maximal exercise testing in children and young adults with sickle cell anaemia. *Br J Haematol*. 2015;171(5):854-861.

8. Joosten YA, Israel TL, Williams NA, et al. Community engagement studios: a structured approach to obtaining meaningful input from stakeholders to inform research. *Acad Med*. 2015;90(12):1646-1650.

9. Nixon PA, Orenstein DM, Kelsey SF, Doershuk CF. The prognostic value of exercise testing in patients with cystic fibrosis. *N. Engl. J. Med*. 1992;327(25):1785-1788.

10. Friedman C, Rubin J, Brown J, et al. Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *J Am Med Inform Assoc*. 2015;22(1):43-50.

DOI 10.1182/blood.2019003651

MYELOID NEOPLASIA

Comment on Borrow et al, pages 2281, and 2291

The curious incident of TdT-mediated mutations in AML

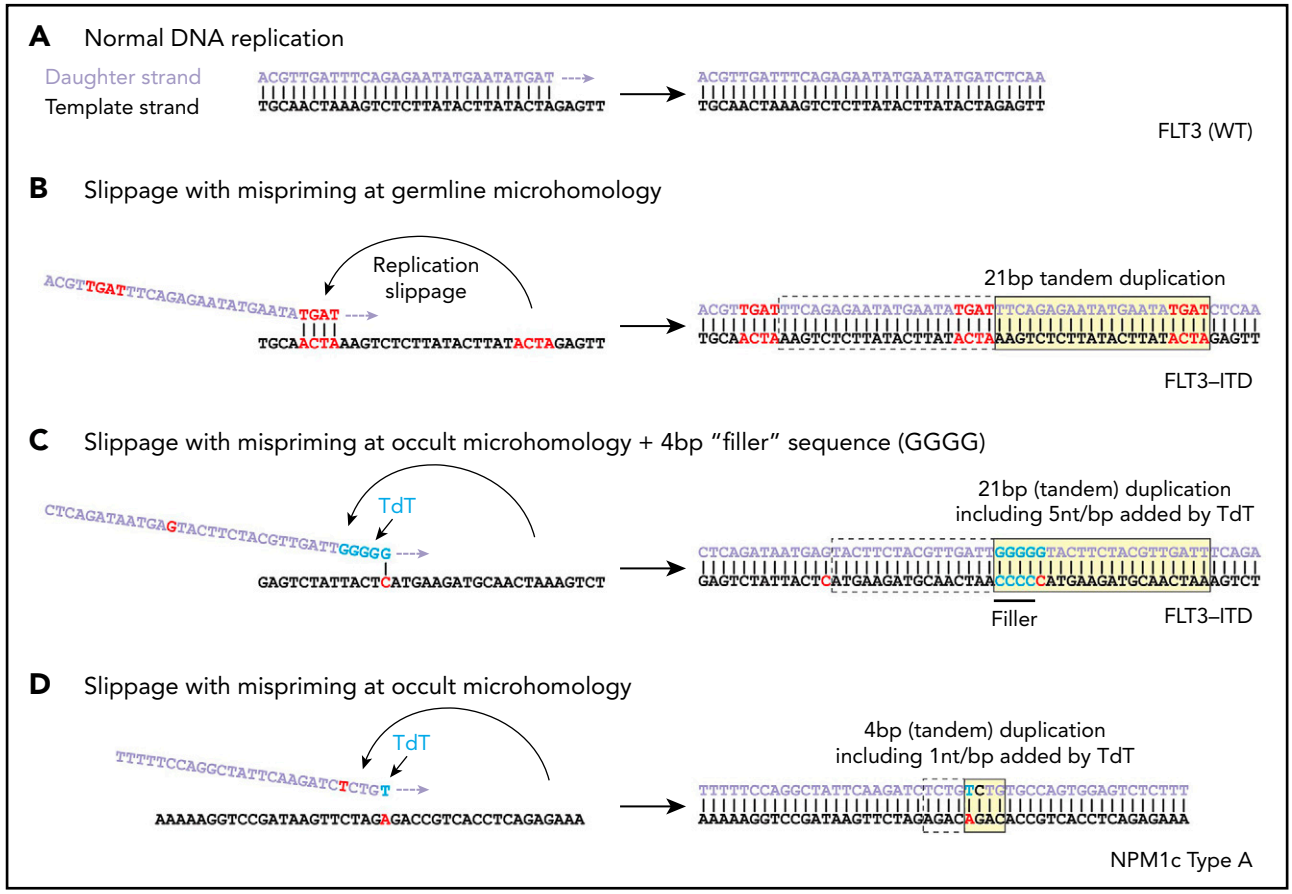
George S. Vassiliou | University of Cambridge

