AWARD NUMBER: W81XWH-15-1-0153

TITLE: Alternative RNA Splicing of CSF3R in Promoting Myelodysplastic Syndromes

PRINCIPAL INVESTIGATOR: Seth J. Corey, MD

CONTRACTING ORGANIZATION: Virginia Commonwealth University
Richmond, VA 23284

REPORT DATE: January 2019

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| **January** 2019 | Annual | 13 Dec 2017 - 12 Dec 2018 |

**4. TITLE AND SUBTITLE**
**Alternative RNA Splicing of CSF3R in Promoting Myelodysplastic Syndromes**

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
W81XWH-15-1-0153

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Dr. Seth Corey

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Virginia Commonwealth
University
800 East Leigh Street
Box 980568
Richmond, VA 23298-0568

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Medical Research
And Material Command
Fort Detrick, Maryland 21702

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

More effective therapies for myelomdysplastic syndromes (MDS) can be developed if we know more about how the disease develops. One of the most exciting advances has been the identification of mutations in genes encoding splicing factors. These occur in 50 - 70% of all adult patients with MDS. These proteins acts as a machine to process instructions (messenger RNA) that lead to the production of a specific protein. We have identified that the receptor for the most important growth factor for the production of granulocytes (the white blood cells most affected in MDS) is subject to splicing. These splicing changes result in a defective receptor, which fails to instruct blood cells to mature. We have developed a test to identify which specific splicing factor is involved in processing the messenger RNA for this receptor. We are identifying that specific splicing factor and whether there is any required post-translational modification of the splicing factor. This knowledge will inform us on how MDS begins and how to interrupt its development and progression to leukemia. Also, we have found that this defective receptor results in too much growth and too little differentiation. We have identified that splicing factors such as U2AF1 and post-translational modification involving phosphorylation contribute to processing of the message for the granulocyte colony stimulating factor receptor. SRSF2 may also play a role in regulating CSF3R. We are developing a mouse model that will allow us to describe in greater, more accurate detail the molecular changes and cell behaviors due to that defective receptor. Our work could allow us to screen for drugs that correct the MDS condition by correcting the faulty splicing and may advance the use of the receptor as a clinical laboratory tool.

**15. SUBJECT TERMS**
Splicing factor, myelodysplastic syndromes, CSF3R

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | USAMRMC |
| Unclassified | Unclassified | Unclassified | | 46 | 19b. TELEPHONE NUMBER *(include area code)* |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

TABLE OF CONTENTS

**Page**

## 1. INTRODUCTION

A major distinguishing feature of myelodysplastic syndromes (MDS), the most common form of acquired bone marrow failure, is the presence of recurrent mutations in one of the genes encoding a component of the splicing machinery. These mutations are found in 50-85% of individuals with MDS. However, little is known of their impact on normal and abnormal hematopoiesis. Our lab studies the signal transduction of Granulocyte Colony Stimulating Factor Receptor (GCSFR, the gene is *CSF3R*). The alternative splicing of *CSF3R*, which is associated with MDS, provides a robust model to reveal the mechanisms by which aberrant splicing promotes myelodysplasia and determine cell fate.

## 2. KEYWORDS

Myelodysplastic Syndromes, Splicing Mutations, CSF3R

## 3. ACCOMPLISHMENTS

- **What were the major goals of the project?**

  *Specific Aim 1.* Determine the splicing mechanism involved in processing the CSF3R gene into transcripts encoding a full-length GCSFR and a truncation, differentiation-impaired GCSFR. We will construct a minigene reporter cassette and test the predicted mechanisms. We will determine which signaling pathways promote intron retention and permit expression of full-length GCSFR so to target this step pharmacologically.

  *Specific Aim 2.* Fully characterize the aberrant proximal phosphoprotein and distal gene regulatory networks and correlate with an in vivo model of a truncated GCSFR. We will compare the signaling and gene expression profiles in murine and human CD34+ hematopoietic stem cells and correlate phenotypically with a retroviral transduction/transplantation model by expressing alternative splice form in the context of Csf3r-/- mice.

- **What was accomplished under these goals?**

  - Major activities: Interrogation of splicing factors known to be recurrently mutated in myelodysplastic syndromes and development of mice, which exclusively express the alternatively spliced CSF3R.
  - Specific objectives*: See "Detailed Accomplishments" below for further information.*
  - Significant results: Identification of CSF3R being affected by SRSF2, U2AF1, S34F, and by post-translational modification.

  **Detailed accomplishments**

  MDS constitute the most common group of bone marrow failure syndromes that are phenotypically characterized by peripheral cytopenias and hypercellular bone marrow. Intensive efforts to understand the molecular basis of myelodysplastic syndromes (MDS) have led to the identification of recurrent somatic mutations in RNA splicing factors 1-6. MDS constitute the most common group of bone marrow failure syndromes that are phenotypically characterized by peripheral cytopenias and hypercellular bone marrow.

Intensive efforts to understand the molecular basis of MDS have led to the identification of recurrent somatic mutations in RNA splicing factors. MDS are molecularly characterized by high frequencies (>50%) of somatic mutations in splicing factors *SF3B1, U2AF1, SRSF2, and ZRSR2*[7-1Ω].

To determine the effect of splicing factor mutations on CSF3R splicing resulting in either the dominant Class I isoform or alternatively spliced Class IV isoform, we performed minigene assays in the cell lines under conditions indicated in Figure 1. Figure 1 is a typical experiment in which the cells are transduced with either minigene alone or minigene plus wild type or myelodysplastic syndrome associated mutant form of a splicing factor. The cells also undergo treatment with either 0.1% DMSO (control) or 10 nM phorbol myristate (in 0.1 % DMSO) or 1 mM sodium orthovanadate (Na3VO4, a pan phosphatase inhibitor). PCR is performed using primers that would identify both Class I or Class IV. Figure 1 shows a representative gel image to indicate Class I and Class IV specific minigene transcripts identified by PCR. Band intensity of Class I and Class IV is determined using Image J and a ratio of Class IV:Class I is determined. Increase in this ratio is indicative of increase splicing resulting in Class IV vs Class I. In Figure 2, we observed that treatment with Na3VO4, results in statistically significant reduction of Class IV levels in reference to Class I. The data suggests that phosphatases may play a role in splicing towards Class IV. In Figure 3, we demonstrate that expression of mutant SRSF2 P95H shows a consistently higher levels of Class IV levels in reference to Class I, even upon Na3VO4 treatment, and is statistically significant. The data demonstrate mutation in SRSF2 results in an increase in Class IV levels and negates the effect of Na3VO4. A similar trend of increased Class IV:Class I is also observed in the K562 cell line (Figure 4). Another splicing factor implicated in MDS and CSF3R splicing is U2AF1. Minigene assay demonstrated a similar increase in Class IV:Class I levels (Figure 5). However, unlike the SRSF2 mutant, the U2AF1 mutant is negatively affected by cellular phosphorylation levels. Both PMA and Na3VO4 treatment show a trend towards reduced Class IV:Class I. The data suggest that cellular phosphorylation may impede the ability of U2AF1 mutant in promoting Class IV splicing. The data also suggest that increased phosphorylation may inhibit mutant U2AF1 activity. We also studied the effect of LuC7L2, a splicing factor present on chromosome 7. We and others have demonstrated monosomy 7 strongly correlates with increase Class IV levels. To achieve effects under conditions of monosomy 7, we performed knockdown of LuC7L2 in HEK293 cells. We were able to achieve maximum knockdown of 80%. We used knockdown of 50% and 80% for our

