, *35*(suppl 2), W460–W464. https://doi.org/10.1093/nar/gkm363

Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C., & Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, *466*(7310), 1129–1133. https://doi.org/10.1038/nature09303

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., … Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, *333*(6047), 1300–1303. https://doi.org/10.1126/science.1210597

Itzykson, R., Kosmider, O., Cluzeau, T., Mansat-De Mas, V., Dreyfus, F., & Beyne-Rauzy, O. (2011). Impact of *TET2* mutations on response to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. *Leukemia*, *25*(7), 1147–1152.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P. V., Mar, B. G., … Ebert, B. L. (2014). Age-related clonal hematopoiesis associated with adverse outcomes. *New England Journal of Medicine*, *371*(26), 2488–2498. https://doi.org/10.1056/NEJMoa1408617

Jaiswal, S., Natarajan, P., Silver, A. J., Gibson, C. J., Bick, A. G., Shvartz, E., … Ebert, B. L. (2017). Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. *New England Journal of Medicine*, *377*(2), 111–121. https://doi.org/10.1056/NEJMoa1701719

Jinek, M., Rehwinkel, J., Lazarus, B. D., Izaurralde, E., Hanover, J. A., & Conti, E. (2004). The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin α. *Nature Structural & Molecular Biology*, *11*(10), 1001–1007. https://doi.org/10.1038/nsmb833

Kafer, G. R., Li, X., Horii, T., Suetake, I., Tajima, S., Hatada, I., & Carlton, P. M. (2016). 5-Hydroxymethylcytosine marks sites of DNA damage and promotes genome stability. *Cell Reports*, *14*(6), 1283–1292. https://doi.org/10.1016/j.celrep.2016.01.035

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., & Xu, J. (2012). Template-based protein structure modeling using the

RaptorX web server. *Nature Protocols*, *7*(8), 1511–1522. https://doi.org/10.1038/nprot.2012.085

Kaur, H., & Raghava, G. P. S. (2003). Prediction of β-turns in proteins from multiple alignment using neural network. *Protein Science*, *12*(3), 627–634.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, *10*, 845–858. https://doi.org/10.1038/nprot.2015.053

Khoueiry, R., Sohni, A., Thienpont, B., Luo, X., Velde, J. V., Bartoccetti, M., … Koh, K. P. (2017). Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nature Genetics*, *49*, 1061–1072. https://doi.org/10.1038/ng.3868

Ko, M., An, J., Bandukwala, H. S., Chavez, L., Aijo, T., & Pastor, W. A. (2013). Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature*, *497*(7447), 122–126.

Ko, M., An, J., Pastor, W. A., Koralov, S. B., Rajewsky, K., & Rao, A. (2015). TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunological Reviews*, *263*(1), 6–21. https://doi.org/10.1111/imr.12239

Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., & Bandukwala, H. S. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant *TET2*. *Nature*, *468*(7325), 839–843.

Kohlmann, A., Grossmann, V., Klein, H.-U., Schindela, S., Weiss, T., & Kazak, B. (2010). Next-generation sequencing technology reveals a characteristic pattern of molecular mutations in 72.8% of chronic myelomonocytic leukemia by detecting frequent alterations in *TET2*, *CBL*, *RAS*, and *RUNX1*. *Journal of Clinical Oncology*, *28*(24), 3858–3865. https://doi.org/10.1200/jco.2009.27.1361

Kosmider, O., Gelsi-Boyer, V., Cheok, M., Grabar, S., Della-Valle, V., Picard, F., … Fontenay, M. (2009). *TET2* mutation is an independent favorable prognostic factor in myelodysplastic syndromes (MDSs). *Blood*, *114*(15), 3285–3291. https://doi.org/10.1182/blood-2009-04-215814

Kosmider, O., Gelsi-Boyer, V., Ciudad, M., Racoeur, C., Jooste, V., Vey, N., … Solary, E. (2009). *TET2* gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia. *Haematologica*, *94*(12), 1676–1681. https://doi.org/10.3324/haematol.2009.011205