In this issue of *Blood*, Borrow et al reanalyze published DNA sequencing data and report, in back-to-back papers,[1,2] that the lymphoid enzyme terminal deoxynucleotide transferase (TdT) appears to be involved in the causation of the 2 most common types of gene mutation in human acute myeloid leukemia (AML),[3] namely internal tandem duplications of *FLT3* (*FLT3-ITD*) and short insertions/duplications within the final exon of *NPM1* (*NPM1c*). The normal function of TdT is to increase the diversity of the immunoglobulin and T-cell receptor (TCR) loci, by adding nontemplated nucleotides to their variable regions. Other enzymes involved in generating this diversity, such as RAG and AID, can act illegitimately to cause oncogenic mutations in acute lymphoblastic leukemia (ALL).[4] Here, the work by Borrow et al proposes TdT as a major mutagen in AML, an unexpected finding that forces a reevaluation of our understanding of myeloid leukemogenesis.

Both *FLT3-ITD* and *NPM1c* mutations are thought to arise during cell division as a result of *DNA replication slippage* (also known as slipped strand mispairing). This is a molecular error in which the DNA polymerase and the attached nascent daughter DNA strand temporarily dissociate from the template strand and move to an upstream location with *microhomology* to the 3′ end of the daughter strand. The DNA polymerase then "resumes" replication from this position and, in so doing, reincorporates the same nucleotides into the elongating daughter strand. Although this error can be detected by excision repair proteins, it is not always correctly repaired, leading to a permanent duplication mutation. Replication slippage is more likely to occur at repetitive regions, where repeats can form hairpin loops and a "slipped" polymerase can readily find a complementary sequence on the template strand at which to relocate. This error has been put forward as an explanation for the presence of repetitive DNA sequences in genomes and for the intergenerational expansion of trinucleotide repeats in disorders such as Huntington disease.[5]

The presence of repetitive DNA sequences at exons 14 and 15 of *FLT3* has been widely thought to promote the occurrence of *FLT3-ITD* mutations. However, by studying the sequences of 300 different *FLT3-ITDs* from published datasets, Borrow et al observed that the mutations were not always perfect duplications, but frequently contained additional *filler* nucleotides, something reported before, but not closely scrutinized. In addition, they noted that, in approximately one-third of cases, the *ITD* could not have been generated by simple replication slippage, because there were no homologous nucleotides at the "slipped" position from which DNA polymerase would "resume" replication, a phenomenon they refer to as *missing microhomology* (see figure). The authors then set out to investigate the origin of the *filler* nucleotides and also to find explanation for the *missing microhomology*. First, they hypothesized that *microhomology* of at least 1 basepair was generated by addition of nucleotides to the end of the daughter strand and inferred what the sequence of these nucleotides had to be. Then they examined the sequence composition of (i) *filler* nucleotides and (ii) nucleotides added to generate *microhomology* and found that both closely matched the pattern of nucleotides added by TdT (in terms of GC content, length, and incidence of homodimers). In their second paper,[2] the authors turned their attention to *NPM1* exon 12 mutations and studied 2430 individual cases. They again observed that the most common form of *NPM1* mutation (type A = TCTG duplication) represents a case of replication slippage with *missing microhomology* (see figure). By examining other less common types of *NPM1* mutation, the authors discovered that almost all of these were consistent with instances of *missing microhomology*. As with *FLT3-ITD*, the patterns of added nucleotides closely matched those expected by the action of TdT.

Although the evidence for the action of TdT is expectedly indirect, TdT is the most template-independent of the Pol X family of DNA polymerases[6] and is expressed by a significant proportion of AMLs, whereas other AMLs have rearranged immunoglobulin or TCR loci (indicating TdT activity in an ancestral cell).[7] Furthermore, the authors allude to unpublished evidence that one of the few other types of recurrent *ITDs* in cancer, namely *BCOR-ITDs* in solid cancers (where TdT is not active), does not harbor additional nucleotides compatible with TdT activity.[2] Also, in mice, where TdT activity is lower than in humans,[8] *Flt3-ITDs* do not occur spontaneously in *Npm1*-mutant animals.[9] In this light, the 2 papers make a strong case that, through its role in causing *NPM1c* and *FLT3-ITD* mutations, TdT is involved in the development of almost half of all cases of human AML, a remarkable finding. AML genomes harbor only small numbers of somatic mutations, mostly nucleotide