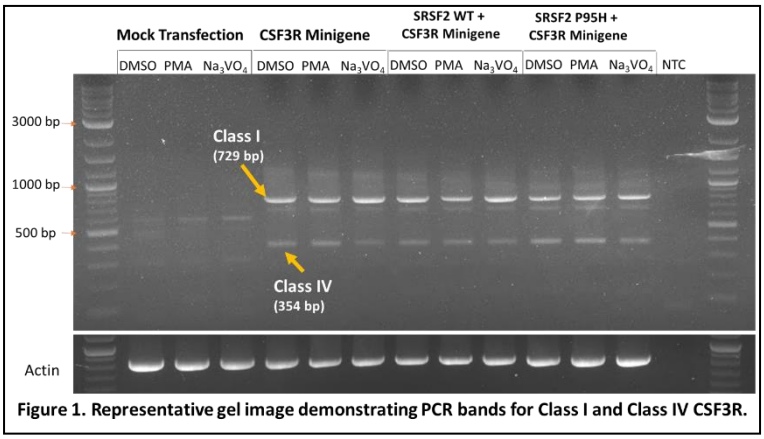


**Figure 1. Representative gel image demonstrating PCR bands for Class I and Class IV CSF3R.**

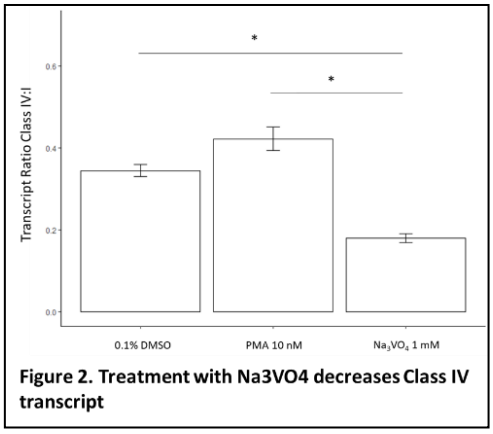


**Figure 2. Treatment with Na3VO4 decreases Class IV transcript**

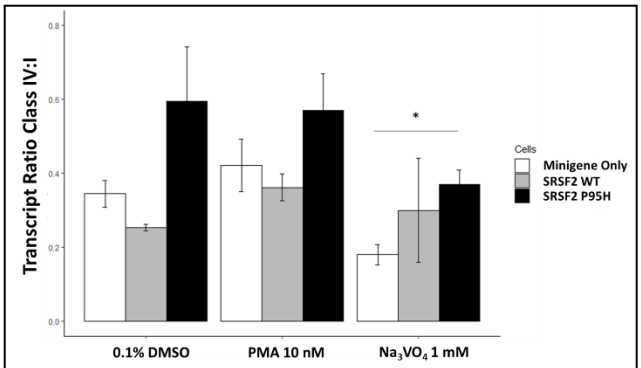



Figure 3. Effect of SRSF2 mutations on the ratio of Class IV:Class I ratio in HEK293 cells.

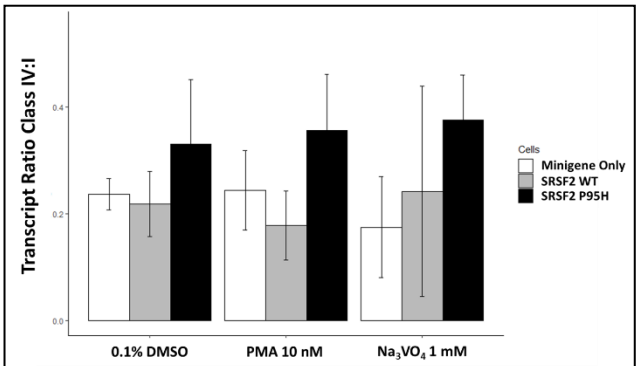



Figure 4. Effect of SRSF2 mutations on the ratio of Class IV:Class I ratio in K562 cells.

minigene assay (Figure 6). The data demonstrate an increasing trend in Class IV:Class I with increasing knockdown of the LUC7L2.

Collectively the data demonstrate that MDS-related mutations in splicing factors drive splicing towards Class IV verses Class I. We are currently carrying out studies to perform real time PCR to obtain more accurate quantitative measure using the minigene assay. In addition, studies are underway to use hematopoietic cell lines such as HL60, NB4, and U937 that would be used to determine effect of splicing factor mutations using both the minigene assay and also the effect on endogenous CSF3R splicing. We have also prepared the lentivirus to carry out the-add back mouse studies proposed in Aim 2.

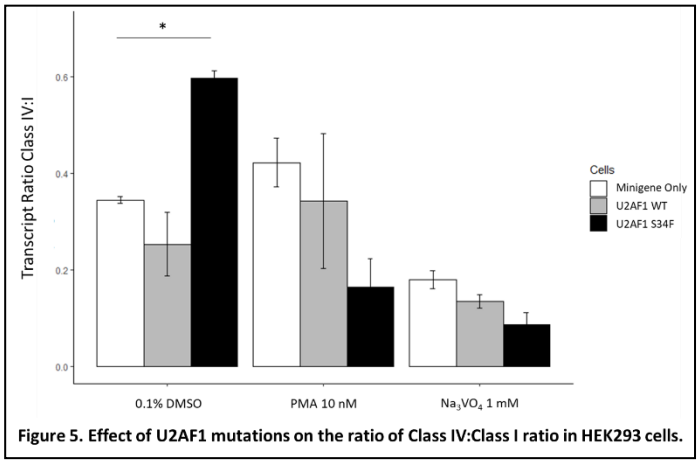



Figure 5. Effect of U2AF1 mutations on the ratio of Class IV:Class I ratio in HEK293 cells.

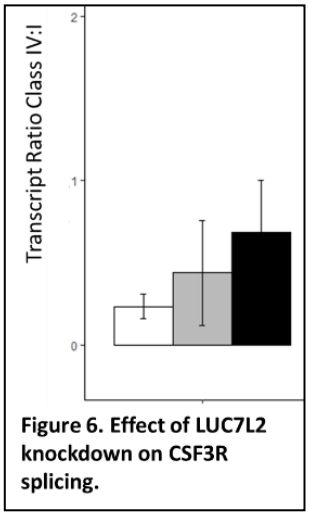



**Figure 6. Effect of LUC7L2 knockdown on CSF3R splicing.**

- **What opportunities for training and professional development has the project provided?**

A doctoral student, Borwyn “Ann” Wang, is performing this study under the PI’s supervision. She successfully defended her qualilfying thesis on October 3, 2018. Also involved in this work was a fellow, Francis Austin, who is now an Assistant Professor of Pediatrics (Hematology/Oncology) at Virginia Commonwealth University. Dr Austin worked on splicing of U2AF1 in CSF3R and during the course of the year project, she identified, described, and characterize splicing of TP53 in a child with acute myeloid leukemia and three other primary cancers.

- **How were the results disseminated to communities of interest?**

  *Publications:* Austin F, Oyarbide U, Massey G, Grimes M, Corey SJ. Synonymous mutation in TP53 results in a cryptic splice site affecting its DNA-binding site in an adolescent with two primary sarcomas. Pediatr Blood Cancer 2017; 64(11). doi: 10.1002/pbc.26584.

  Oyarbide U, Corey SJ. SRP54 and a need for a new neutropenia nosology **Blood**, 132:1220-1222, 2018.

  Oyarbide U, Topczewski J, Corey SJ. Peering Through Zebrafish to Understand Inherited Bone Marrow Failure Syndromes. **Haematologica** 104(1):13-24, 2019.

  Wojdyla T, Mehta H, Glaubach T, Bertolusso R, Iwanaszko M, Braun R, Corey SJ*, Kimmel M* corresponding authors. Mutation, Drift and Selection in Single-Driver Hematologic Malignancy: Example of Secondary Myelodysplastic Syndrome Following Treatment of Inherited Neutropenia**. PLoS Comp Bio**, Jan 7;15(1):e1006664, 2019.

  *Presentations:* Ann Wang poster presentation at the Massey Cancer Retreat June 2018; seminars by Seth Corey at Cleveland Clinic Cancer Biology/Translational Hematology and Oncology Research, University of Virginia Cancer Center, University of Pittsburgh School of Pharmacy, Institute of Molecular Genetics and Biochemistry, Seoul National University, University of North Texas, and 6$^{th}$ International AA&MDS International Symposium.

- **What do you plan to do during the next reporting period to accomplish the goals?**

  Make the mouse expressing only class IV CSF3R and characterize the effects on hematopoiesis and identify the effects of LUC7L2 haploinsufficiency in monosomy 7 with CSF3R splicing. Write a manuscript.

## 4. IMPACT

- **What was the impact on the development of the principal discipline(s) of the project?**

  We have tentatively identified U2AF1 and tyrosine phosphorylation as a splicing factor and post-translational modification that regulate the processing of the CSF3R transcript. We have also identified possible role for SRSF2. This may identify a pathway for therapeutic targeting in MDS.

- **What was the impact on other disciplines?**

  Nothing to report.

- **What was the impact on technology transfer?**

None

  Nothing to report.

**What was the impact on society beyond science and technology?**

  Nothing to report.

## 5. CHANGES/PROBLEMS

- **Changes in approach and reasons for change**

  No changes to report.

- **Actual or anticipated problems or delays and actions or plans to resolve them**

  Dr. Corey and his lab has been recruited to Cleveland Clinic and its world renown biological and clinical center for acquired bone marrow failure and myelodysplastic syndromes, headed by Drs. Jaroslaw Maciejewski and Mikkael Sekeres. The lab relocated November 1, 2018, and it is operative as of December 2018. Furthermore, we have already benefited from establishing a new collaboration with Dr. Richard Padgett, a well-established expert in RNA splicing. There is a slight delay in getting the mice bred for experiments (the animal protocol has been submitted).

- **Changes that had a significant impact on expenditures**

  We have carry over due to a delay in getting the grant transferred from Northwestern to Virginia Commonwealth University and due to salary coverage of a doctoral student who was covered initially by other departmental resources.

- **Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents**

  None.

- **Significant changes in use or care of human subjects**

  None.

- **Significant changes in use or care of vertebrate animals**

  None.

- **Significant changes in use of biohazards and/or select agents**

  None.

## 6. PRODUCTS

- **Publications, conference papers, and presentations**

  - **Journal publications**
    1. Austin F, Oyarbide U, Massey G, Grimes M, Corey SJ. Synonymous mutation in TP53 results in a cryptic splice site affecting its DNA-binding site in an adolescent with two primary sarcomas. **Pediatr Blood Cancer** 2017; 64(11). doi: 10.1002/pbc.26584.
    2. Oyarbide U, Corey SJ. SRP54 and a need for a new neutropenia nosology **Blood**, 132:1220-1222, 2018.