Kumar, S., van Raam, B. J., Salvesen, G. S., & Cieplak, P. (2014). Caspase cleavage sites in the human proteome: CaspDB, a database of predicted substrates. *PLoS ONE*, *9*(10), e110539. https://doi.org/10.1371/journal.pone.0110539

Langemeijer, S. M. C., Kuiper, R. P., Berends, M., Knops, R., Aslanyan, M. G., Massop, M., … Jansen, J. H. (2009). Acquired mutations in *TET2* are common in myelodysplastic syndromes. *Nature Genetics*, *41*(7), 838–842. https://doi.org/10.1038/ng.391

Lemonnier, F., Couronne, L., Parrens, M., Jais, J.-P., Travert, M., Lamant, L., … Gaulard, P. (2012). Recurrent *TET2* mutations in peripheral T-cell lymphomas correlate with T_{FH}-like features and adverse clinical parameters. *Blood*, *120*(7), 1466–1469. https://doi.org/10.1182/blood-2012-02-408542

Li, Z., Cai, X., Cai, C.-L., Wang, J., Zhang, W., Petersen, B. E., … Xu, M. (2011). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*, *118*(17), 4509–4518. https://doi.org/10.1182/blood-2010-12-325241

Lin, K., Simossis, V. A., Taylor, W. R., & Heringa, J. (2005). A simple and fast secondary structure prediction method using hidden neural networks. *Bioinformatics*, *21*(2), 152–159. https://doi.org/10.1093/bioinformatics/bth487

Liu, X., Wang, L., Zhao, K., Thompson, P. R., Hwang, Y., Marmorstein, R., & Cole, P. A. (2008). The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature*, *451*, 846–850. https://doi.org/10.1038/nature06546

Maksimoska, J., Segura-Peña, D., Cole, P. A., & Marmorstein, R. (2014). Structure of the p300 histone acetyltransferase bound to acetyl-coenzyme A and its analogues. *Biochemistry*, *53*(21), 3415–3422. https://doi.org/10.1021/bi500380f

Mehdi, A. M., Sehgal, M. S. B., Kobe, B., Bailey, T. L., & Bodén, M. (2013). DLocalMotif: A discriminative approach for discovering local motifs in protein sequences. *Bioinformatics*, *29*(1), 39–46. https://doi.org/10.1093/bioinformatics/bts654

Metzeler, K. H., Maharry, K., Radmacher, M. D., Mrózek, K., Margeson, D., & Becker, H. (2011). *TET2* mutations improve the new European LeukemiaNet risk classification of acute myeloid leukemia: A cancer and leukemia group B study. *Journal of Clinical Oncology*, *29*(10), 1373–1381. https://doi.org/10.1200/jco.2010.32.7742

Montagner, S., Leoni, C., Emming, S., Della Chiara, G., Balestrieri, C., Barozzi, I., … Monticelli, S. (2016). TET2 regulates mast cell differentiation and proliferation through catalytic and non-catalytic activities. *Cell Reports*, *15*(7), 1566–1579. https://doi.org/10.1016/j.celrep.2016.04.044

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., … Levine, R. L. (2011). *Tet2* loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell*, *20*(1), 11–24. https://doi.org/10.1016/j.ccr.2011.06.001

Mouradov, D., Sloggett, C., Jorissen, R. N., Love, C. G., Li, S., Burgess, A. W., … Sieber, O. M. (2014). Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer. *Cancer Research*, *74*(12), 3238–3247. https://doi.org/10.1158/0008-5472.can-14-0013

Munoz, L., Aventin, A., Villamor, N., Junca, J., Acebedo, G., & Domingo, A. (2003). Immunophenotypic findings in acute myeloid leukemia with FLT3 internal tandem duplication. *Haematologica*, *88*(6), 637–645.

Munoz, L., Nomdedeu, J. F., Lopez, O., Carnicer, M. J., Bellido, M., & Aventin, A. (2001). Interleukin-3 receptor alpha chain (CD123) is widely expressed in hematologic malignancies. *Haematologica*, *86*(12), 1261–1269.