Role of TdT in the causation of *FLT3-ITD* and *NPM1c* mutations in AML. (A) In normal DNA replication, the daughter strand (gray) is a perfect complement of the template strand (black). (B) FLT3-ITDs can arise as a result of replication slippage during which the DNA polymerase resumes replication from an upstream position microhomologous with the end of the elongating daughter strand (red). (C) In approximately one-third of FLT3-ITDs, there is no visible microhomology. Instead, TdT is predicted to add nucleotides (blue) to the end of the daughter strand in order to provide the occult (or "missing") microhomology (of at least 1 basepair) between the daughter and template strands. In this specific instance, a run of 5 Gs was added, only the final one of which was used for microhomology. The first 4 Gs then appear as a "filler" sequence in the final mutant DNA and also serve to keep the FLT3 mRNA reading frame open (the total number of additional nucleotides [21] is divisible by 3). (D) Similarly, a single T nucleotide is predicted to be added to the daughter strand during the formation of the common NPM1 mutation, a 4-bp tandem duplication. Borrow et al provide evidence that the extra nucleotides facilitating the formation of approximately one-third of FLT3-ITDs and almost all NPM1c mutations[2] follow a pattern that strongly suggests illegitimate TdT activity. Panels B to D are derived from real examples of mutations depicted by Borrow et al. bp, basepair; nt, nucleotide. Professional illustration by Patrick Lane, ScEYEnce Studios.

substitutions, that arise stochastically through rare mistakes in genome maintenance and accumulate slowly with age.[10] *NPM1c* and *FLT3-ITD* mutations never fitted this pattern, and the work by Borrow et al provides a plausible explanation for their causation.

Incriminating TdT in the causation of the 2 most common mutations in AML may have important mechanistic and by extension clinical implications. For example, mutations in either *NPM1* or *FLT3* have not been described in clonal hematopoiesis (CH) and are usually acquired in cells harboring preexisting mutations in genes such as *DNMT3A*, leading to AML. It has been hypothesized that such CH mutations "synergize" with NPM1 to drive leukemogenesis. Could they instead, or in addition, be involved in epigenetic activation of TdT activity in stem/progenitor cells? Alternatively, could TdT activation be transient and related to environmental exposures or interim illnesses, as has been proposed for pediatric ALL?[4,11] Furthermore, could the shared causation by TdT underlie the preferred cooccurrence of *NPM1c* with *FLT3-ITD* (compared to other synergistic mutations)?

Finally, beyond its scientific impact, the authors' investigation stands out as an elegant demonstration of how a subtle observation can give unexpected insights with far-reaching implications. In this manner, their scrutiny of the previously inconspicuous nucleotides and their conclusions about their significance are reminiscent of the fabled adventure alluded to in this piece's title: the fact that the dog did not bark in the nighttime was easy to ignore, but to Sherlock Holmes it was the key evidence leading him to the identity of the intruder.[12]

Conflict-of-interest disclosure: G.S.V. is a consultant to Kymab and Oxstem. ■

REFERENCES

1. Borrow J, Dyer SA, Akiki S, Griffiths MJ. Terminal deoxynucleotidyl transferase promotes acute myeloid leukemia by priming *FLT3*-ITD replication slippage. *Blood*. 2019; 2281-2290.

2. Borrow J, Dyer SA, Akiki S, Griffiths MJ. Molecular roulette: nucleophosmin mutations in AML are orchestrated through N-nucleotide addition by TdT. *Blood*. 2019;2291-2303.

3. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016; 374(23):2209-2221.

4. Greaves M. A causal mechanism for childhood acute lymphoblastic leukaemia [published correction appears in Nat Rev Cancer. 2018; 18(8):426]. *Nat Rev Cancer*. 2018;18(8): 471-484.

5. McMurray CT. Mechanisms of trinucleotide repeat instability during human development [published correction appears in Nat Rev Genet. 2010;11(12):886]. *Nat Rev Genet*. 2010;11(11):786-799.

6. Nick McElhinny SA, Havener JM, Garcia-Diaz M, et al. A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol Cell*. 2005;19(3):357-366.

7. Drexler HG, Sperling C, Ludwig WD. Terminal deoxynucleotidyl transferase (TdT) expression in acute myeloid leukemia. *Leukemia*. 1993; 7(8):1142-1150.

8. Bentolila LA, Fanton d'Andon M, Nguyen QT, Martinez O, Rougeon F, Doyen N. The two isoforms of mouse terminal deoxynucleotidyl transferase differ in both the ability to add N regions and subcellular localization. *EMBO J*. 1995;14(17):4221-4229.

9. Dovey OM, Cooper JL, Mupo A, et al. Molecular synergy underlies the co-occurrence patterns and phenotype of *NPM1*-mutant acute myeloid leukemia. *Blood*. 2017; 130(17):1911-1922.

10. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 2012;150(2):264-278.

11. Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet*. 2014;46(2):116-125.

12. Doyle AC. *Adventure 1: Silver Blaze. The Memoirs of Sherlock Holmes*. London, United Kingdom: George Newes; 1894.

DOI 10.1182/blood.2019003619

THROMBOSIS AND HEMOSTASIS

Comment on Bellio et al, page 2304

SHP2: when cardiology meets hematology

Cécile Oury and Patrizio Lancellotti | University of Liège

In this issue of *Blood*, Bellio et al identified a causal link between a platelet signaling defect and bleeding in patients with Noonan syndrome (NS) who have activating mutations of Src homology 2 (SH2) domain-containing phosphatase 2 (SHP2).[1]

NS is a developmental disorder affecting 1 in every 1000 to 2500 people.[2] This clinically and genetically heterogeneous syndrome is commonly characterized by defects in the RAS/MAPK cell signaling pathway. Although multiple gene mutations can cause NS, about half the cases are caused by mutations in the *PTPN11* gene. Affected individuals most often present with distinctive facial features, short stature, chest deformity, and a wide spectrum of congenital heart diseases.

Disordered bleeding has been reported for 30% to 65% of individuals with NS and has consistently been associated with mutations in the *PTPN11* gene. Bleeding is often mild with symptoms such as excessive bruising, epistaxis, or menorrhagia, but it can become severe after surgical procedures. Defects of 1 or several coagulation factors or thrombocytopenia have been described but without any consistent association with bleeding episodes.

It is noteworthy that the study by Bellio et al is the first to describe a genetic cause of thrombopathy resulting from a platelet signaling defect linked to disordered bleeding. So far, only a few inherited platelet disorders have been related to a platelet signaling defect such as autosomal recessive gene mutations in *PLA2G4A* (Ghosal syndrome), *TBXSA1*, and *GNAS*.[3] However, whether these specific platelet defects are responsible for bleeding remains unknown.

The study by Bellio et al provides the first mechanistic explanation for bleeding in patients with *PTPN11* mutations. The authors used genetically engineered mouse models and blood samples from NS patients to demonstrate that mutations of this gene can cause a platelet signaling defect. The protein encoded by *PTPN11* is the ubiquitously expressed classical nonreceptor protein tyrosine phosphatase (PTP) SHP2.[4] This protein has 2 tandem N-terminal SH2 domains, a catalytic PTP domain, a C-terminal tail with 2 tyrosine phosphorylation sites (Y542 and Y580), and a proline-rich region. At resting state, SH2 domain blocks access to SHP2 substrates by binding to its active site pocket. Upon ligand-receptor interaction, the N-SH2 domain preferentially binds to tyrosine-phosphorylated proteins such as receptor tyrosine kinase or scaffold proteins to open up the phosphatase active site, resulting in catalytic activation of SHP2. Mutations causing NS result in SHP2 gain-of-function by destabilizing the catalytically inactive conformation of the protein, which prolongs ligand-dependent signal transduction. In addition to NS, loss-of-function mutations of SHP2 can cause a rare NS-related syndrome, NS with multiple lentigines (NSML). Intriguingly, patients with NSML can also present with disordered bleeding.

Previous studies in mice with megakaryocyte- and platelet-specific deficiencies indicate that SHP2 contributes to thrombus stabilization under high shear stress without dramatically interfering with hemostasis.[5] Indeed, platelet aggregation induced by collagen, adenosine diphosphate, or thrombin occurred normally in the absence of SHP2. In sharp contrast to these findings, Bellio et al showed a proximal collagen receptor signaling defect in mice heterozygous for the gain-of-function D61G mutation in SHP2. Ex vivo and in vivo thrombus formation were also impaired, and tail bleeding time was prolonged in most animals, proving the defective primary hemostasis. Hence, these data nicely recapitulated the bleeding diathesis of NS patients with gain-of-function mutations of SHP2. These results were validated with blood samples from NS patients who bear different gain-of-function mutations of SHP2. Despite variable Tosetto bleeding scores ranging from normal to high, platelets from all patients displayed defective aggregation in response to low concentrations of collagen as well as impaired thrombus formation on a collagen-coated surface under flow. Conversely, platelets from mice expressing a loss-of-function SHP2 mutant showed enhanced response to low concentrations of collagen, which further confirmed the role of SHP2 in regulating collagen receptor signaling. Thrombus formation was also increased under high shear stress. Ferric chloride–induced occlusive thrombosis and bleeding time were normal, in agreement with the existence of a mild shear-dependent prothrombotic tendency associated with loss-of-function mutations of SHP2. Accordingly, blood samples from 3 NSML patients depicted enhanced thrombus growth on a collagen