3. Oyarbide U, Topczewski J, Corey SJ. Peering Through Zebrafish to Understand Inherited Bone Marrow Failure Syndromes. **Haematologica** 104(1):13-24, 2019.
4. Wojdyla T, Mehta H, Glaubach T, Bertolusso R, Iwanaszko M, Braun R, Corey SJ*, Kimmel M* corresponding authors. Mutation, Drift and Selection in Single-Driver Hematologic Malignancy: Example of Secondary Myelodysplastic Syndrome Following Treatment of Inherited Neutropenia**. PLoS Comp Bio**, Jan 7;15(1):e1006664, 2019.

- **Books or other non-periodical, one-time publications**

  Nothing to report.

- **Other publications, conference papers and presentations**
  - 2017, Jan 6: Columbia Univ Dept of Pediatrics
  - 2017, Mar 14: University of Pittsburgh School of Pharmacy
  - 2017, Oct 13: University of Virginia Cancer Center Grand Rounds
  - 2017, Dec 10,11: American Society Hematology Education Session, Chair and Speaker on “Neutropenias”*
  - 2018, March 14: Inst of Molecular Genetics & Biochemistry, Seoul National University
  - 2018, March 22: AA&MDS International Symposium Debate: Should you alter management/treatment based on the presence of acquired mutations in someone with inherited BMF or leukemia
  - 2018, Oct 3: The Role of Alternatively-Spliced CSF3R in Promoting Myelodysplastic Syndromes (successful qualifying presentation of doctoral student Borwyn Wang)
  - 2018, Oct 5: University of North Texas Dept of Biology

- **Website(s) or other Internet site(s)**

  None.

- **Technologies or techniques**

  Minigene for CSF3R.

- **Inventions, patent applications, and/or licenses**

  None.

- **Other Products**

  None.

## 7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

- **What individuals have worked on the project?**

| Name: | Seth Corey, MD |
|---|---|
| Project Role: | Principal Investigator |
| Researcher Identifier: | sjcorey (NIH commons) |
| Nearest person month worked: | 1 |
| Contribution to Project: | Seth Corey supervises the entire project |
| Funding Support: | NIH, LLS, DOD, ALSF, Hyundai |

| Name: | Rishi Mehta, PhD |
|---|---|
| Project Role: | co-investigator |
| Researcher Identifier: | HRISHIKESH.MEHTA (NIH commons) |
| Nearest person month worked: | 2 |
| Contribution to Project: | Rishi Mehta makes lentiviral particles, breeds CSF3R-/- mice, and underwent security clearance so to use irradiator to do the lentiviral transduction/transplantation |
| Funding Support: | DOD, NIH, start-up funds |

| Name: | Conghui Cheng, MD, PhD |
|---|---|
| Project Role: | Co-PI at Baylor College of Medicine |
| Researcher Identifier: | |
| Nearest person month worked: | <1 |
| Contribution to Project: | Provides expertise related to alternative splicing mechanisms |
| Funding Support: | |

| Name: | Ann Wang, BS MS |
|---|---|
| Project Role: | doctoral student |
| Researcher Identifier: | Wangba |
| Nearest person month worked: | 12 |
| Contribution to Project: | Ann Wang performs transfections of cell lines, PCR and qPCR for *CSF3R* in minigene assay |
| Funding Support: | Graduate School, DOD |

- **Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**

  No

- **What other organizations were involved as partners?**

  Dr. Chonghui Cheng at the Baylor College of Medicine transitioned from serving on the project as a Co-PI to serving as a consultant. Dr. Cheng continues to interact with Dr. Corey by providing expertise in alternative splicing mechanisms and by helping to guide the characterization of the CSF3R splicing mechanism. Since relocating to Cleveland Clinic, Dr.

Corey has engaged RNA splicing expert Dr. Richard Padgett as a consultant. They participate in joint lab meeting. Dr. Corey and Ms Wang are members of the Cancer Biology department at Cleveland Clinic, and they interact closely with Dr Jaroslaw Maciejewski, world renown expert in MDS and splicing.

## 8. SPECIAL REPORTING REQUIREMENTS

- **COLLABORATIVE AWARDS**

  Not applicable.

- **QUAD CHARTS**

  Not applicable.

## 9. APPENDICES


1. Austin F, Oyarbide U, Massey G, Grimes M, Corey SJ. Synonymous mutation in TP53 results in a cryptic splice site affecting its DNA-binding site in an adolescent with two primary sarcomas. **Pediatr Blood Cancer** 2017; 64(11). doi: 10.1002/pbc.26584.
2. Oyarbide U, Corey SJ. SRP54 and a need for a new neutropenia nosology **Blood**, 132:1220-1222, 2018.
3. Oyarbide U, Topczewski J, Corey SJ. Peering Through Zebrafish to Understand Inherited Bone Marrow Failure Syndromes. **Haematologica** 104(1):13-24, 2019.
4. Wojdyla T, Mehta H, Glaubach T, Bertolusso R, Iwanaszko M, Braun R, Corey SJ*, Kimmel M* corresponding authors. Mutation, Drift and Selection in Single-Driver Hematologic Malignancy: Example of Secondary Myelodysplastic Syndrome Following Treatment of Inherited Neutropenia**. PLoS Comp Bio**, Jan 7;15(1):e1006664, 2019.

Received: 6 February 2017 | Revised: 7 March 2017 | Accepted: 8 March 2017

DOI: 10.1002/pbc.26584

BRIEF REPORT

WILEY Pediatric Blood & Cancer

SIOP INTERNATIONAL SOCIETY OF PAEDIATRIC ONCOLOGY

aspho The American Society of Pediatric Hematology/Oncology

# Synonymous mutation in *TP53* results in a cryptic splice site affecting its DNA-binding site in an adolescent with two primary sarcomas

**Frances Austin**[1] | **Usua Oyarbide**[1] | **Gita Massey**[1] | **Margaret Grimes**[2] | **Seth J. Corey**[1,3]

[1]Division of Pediatric Hematology, Oncology, and Stem Cell Transplantation, Children's Hospital of Richmond at Virginia Commonwealth University, Richmond, Virginia

[2]Department of Pathology, Virginia Commonwealth University, Richmond, Virginia

[3]Cancer Molecular Genetics Program, Massey Cancer Center at Virginia Commonwealth University, Richmond, Virginia

**Correspondence**
Frances Austin, Division of Pediatric Hematology, Oncology, and Stem Cell Transplantation, Children's Hospital of Richmond at VCU, 401 College Street, PO Box 980037, Richmond, VA.
Email: frances.austin@vcuhealth.org

Grant sponsor: Department of Defense Bone Marrow Failure Idea Grant; Grant sponsor: Children's Hospital Foundation; Grant sponsor: Connors' Heroes Foundation.

## Abstract

Pathologic variants in *TP53* are known risk factors for the development of cancer. We report a 17-year-old male who presented with two primary sarcomas. Germline sequencing revealed a novel *TP53* c.672 G>A mutation. Sequencing revealed wild-type *TP53* in the parents, and there was no history of cancer in first-degree relatives. This de novo synonymous germline mutation results in a 5′ cryptic splice site that is bound by U1, resulting in a shift of the splice site by 5 base pairs. The frame shift results in a truncated protein at residue 246, which disrupts the DNA-binding domain of p53.

**KEYWORDS**

Li-Fraumeni syndrome, sarcoma, splicing

## 1 | INTRODUCTION

*TP53* is the most frequently mutated gene in human cancer.[1] Encoded by 13 exons, this gene has many genetic polymorphisms leading to more than 100 haplotypes and through alternative splicing 12 nonpathogenic protein isoforms.[2,3] To date, none of the nonpathogenic isoforms have splicing alterations in the exons 6 and 7. Most *TP53* mutations associated with cancer result from a missense substitution (73%) leading to an amino acid substitution in the DNA-binding domain. Other changes include frameshift insertions and deletions (9%), nonsense mutations (8%), silent mutations (4%), and splice site mutations (2%; IARC TP53 database).[4] These mutations typically result in a full-length protein with either a gain of function or loss of wild-type tumor suppressive function. Exons 4 through 8 of *TP53* comprise the DNA-binding domain with exons 5 through 7 constituting a hotspot for cancer-associated mutations.

The classic definition of Li-Fraumeni syndrome includes one proband with sarcoma less than 45 years of age, a first-degree relative with cancer prior to the age of 45, and a first- or second-degree relative with any cancer before 45 years of age or a sarcoma at any age.[5] Birch and Eeles modified the criteria, leading to the nomenclature of "Li-Fraumeni-like syndrome," which still requires two relatives with cancer. *TP53* mutations (7–24%) are de novo mutations.[6,7] Carriers with *TP53* pathogenic variants have a 40% probability of cancer by 20 years and greater than 90% by 70 years. These mutations produce an 83-fold risk of developing multiple tumors.[8]

We describe an adolescent who presented with two different sarcomas. Sequencing revealed a germline synonymous mutation in the DNA-binding domain of *TP53*, which produces a 5′ cryptic splice site that results in a frame shift in the amino acids translated resulting in a premature termination at amino acid 246, disrupting the DNA-binding domain of p53.