Muñoz, L., Nomdedéu, J. F., Villamor, N., Guardia, R., Colomer, D., Ribera, J. M., … Sierra, J. (2003). Acute myeloid leukemia with MLL rearrangements: Clinicobiological features, prognostic impact and value of flow cytometry in the detection of residual leukemic cells. *Leukemia*, *17*(1), 76–82. https://doi.org/10.1038/sj.leu.2402708

Nakagawa, T., Lv, L., Nakagawa, M., Yu, Y., Yu, C., D'Alessio, A. C., … Xiong, Y. (2015). $CRL4^{VprBP}$ E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases. *Molecular Cell*, *57*(2), 247–260. https://doi.org/10.1016/j.molcel.2014.12.002

Nibourel, O., Kosmider, O., Cheok, M., Boissel, N., Renneville, A., Philippe, N., … Preudhomme, C. (2010). Incidence and prognostic value of *TET2* alterations in de novo acute myeloid leukemia achieving complete remission. *Blood*, *116*(7), 1132–1135. https://doi.org/10.1182/blood-2009-07-234484

Nomdedéu, J., Hoyos, M., Carricondo, M., Esteve, J., Bussaglia, E., Estivill, C., … Sierra, J. (2012). Adverse impact of IDH1 and IDH2 mutations in primary AML: Experience of the Spanish

CETLAM group. *Leukemia Research*, *36*(8), 990–997. https://doi.org/10.1016/j.leukres.2012.03.019

Nomdedéu, J. F., Puigdecanet, E., Bussaglia, E., Hernández, J. J., Carricondo, M., & Estivill, C. (2017). Feasibility of the AML profiler (Skyline™ Array) for patient risk stratification in a multicentre trial: A preliminary comparison with the conventional approach. *Hematological Oncology*, *35*(4), 778–788. https://doi.org/10.1002/hon.2304

Ohgami, R. S., Ma, L., Merker, J. D., Gotlib, J. R., Schrijver, I., Zehnder, J. L., & Arber, D. A. (2015). Next-generation sequencing of acute myeloid leukemia identifies the significance of TP53, U2AF1, ASXL1, and TET2 mutations. *Modern Pathology*, *28*(5), 706–714. https://doi.org/10.1038/modpathol.2014.160

Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., … Campbell, P. J. (2016). Genomic classification and prognosis in acute myeloid leukemia. *New England Journal of Medicine*, *374*(23), 2209–2221. https://doi.org/10.1056/NEJMoa1516192

Papaemmanuil, E., Gerstung, M., Malcovati, L., Tauro, S., Gundem, G., Van Loo, P., … Campbell, P. J. (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*, *122*(22), 3616–3627. https://doi.org/10.1182/blood-2013-08-518886

Parthiban, V., Gromiha, M. M., & Schomburg, D. (2006). CUPSAT: Prediction of protein stability upon point mutations. *Nucleic Acids Research*, *34*, W239–W242. https://doi.org/10.1093/nar/gkl190

Pastor, W. A., Aravind, L., & Rao, A. (2013). TETonic shift: Biological roles of TET proteins in DNA demethylation and transcription. *Nature Reviews Molecular Cell Biology*, *14*(6), 341–356. https://doi.org/10.1038/nrm3589

Quivoron, C., Couronné, L., Della Valle, V., Lopez, C. K., Plo, I., Wagner-Ballon, O., … Bernard, O. A. (2011). TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell*, *20*(1), 25–38. https://doi.org/10.1016/j.ccr.2011.06.003

Rampal, R., Alkalin, A., Madzo, J., Vasanthakumar, A., Pronier, E., Patel, J., … Levine, R. L. (2014). DNA hydroxymethylation profiling reveals that *WT1* mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Reports*, *9*(5), 1841–1855. https://doi.org/10.1016/j.celrep.2014.11.004

Rasmussen, K. D., & Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes & Development*, *30*(7), 733–750. https://doi.org/10.1101/gad.276568.115

Sano, S., Ohshima, K., Wang, Y., MacLauchlan, S., Katanasaka, Y., & Sano, M. (2018). Tet2-mediated clonal hematopoiesis accelerates heart failure through a mechanism involving the IL-1β/NLRP3 inflammasome. *Journal of the American College of Cardiology*, *71*(8), 875–886. https://doi.org/10.1161/circresaha.118.313225