## 2 | CASE REPORT

A 17-year-old fraternal twin male, the product of in vitro fertilization (IVF), presented with hematuria. He was initially diagnosed with pleomorphic sarcoma, most likely a leiomyosarcoma, of the pelvis (Fig. 1A

Abbreviations: bm, bone marrow; bp, base pair

 © 2017 Wiley Periodicals, Inc. 

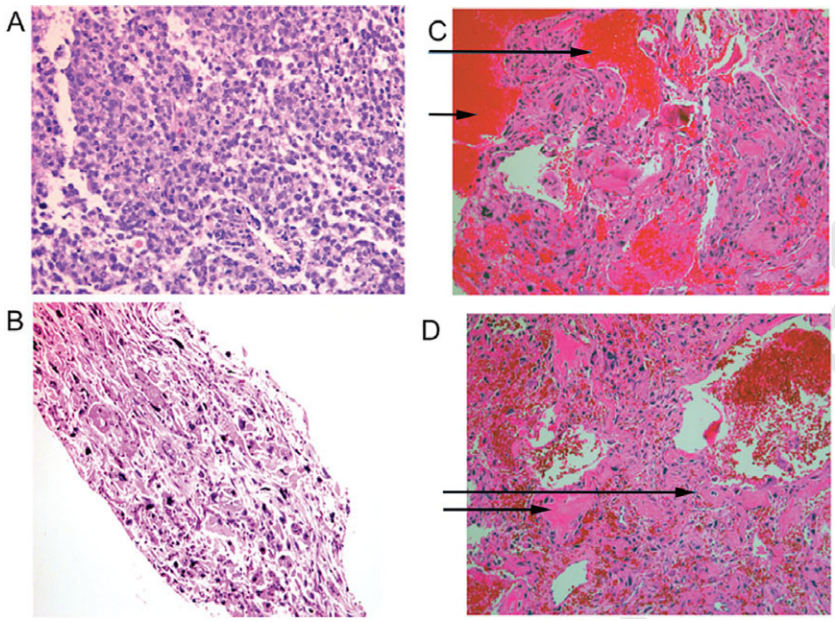



**FIGURE 1** Histopathology of the two primary tumors. Hematoxylin and eosin stain (A and B). The abdominal tumor showed both round cell (A) and pleomorphic spindle cell (B) morphology. No osteoid was identified. Smooth muscle actin immunohistochemical stain was strongly positive, favoring leiomyosarcoma. The bone tumor consisted of (C) spindle cells with blood-filled cystic spaces (arrows) and (D) osteoid production (arrows), consistent with an aggressive telangiectatic osteosarcoma.

and 1B). During the workup for metastasis, he was found to have a left humeral mass, which was biopsied and determined to be a telangiectatic osteosarcoma (Fig. 1C and 1D). His bone marrow (bm) at diagnosis had no evidence of disease. He completed five cycles of ifosfamide/doxorubicin and local radiation, followed by surgical resection, of the pelvic tumor. His disease progressed postoperatively, and he died 13 months from his initial diagnosis. There was no family history of cancer, other than a maternal uncle diagnosed with a glioblastoma multiforme in his seventh decade.

## 3 | METHODS AND RESULTS

Following informed consent, the patient and his parents provided peripheral blood lymphocytes for *TP53* DNA sequencing, performed by GeneDx. Sequencing of DNA from peripheral blood lymphocytes from the patient, but not his parents, revealed c.672 G>A germline variant preserving the glutamate at p. 224. However, this mutation is located at the last base pair (bp) of exon 6. An in silico algorithm (http://in-silico.net/tools/biology/sequence_conversion) predicts that this novel mutation produces a 5′ cryptic splice site at an alternative location.

Total RNA was then extracted from formaldehyde-fixed paraffin-embedded bm and leiomyosarcoma.[9] RNA was then transcribed into cDNA using SuperScript VILO master mix containing random primers (Invitrogen, Carlsbad, CA). The RNA yield, as well as the quality, was low (0.02 and 0.01 $\mu g/\mu l$). Short fragments of <150 bp were amplified and the best sets amplified 131 bp using nested polymerase chain reaction (PCR). Primers were designed to amplify the region spanning exons 6 and 7 junction (Fig. 2A). Amplified fragments were resolved by MetaPhor agarose gel, which demonstrated doublets (Fig. 2B). The higher molecular weight band seen in the samples is consistent with intronic retention (arrow in Fig. 2B). We believe the mutation is germline due to its presence in the bm where he had no evidence of disease. However, the expression of the mutant mRNA seems to be more stable in the bm, whereas the doublet in the tumor has less definition. The two amplified fragments were isolated from the gel and sequenced, using Sanger sequencing (GENEWIZ, South Plainfield, NJ). Sequencing of the 131 bp fragment proved challenging. To better delineate the mutation in such a small fragments, the fragments were inserted into the pcDNA3 plasmid (Invitrogen) resulting in sequencing of the entire fragment. Figure 2C shows the insertion of 5 bp from the intron in the mutated strand. Supplementary Figure S1 shows sequencing from the controls samples.

## 4 | DISCUSSION

We performed a search and surveyed the online *TP53* data in IARC *TP53*, NCBI dbSNP, and ExAc for mutations and their effects on the structure of p53 and its DNA-binding site. Other patients harbored synonymous mutations G>T and G>C at this site. His novel germline mutations (G>A), we predicted, would result in a splice variant based on the conserved sequences for splicing at the end of the exon, but has not been previously described in the literature. The U1 splicing complex binds to the 5′ intron–exon border based on conserved sequences. A mutation in these conserved sequences can result in a shift of the spliceosome binding site and insertion of intronic sequence into the mRNA transcript. We have found no reports of our patient's mutation. Review of the literature found a reported family with a G>A at

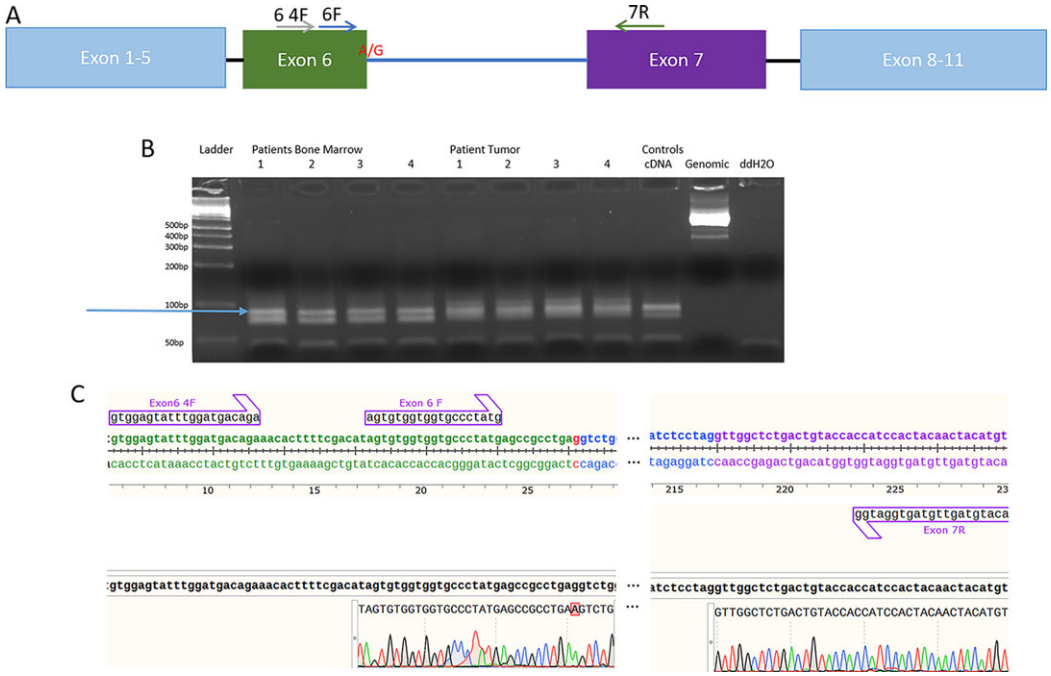



**FIGURE 2** Genomic organization of *TP53* and strategy of primer design. (A) Diagram to illustrate exons 2–11 of *TP53*, where the primers are in relation to the mutation. (B) Nested PCR product (lane 1): the DNA ladder; lanes 2–5: bone marrow (unaffected tissue) PCR done in quadruplicate; lanes 6–9 tumor: leimyosarcoma in quadruplicate; lane 10: human cDNA; and lane 11: human genomic DNA from an unaffected individual. A doublet is present in the bone marrow and tumor, but the tumor sample has shifted up, indicating a larger fragment. (C) Sanger sequence of the tumor depicting cryptic splice site after 5 bp intronic insertion, green is the 5′ of exon 6, red is the mutation, blue are the ends of the intron, purple is the 3′ of exon 7, and all three primers

+1 of the intron 6 resulting in an 18 bp insertion.[10] Given the conserved nature of this region and that it is a hotspot for oncogenic mutations, we believe our patient's pathogenic variant would have truncated the p53 protein in the DNA-binding site resulting in a change in the binding affinity to the DNA. Truncation mutations 5′ and 3′ of this site have been reported in adrenocortical carcinoma, sarcomas, and breast cancer[11,12] (http://p53.iarc.fr). Most *TP53* mutations result in an amino acid substitution in the DNA-binding domain. The mutations result in a full-length protein produced with a gain of function. Most splice variants lead to a loss of function due to exon skipping or early terminations. Interestingly, this de novo pathologic variant occurred in an adolescent, who was a product of IVF, but most evidence points to no increased risk of cancer in children conceived by IVF.[13]

Limited availability of tissue prevents us from further investigation whether there are more splice variants in the tumors. The cDNA tumor bands in the agarose gel were not as defined as they were in the unaffected bm. With the limited amount and quality of the tissue, we have not been able to answer this question. However, based on the bioinformatic analysis, we believe that the de novo synonymous germline mutation in *TP53* resulted in a shift of the exon 6 splice site by 5 bp, producing a frameshift and premature stop codon at residue 246 in exon 7. This pathogenic variant likely served as a driver for the two primary tumors in this patient. A recent report suggests that ~9% of pediatric cancers have a predisposition due to pathogenic variants. Of the genes affected, *TP53* is the most common.[14] Thus, identification of *TP53* germline mutations are important in pediatric cancers and cancer surveillance. Any child or adolescent with two primary cancers should undergo genetic testing and subsequent genetic counseling.

## ACKNOWLEDGMENTS

This work was supported partially by Department of Defense Bone Marrow Failure Idea Grant, the Children's Hospital Foundation, and Connors' Heroes Foundation (SJC). We thank Dr. Hrishikesh Mehta for technical advice on alternative splicing.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## REFERENCES

1. Joruiz SM, Bourdon J-C. p53 Isoforms: key regulators of the cell fate decision. *Cold Spring Harb Perspect Med*. 2016;6:1–20.
2. Wu D, Zhang Z, Chu H, et al. Intron 3 sixteen base pairs duplication polymorphism of p53 contributes to breast cancer susceptibility: evidence from meta-analysis. *PLoS One*. 2013;8:1–8.
3. Marcel V, Vijayakumar V, Fernández-Cuesta L, et al. p53 regulates the transcription of its Delta133p53 isoform through specific response elements contained within the TP53 P2 internal promoter. *Oncogene*. 2010;29:2691–2700.
4. Lawrence MS, Stojanov P, Mermel CH, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature*. 2014;505:495–501.
5. Li FP, Fraumeni JF, Mulvihill JJ, et al. A cancer family syndrome in twenty-four kindreds. *Cancer Res*. 1988;48:5358–5362.
6. Gonzalez KD, Noltner KA, Buzin CH, et al. Beyond Li-Fraumeni syndrome: clinical characteristics of families with p53 germline mutations. *J Clin Oncol*. 2009;27:1250–1256.

7. Chompret A, Brugières L, Ronsin M, et al. P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *Br J Cancer*. 2000;82:1932–1937.
8. Malkin D. Li-Fraumeni syndrome and p53 in 2015: celebrating their silver anniversary. *Clin Investig Med Medecine Clin Exp*. 2016;39:E37–E47.
9. Specht K, Richter T, Müller U, Walch A, Werner M, Höfler H. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol*. 2001;158:419–429.
10. Sakurai N, Iwamoto S, Miura Y, et al. Novel p53 splicing site mutation in Li-Fraumeni-like syndrome with osteosarcoma. *Pediatr Int*. 2013;1:107–111.
11. Ognjanovic S, Olivier M, Bergemann TL, Hainaut P. Sarcomas in TP53 germline mutation carriers: a review of the IARC TP53 database. *Cancer*. 2012;118:1387–1396.
12. Varley JM, McGown G, Thorncroft M, et al. Are there low-penetrance TP53 Alleles? evidence from childhood adrenocortical tumors. *Am J Hum Genet*. 1999;65:995–1006.
13. Williams CL, Bunch KJ, Stiller CA, et al. Cancer risk among children born after assisted conception. *N Engl J Med*. 2013;369:1819–1827.
14. Zhang J, Walsh MF, Wu G, et al. Germline mutations in predisposition genes in pediatric cancer. *N Engl J Med*. 2015;373:2336–2346.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Austin F, Oyarbide U, Massey G, Grimes M, Corey SJ. Synonymous mutation in *TP53* results in a cryptic splice site affecting its DNA binding site in an adolescent with two primary sarcomas. *Pediatr Blood Cancer*. 2017;64:e26584. https://doi.org/10.1002/pbc.26584

From www.bloodjournal.org by guest on January 10, 2019. For personal use only.

and deneddylation, which removes the Nedd8 moiety and requires the isopeptidase activity of the COP9 signalosome.[9] Among the 6 human cullins family, the cullin4 (CUL4) subfamily comprises 2 members, CUL4A and CUL4B, which share 83% sequence identity and functional redundancy. It consists of a RING finger domain protein, CUL4 scaffold protein, and DDB1-CUL4 associate substrate receptors. Recent studies have highlighted the role of CUL4A complexes in regulating substrates involved in the cell cycle, signaling, tumor suppression, DNA damage response, and chromatin remodeling and suggested CRL4A as a promising novel target for cancer therapy.[10]

CRBN is a substrate receptor protein for the CRL4A E3 ubiquitin ligase complex, and drugs like IMiDs have been reported to be able to inhibit or alter the substrate specificity of the E3 ligase activity of $CRL4A^{CRBN}$. Therefore, the identification of the molecular components that regulate IMiDs-dependent activity of $CRL4^{CRBN}$ will allow a better understanding of the parameters controlling their therapeutic efficacy and help to identify the mechanisms underlying their resistance. As such, through an elegant genome-wide CRISPR-CAS9 screen, the authors identified regulators of cullin-RING ligase neddylation as well as the elusive $CRL4A^{CRBN}$ E2 conjugating enzymes. By using functional genetics and in vitro assays, they demonstrated that UBE2D3 primes $CRL4^{CRBN}$ target substrates via monoubiquitination, after which UBE2G1 polyubiquinates them via K48-linked ubiquitin chains. Furthermore, loss of UBE2M or members of the COP9 signalosome resulted in altered neddylation of CUL4A and impaired lenalidomide-dependent $CRL4^{CRBN}$ activity (see figure).

Overall, the results presented here are novel and pivotal to expand our understanding of IMiDs' mechanisms of action and likely to elucidate mediators of acquired resistance. Additional questions remain, including whether other CRL4-containing E3 ligases utilize UBE2G1/UBE2D3 conjugating enzymes and what structural motifs on $CRL4^{CRBN}$ or its substrates dictate E2 usage and specificity. Nevertheless, Sievers et al, by CRISPRing the $CRL4^{CRBN}$ ring, established key proteins required for lenalidomide-dependent $CRL4^{CRBN}$ function in MM and provided us with a deeper understanding of this ubiquitin ligase function and regulation. The next step will be to therapeutically exploit these newly identified targets to augment or restore the activities of IMiDs in MM.

*Conflict-of-interest disclosure: The author declares no competing financial interests.* ■

## REFERENCES


1. Sievers QL, Gasser JA, Cowley GS, Fischer ES, Ebert BL. Genome-wide screen identifies cullin-RING ligase machinery required for lenalidomide-dependent $CRL4^{CRBN}$ activity. *Blood*. 2018;132(12):1293-1303.
2. Ito T, Ando H, Suzuki T, et al. Identification of a primary target of thalidomide teratogenicity. *Science*. 2010;327(5971):1345-1350.
3. Krönke J, Udeshi ND, Narla A, et al. Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science*. 2014;343(6168):301-305.
4. Petzold G, Fischer ES, Thomä NH. Structural basis of lenalidomide-induced CK1α degradation by the CRL4(CRBN) ubiquitin ligase. *Nature*. 2016;532(7597):127-130.
5. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821.
6. Komander D, Rape M. The ubiquitin code. *Annu Rev Biochem*. 2012;81(1):203-229.
7. Pickart CM. Mechanisms underlying ubiquitination. *Annu Rev Biochem*. 2001;70(1):503-533.
8. Duda DM, Borg LA, Scott DC, Hunt HW, Hammel M, Schulman BA. Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell*. 2008;134(6):995-1006.
9. Lyapina S, Cope G, Shevchenko A, et al. Promotion of NEDD-CUL1 conjugate cleavage by COP9 signalosome. *Science*. 2001;292(5520):1382-1385.
10. Hannah J, Zhou P. Distinct and overlapping functions of the cullin E3 ligase scaffolding proteins CUL4A and CUL4B. *Gene*. 2015;573(1):33-45.