Shen, Q., Zhang, Q., Shi, Y., Shi, Q., Jiang, Y., Gu, Y., … Cao, X. (2018). Tet2 promotes pathogen infection-induced myelopoiesis through mRNA oxidation. *Nature*, *554*, 123–127. https://doi.org/10.1038/nature25434

Sigrist, C. J. A., de Castro, E., Cerutti, L., Cuche, B. A., Hulo, N., Bridge, A., … Xenarios, I. (2012). New and continuing developments at PROSITE. *Nucleic Acids Research*, *41*(D1), D344–D347. https://doi.org/10.1093/nar/gks1067

Smith, Z. D., & Meissner, A. (2013). DNA methylation: Roles in mammalian development. *Nature Reviews Genetics*, *14*(3), 204–220. https://doi.org/10.1038/nrg3354

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., … Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, *324*(5929), 930–935. https://doi.org/10.1126/science.1170116

Tefferi, A. (2010). Novel mutations and their functional and clinical relevance in myeloproliferative neoplasms: *JAK2*, *MPL*, *TET2*, *ASXL1*, *CBL*, *IDH* and *IKZF1*. *Leukemia*, *24*, 1128–1138. https://doi.org/10.1038/leu.2010.69

Tefferi, A., Levine, R. L., Lim, K.-H., Abdel-Wahab, O., Lasho, T. L., Patel, J., … Gilliland, D. G. (2009). Frequent *TET2* mutations in systemic mastocytosis: Clinical, *KITD816V* and *FIP1L1-PDGFRA* correlates. *Leukemia*, *23*(5), 900–904. https://doi.org/10.1038/leu.2009.37

Vella, P., Scelfo, A., Jammula, S. G., Chiacchiera, F., Williams, K., Cuomo, A., … Pasini, D. (2013). Tet proteins connect the O-Linked N-acetylglucosamine transferase OGT to chromatin in embryonic stem cells. *Molecular Cell*, *49*(4), 645–656. https://doi.org/10.1016/j.molcel.2012.12.019

Wang, Y., & Zhang, Y. (2014). Regulation of TET protein stability by calpains. *Cell Reports*, *6*(2), 278–284. https://doi.org/10.1016/j.celrep.2013.12.031

Watson, P. J., Millard, C. J., Riley, A. M., Robertson, N. S., Wright, L. C., Godage, H. Y., … Schwabe, J. W. R. (2016). Insights into the activation mechanism of class I HDAC complexes by inositol phosphates. *Nature Communications*, *7*, 11262. https://doi.org/10.1038/ncomms11262

Wu, D., Hu, D., Chen, H., Shi, G., Fetahu, I. S., & Wu, F. (2018). Glucose-regulated phosphorylation of TET2 by AMPK reveals a pathway linking diabetes to cancer. *Nature*, *559*(7715), 637–641.

Xu, D., & Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical Journal*, *101*(10), 2525–2534. https://doi.org/10.1016/j.bpj.2011.10.024

Zang, S., Li, J., Yang, H., Zeng, H., Han, W., Zhang, J., … Huang, Y. (2017). Mutations in 5-methylcytosine oxidase TET2 and RhoA cooperatively disrupt T cell homeostasis. *The Journal of Clinical Investigation*, *127*(8), 2998–3012. https://doi.org/10.1172/JCI92026

Zhang, Q., Zhao, K., Shen, Q., Han, Y., Gu, Y., Li, X., … Cao, X. (2015). Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature*, *525*, 389–393. https://doi.org/10.1038/nature15252

Zhang, Y. W., Wang, Z., Xie, W., Cai, Y. I., Xia, L., Easwaran, H., … Baylin, S. B. (2017). Acetylation enhances TET2 function in protecting against abnormal DNA methylation during oxidative stress. *Molecular Cell*, *65*(2), 323–335. https://doi.org/10.1016/j.molcel.2016.12.013

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Bussaglia E, Antón R, Nomdedéu JF, Fuentes-Prior P. *TET2* missense variants in human neoplasia. A proposal of structural and functional classification. *Mol Genet Genomic Med*. 2019;7:e772. https://doi.org/10.1002/mgg3.772