DOI 10.1182/blood-2018-08-867069

© 2018 by The American Society of Hematology

**PHAGOCYTES, GRANULOCYTES, AND MYELOPOIESIS**

*Comment on Bellanné-Chantelot et al, page 1318*

# SRP54 and a need for a new neutropenia nosology

**Usua Oyarbide and Seth J. Corey** | Virginia Commonwealth University School of Medicine

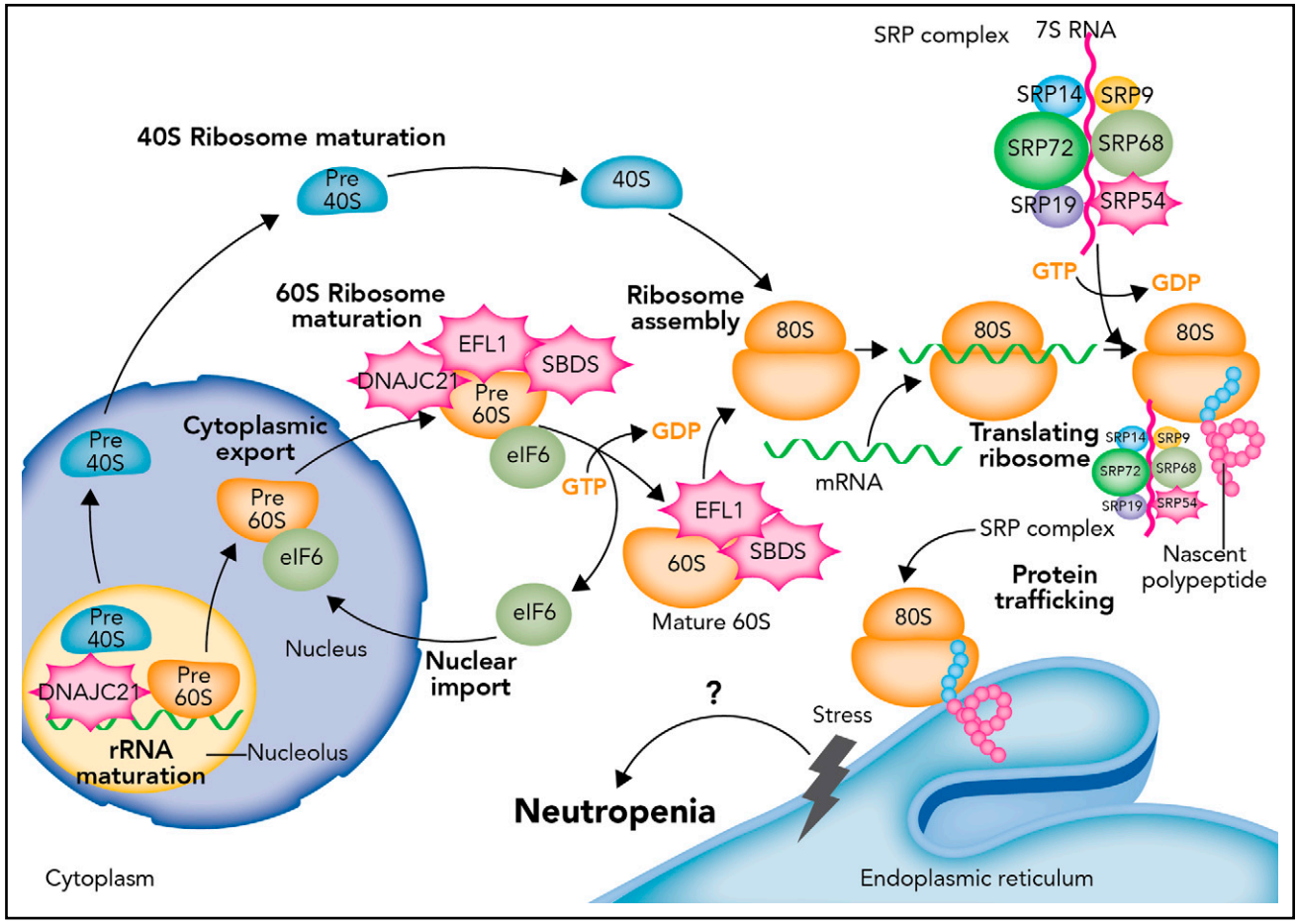
**In this issue of *Blood*, Bellanné-Chantelot et al noted that mutations in signal recognition particle 54 (*SRP54*) cause an inherited neutropenia in 23 individuals with features of both severe congenital neutropenia (SCN) and Shwachman-Diamond syndrome (SDS).[1]**

In fact, after *ELANE*, mutations in *SRP54* are the second most common cause of inherited neutropenia in the French Congenital Neutropenia Registry. Together with the 3 patients reported by Carapito et al,[2] the work of Bellanné-Chantelot et al adds to our understanding of genetic pathways and phenotype correlations of inherited neutropenias, but it raises the question of how to classify these and other genetic cytopenias: molecular pathway-based classification or phenotype-based classification?

The inherited bone marrow failure syndromes constitute a heterogeneous group of blood disorders that are phenotypically distinct.[3] They are true experiments of nature, enlightening our understanding of normal and pathologic hematopoiesis and often non-hematopoietic tissue development. Their importance goes further with their acting as leukemia and/or cancer predisposition syndromes. Molecular cloning has reinforced phenotype distinctions by revealing that mutations in pathways lead to the separate clinical entities. Fanconi anemia is caused by at least 20 different genes involved in DNA damage response, Diamond-Blackfan anemia by at least 15 genes involved in ribosomal structure, and dyskeratosis congenita by at least 10 genes that promote telomere maintenance and stability.

The inherited neutropenias include SCN, cyclic neutropenia (CyN), SDS, and a motley

From www.bloodjournal.org by guest on January 10, 2019. For personal use only.




Defects in the proteostasis pathway lead to inherited neutropenia. Molecular cloning of genetic lesions in individuals with moderate to severe neutropenias have centered on the biosynthetic pathway for proteins and their folding and trafficking (proteostasis for protein homeostasis). The most common genetic lesion in SDS is the biallelic mutation of SBDS. SBDS interacts with EFL1 to displace eukaryotic initiation factor 6 (eIF6). DNAJC21 stabilizes the 60S ribosome. As part of the SRP complex, SRP54 escorts the nascent polypeptide to the ER to complete translation and possible posttranslational modification. When defects arise in this prodigious and continuous process, unfolded protein response and ER stress follows. The most commonly mutated gene in SCN is *ELANE*, which encodes a serine protease. It is thought that mutated ELANE results in proteins that misfold and cause unfolded protein response. Through unknown mechanisms of apoptosis and/or differentiation impairment, granulopoiesis cannot be completed in SCN or sufficiently produced in SDS. How SCN and SDS transform frequently to MD or AML remains inadequately understood. GDP, guanosine diphosphate; GTP, guanosine triphosphate. Professional illustration by Somersault18:24.

group of monogenic disorders.[4] SCN results from more than 8 different genes, which encode a variety of proteins that have defied easy categorization. Approximately 50% of patients with SCN harbor mutations in the neutrophil serine protease ELANE. Individuals with SCN typically have recurrent life-threatening bacterial infections beginning during infancy. The absolute neutrophil count (ANC) is chronically <200 cells per μL. Recombinant human granulocyte colony-stimulating factor (G-CSF [filgrastim]) improves the severe neutropenia and dramatically reduces the occurrence of infections. For SCN patients, there is concern that filgrastim may accelerate the transformation of SCN to myelodysplastic syndrome (MDS) or acute myeloid leukemia (AML) because of acquired mutations in *CSF3R* (G-CSF receptor). CyN is another inherited disorder characterized by severe neutropenia that occurs every 21 days (the nadir may occur in 14- to 35-day cycles). Occasionally, erythrocyte and platelet numbers display cyclic variations. Almost all of the CyN is a result of germ line *ELANE* mutations. These individuals typically present in the first decade of life or later, with less severe or frequent infections but annoying mouth sores that resolve when the neutropenia abates.

SDS is characterized by short stature and thin habitus, neutropenia, pancreatic insufficiency, and skeletal anomalies. The neutropenia is classically moderate (ANC >500 cells per μL) and can fluctuate. A minority of patients have infections and/or depend on filgrastim to maintain ANC >750 cells per μL. SDS patients also have a high risk for transformation to MDS or AML. In up to 90% of patients, SDS results from biallelic mutations in the eponymous *SBDS* gene that encodes an assembly factor for the 80S ribosome. Mutational gene discoveries over the past 2 years in individuals who seem to have SDS but actually have wild-type SBDS have been identified: the chaperone *DNAJC21*,[5] the elongation factor-like GTPase 1 (*EFL1*),[6] and now *SRP54*. All these proteins are, in one way or another, involved in protein synthesis: ribosomal RNA (rRNA) or ribosomal subunit maturation and protein trafficking to the endoplasmic reticulum (ER). SBDS interacts with EFL1 on the pre-60S ribosome to promote assembly of the 80S ribosome from combining the 40S and 60S subunits, DNAJC21 is implicated in rRNA biogenesis and late cytoplasmic maturation of the 60S subunit, and SRP54 helps to direct proteins coming off the 80S ribosome and transfer them to the ER for correct trafficking. The

From www.bloodjournal.org by guest on January 10, 2019. For personal use only.

signal recognition particle (SRP) itself is a ribonucleoprotein complex of 6 proteins: SRP9, SRP14, SRP19, SRP54, SRP68, and SRP72 (see figure). SRP72 has been associated with familial aplastic and myelodysplastic syndromes.[7]

Bellanné-Chantelot and colleagues showed that *SRP54* mutants lead to ER stress and autophagy. Activation of the unfolded protein response pathways has also been described in patients who have SCN with *ELANE* mutations. They also observed a high reduction of proliferation of granulocytic cells and apoptosis dependent on the P53 pathway, characteristics also observed in patients with SDS and Diamond Blackfan anemia. Unlike SCN or SDS, none of the patients with SRP54 mutations (23 cases) transformed to leukemia, despite high doses of filgrastim, after a median follow-up of 15 years. Unlike in SCN or SDS, none of the 17 patients analyzed by deep sequencing harbored mutations in *CSF3R*, *RUNX1*, or *TP53*, which may be associated with risk of leukemogenesis.[8] How defects in these different pathways and cellular stress responses cause neutropenia that frequently transforms to AML or MDS is still not well understood. Might the risk of leukemic transformation and the identities of cooperating mutants exist in mutant *SRP54*-associated neutropenia? Must they await their emergence with a longer follow-up?

The title of Bellanné-Chantelot's article suggests ambivalence regarding which class of inherited neutropenia *SRP54* mutations fit in: SCN, SDS, or both. In this series, the median ANC was 220 cells per μL, and the use of filgrastim was high (87%) but not universal, unlike with SCN. In this series, only 2 individuals had exocrine pancreatic insufficiency, unlike with SDS. Six (of 23) had clinically significant neurologic abnormalities. Morphologically, there was a promyelocytic arrest, like with SCN and unlike with SDS. Dysplastic features were common, perhaps more like SDS.[9]

Classification of human diseases has advanced considerably since the ancient Greeks divided them mechanistically into the four humors. Next-generation sequencing has offered both precision and complexity to hematologic diseases. Should one lump or split syndromes? Even in monogenically defined blood diseases (eg, sickle cell anemia, hemophilia A, or SBDS-mutated SDS), how does one account for phenotypic diversity? Could a subatomic genetic classification be looming? Apropos, does an *SRP54* mutation associated with neutropenia belong to SCN or SDS? Can there be a useful classification scheme for the inherited neutropenias, the list of which grows seasonally?

*Conflict-of-interest disclosure: The authors declare no competing financial interests.* ■

## REFERENCES


1. Bellanné-Chantelot C, Schmaltz-Panneau B, Marty C, et al. Mutations in the *SRP54* gene cause severe congenital neutropenia as well as Shwachman-Diamond–like syndrome. *Blood*. 2018;132(12):1318-1331.
2. Carapito R, Konantz M, Paillard C, et al. Mutations in signal recognition particle SRP54 cause syndromic neutropenia with Shwachman-Diamond-like features. *J Clin Invest*. 2017;127(11):4090-4103.
3. Bluteau O, Sebert M, Leblanc T, et al. A landscape of germ line mutations in a cohort of inherited bone marrow failure patients. *Blood*. 2018;131(7):717-732.
4. Corey SJ, Oyarbide U. New monogenic disorders identify more pathways to neutropenia: from the clinic to next-generation sequencing. *Hematology Am Soc Hematol Educ Program*. 2017;2017(1):172-180.
5. Tummala H, Walne AJ, Williams M, et al. DNAJC21 mutations link a cancer-prone bone marrow failure syndrome to corruption in 60S ribosome subunit maturation. *Am J Hum Genet*. 2016;99(1):115-124.
6. Stepensky P, Chacón-Flores M, Kim KH, et al. Mutations in *EFL1*, an *SBDS* partner, are associated with infantile pancytopenia, exocrine pancreatic insufficiency and skeletal anomalies in a Shwachman-Diamond like syndrome. *J Med Genet*. 2017;54(8):558-566.
7. Kirwan M, Walne AJ, Plagnol V, et al. Exome sequencing identifies autosomal-dominant SRP72 mutations associated with familial aplasia and myelodysplasia. *Am J Hum Genet*. 2012;90(5):888-892.
8. Xia J, Miller CA, Baty J, et al. Somatic mutations and clonal hematopoiesis in congenital neutropenia. *Blood*. 2018;131(4):408-416.
9. Lesesve JF, Broséus J. Dysplastic neutrophils in the bone marrow of a Shwachman-Diamond syndrome patient. *Blood*. 2017;130(1):96.


DOI 10.1182/blood-2018-07-859959

© 2018 by The American Society of Hematology

**PLATELETS AND THROMBOPOIESIS**

*Comment on Kho et al, page 1332*

# Platelets in malaria pathogenesis

**Jamie M. O'Sullivan and James S. O'Donnell** | Royal College of Surgeons in Ireland

**In this issue of *Blood*, Kho et al provide further evidence that platelets play a major role in the pathogenesis of malaria infection.[1] In particular, they demonstrate that platelets can kill circulating parasites of all major *Plasmodium* species in human malaria. Elucidating the molecular mechanisms underpinning this platelet-directed killing mechanism may offer the opportunity to develop novel adjunctive antimalarial therapies.**

Human malaria continues to be a leading cause of mortality, with an estimated 500 000 deaths per year. Unfortunately, the majority of these deaths occur in sub-Saharan children under 5 years of age. In addition, it is important to recognize that malaria is also associated with significant global morbidity. For example, many children who survive cerebral malaria (CM) suffer secondary long-term neurologic sequelae. Despite this clinical burden, perhaps surprisingly, the biological mechanisms involved in the pathophysiology of severe malaria remain relatively poorly defined. Interestingly, thrombocytopenia is a common finding in humans and mice with malaria infection and is most marked in patients with severe *Plasmodium falciparum* infection.[2] Moreover, previous studies have reported that extent of thrombocytopenia correlates with parasite density, severity of malaria infection, and clinical outcomes. Together, these data support the hypothesis that platelets are important in malaria pathogenesis. Critically, however, accumulating

From www.bloodjournal.org by guest on January 10, 2019. For personal use only.

2018 132: 1220-1222
doi:10.1182/blood-2018-07-859959

# SRP54 and a need for a new neutropenia nosology

Usua Oyarbide and Seth J. Corey

Updated information and services can be found at:
**http://www.bloodjournal.org/content/132/12/1220.full.html**

Articles on similar topics can be found in the following Blood collections
Free Research Articles (5262 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
**http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests**

Information about ordering reprints may be found online at:
**http://www.bloodjournal.org/site/misc/rights.xhtml#reprints**

Information about subscriptions and ASH membership may be found online at:
**http://www.bloodjournal.org/site/subscriptions/index.xhtml**

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.
Copyright 2011 by The American Society of Hematology; all rights reserved.

# Peering through zebrafish to understand inherited bone marrow failure syndromes

**Usua Oyarbide,[1]* Jacek Topczewski[2,3] and Seth J. Corey[1,4,5]***

[1]Department of Pediatrics, Children's Hospital of Richmond and Massey Cancer Center at Virginia Commonwealth University, Richmond, VA, USA; [2]Department of Pediatrics, Stanley Manne Children's Research Institute, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; [3]Department of Biochemistry and Molecular Biology, Medical University of Lublin, Poland; [4]Department of Microbiology/Immunology, Virginia Commonwealth University, USA and [5]Department of Human and Molecular Genetics, Virginia Commonwealth University, Richmond, USA

**Current address: Departments of Pediatrics, Translational Hematology and Oncology Research, and Cancer Biology, Cleveland Clinic, Cleveland, OH, USA*

Ferrata Storti Foundation

**Haematologica** 2019
Volume 104(1):13-24

## ABSTRACT

Inherited bone marrow failure syndromes are experiments of nature characterized by impaired hematopoiesis with cancer and leukemia predisposition. The mutations associated with inherited bone marrow failure syndromes affect fundamental cellular pathways, such as DNA repair, telomere maintenance, or proteostasis. How these disturbed pathways fail to produce sufficient blood cells and lead to leukemogenesis are not understood. The rarity of inherited cytopenias, the paucity of affected primary human hematopoietic cells, and the sometime inadequacy of murine or induced pluripotential stem cell models mean it is difficult to acquire a greater understanding of them. Zebrafish offer a model organism to study gene functions. As vertebrates, zebrafish share with humans many orthologous genes involved in blood disorders. As a model organism, zebrafish provide advantages that include rapid development of transparent embryos, high fecundity (providing large numbers of mutant and normal siblings), and a large collection of mutant and transgenic lines useful for investigating the blood system and other tissues during development. Importantly, recent advances in genomic editing in zebrafish can speedily validate the new genes or novel variants discovered in clinical investigation as causes for marrow failure. Here we review zebrafish as a model organism that phenocopies Fanconi anemia, Diamond-Blackfan anemia, dyskeratosis congenita, Shwachman-Diamond syndrome, congenital amegakaryocytic thrombocytopenia, and severe congenital neutropenia. Two important insights, provided by modeling inherited cytopenias in zebrafish, widen understanding of ribosome biogenesis and TP53 in mediating marrow failure and non-hematologic defects. They suggest that TP53-independent pathways contribute to marrow failure. In addition, zebrafish provide an attractive model organism for drug development.

## Introduction

The inherited bone marrow failure syndromes (IBMFs) comprise a diverse group of rare monogenic disorders that are phenotypically heterogeneous. They may involve a single or multiple lineage(s). The classic disorders are: Fanconi anemia (FA), Diamond-Blackfan anemia (DBA), Shwachman-Diamond syndrome (SDS), dyskeratosis congenita (DC), severe congenital neutropenia (SCN), and congenital amegakaryocytic thrombocytopenia (CAMT). Besides their phenotypic characterizations, these syndromes correlate strongly with mutations involving a specific pathway. FA results from mutations in genes encoding components of the DNA damage response,[1] DC in telomere maintenance,[2] and DBA in ribosome function.[3]

**Correspondence:**
coreys2@ccf.org

Received: August 20, 2018.
Accepted: November 14, 2018.
Pre-published: December 20, 2018.

doi:10.3324/haematol.2018.196105

*Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/104/1/13*

©2019 Ferrata Storti Foundation

*Material published in Haematologica is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode. Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.*

SDS is emerging as a disorder in proteostasis and ribosome maturation (Table 1).[4] The molecular basis for how these phenotypically and genotypically heterogeneous conditions result in single or multiple cytopenias remains poorly understood. No common pathway has yet been established, but zebrafish studies have suggested TP53 responses. Activation of the TP53 pathway in mediating marrow failure has been reported for DC,[5] FA,[6] and a novel bone marrow failure syndrome.[7] The TP53 pathway has been suggested to mediate marrow failure for other inherited neutropenias such as SCN and SDS.[8] Environmental exposures can accelerate marrow failure, for example, aldehydes producing DNA crosslinks in FA.[9] How epigenetics and genetic co-modifiers contribute to these diseases is even less understood. Investigating the molecular basis of the IBMFS will lead to a greater understanding of hematopoiesis, and development and maintenance of non-hematologic tissues. Since the IBMFS constitute leukemia or cancer predisposition syndromes, insights into their pathophysiology will also benefit our understanding, prevention, and perhaps treatment of cancer and age-related genetic changes.

## Zebrafish model to study inherited bone marrow failure syndromes

Zebrafish (*Danio rerio*) have gained popularity as a model organism for a number of reasons. Approximately 70% of all human genes have a zebrafish ortholog.[10] Genes are orthologs if they evolved from a common gene, and orthologs typically share similar function. (The Human Genome Organization has adopted a nomenclature for gene and protein expression among different species, which we show using SDS as an example in Table 2.) In addition to lower maintenance and breeding costs, zebrafish provide major advantages to mice: their large clutch size of externally fertilized eggs, transparent embryos, quicker development (all major organs develop and begin functioning during the first 5 days), and short generational time to gamete formation.[11] A high degree of genetic and morphological similarity in hematopoiesis between zebrafish and humans suggests that zebrafish can provide valuable insights into the pathogenesis of IBMFS. Developmental hematopoiesis in the zebrafish is comparable to that observed in mice or humans (Figure 1).[12-15] One notable difference is that the site of definitive hematopoiesis lies in the zebrafish kidney perivascular space, not the bone marrow. Since the hematopoietic stem cell (HSC) niche provides protection and regulation of self-renewal and differentiation of HSC into blood cells, this difference may be important in non-cell autonomous processes.

Studies using zebrafish have facilitated our understanding of vertebrate hematopoiesis and aberrant hematopoiesis in diseases. Hematopoietic and non-hematopoietic lineage-specific transgenic reporter strains are available. They have been useful for the identification and characterization of genes for embryonic hematopoiesis, erythropoiesis, and modeling of human blood diseases (Table 3).[16-19] In addition to a collection of zebrafish mutants induced by *N*-ethyl-N-nitrosourea or

**Table 1.** Inherited bone marrow failure syndromes.

| Disease | Prevalence per 1,000,000 | Male-to-female ratio | Symptoms | Genes involved and their estimated frequency | Cancer predisposition |
|---|---|---|---|---|---|
| Diamond-Blackfan anemia (DBA) | 5-7 | 1:1 | Erythroid failure, congenital malformations, growth retardation, short stature. Thumbs, upper limbs, hands, and craniofacial, urogenital, and cardiovascular anomalies are also common | *RPS19 (25%), RPL5 (7%), RPS26 (6.6), RPL11 (5%), RPL35a (3%), RPS10 (3%), RPS24 (2.4%), RPS17 (1%), RPL15, RPS28, RPS29, RPS7, RPS15, RPS27a, RPS27, RPL9, RPL18, RPL26, RPL27, RPL31, TSR2, GATA1, EPO* | AML, MDS, ALL, Hodgkin and non-Hodgkin lymphomas, osteogenic sarcoma, breast cancer, hepatocellular carcinoma, melanoma, fibrohistiocytoma, gastric cancer, colon cancer |
| Dyskeratosis congenita (DC) | 1 | 3:1 | Abnormal skin pigmentation, nail dystrophy, mucosal leukoplakia, pulmonary fibrosis, and bone marrow failure | *DKC1 (17-36%), TERC (6-10%), TERT(1-7%), NHP2 (<1%), NOP10 (<1%), CTC1 (1-3%), WRAP53 (3%) and TINF2 (11-24%), ACD, PARN, RTEL1, USB1, TCAB1, POT1, TPP1, WRD79, TR, NOLA2, NOLA3* | AML, solid tumors |
| Fanconi anemia (FA) | 3 | 1.2:1 | Developmental abnormalities in a number of organ systems and bone marrow failure | *FANCA (65%), FANCB (<1%), FANCC (14%), FANCG (10%), FANCD1/BRCA2 (<1%), FANCD2 (<1%), FANCE (4%), FANCF (4%), RAD51, FANCC1, FANL, FANCL, FANC, PALPB2, RADC51C, SLX4, FANCQ. BRCA1, FANCT* | AML, solid tumors |
| Shwachman-Diamond syndrome (SDS) | 13 | 1.7:1 | Exocrine pancreatic insufficiency, bone marrow dysfunction and skeletal abnormalities | *SBDS (90%) DNAJC21 EFL1, SRP54* | AML, MDS |
| Congenital amegakaryocytic thrombocytopenia (CAMT) | Unknown (less than 100 cases reported) | | Thrombocytopenia and megakaryocytopenia | *MPL* | AML, MDS |
| Severe congenital neutropenia (SCN) | 5 | | Neutropenia | *ELANE, GFI1, HAX1, G6PC3, VPS45, JAG1, CSF3R, WAS, SRP54* | AML, MDS |

AML: acute myeloid leukemia; ALL: acute lymphocytic leukemia; MDS: myelodysplastic syndromes.

viral insertion,[20,21] gene function can be studied by transgenic expression or genome editing by transcription activator-like effector nucleases (TALEN) or Cas nucleases acting on clustered, regularly interspaced, short palindromic repeats (CRISPR). Gene expression can be silenced temporarily and early during development by injection of morpholino antisense nucleotides (MO).

Zebrafish have provided a useful model organism for a quick validation and study of human disease candidate genes, including those involved in the pathophysiology of IBMFS (Table 4). MO-mediated knockdown was widely used to probe gene function, though this method has limitations. Phenotype of morphants (MO-injected animals) can differ and is often more severe than those of the corresponding mutants. There could be different reasons for this: 1) phenotypic rescue of zygotic mutants by maternal wild-type mRNA; 2) off-target effects of the MO; 3) hypomorphic nature of the mutant allele analyzed; or 4) genetic compensation in mutants but not in morphants (see Stainier *et al.*[22]). Moreover, injection of MO can cause Tp53 activation and cell death.[23] In some instances, cell death can be prevented by simultaneous blocking of p53 by a second MO. This may lead to a misinterpretation of results, particularly in processes that depend on the Tp53 DNA damage response pathway (reviewed below). In some cases, results of MO knockdowns were not recapitulated with the genome editing techniques.[24] Close examination of the differences in gene expression revealed a novel compensation mechanism that operates only after mutation but not after MO knockdown (Table 5).[25]

## Diamond-Blackfan anemia

Diamond-Blackfan anemia is characterized by red cell hypoplasia, erythroid macrocytosis, and markedly reduced erythroid precursors in the bone marrow. Other hematopoietic lineages are usually normal at birth,[26] but they may be affected later in childhood/adolescence.[27] In

**Table 2.** Gene and protein nomenclature among species

| | Gene symbol | Protein symbol |
|---|---|---|
| Human | *sbds* | SBDS |
| Mouse | *sbds* | SBDS |
| Zebrafish | *sbds* | Sbds |

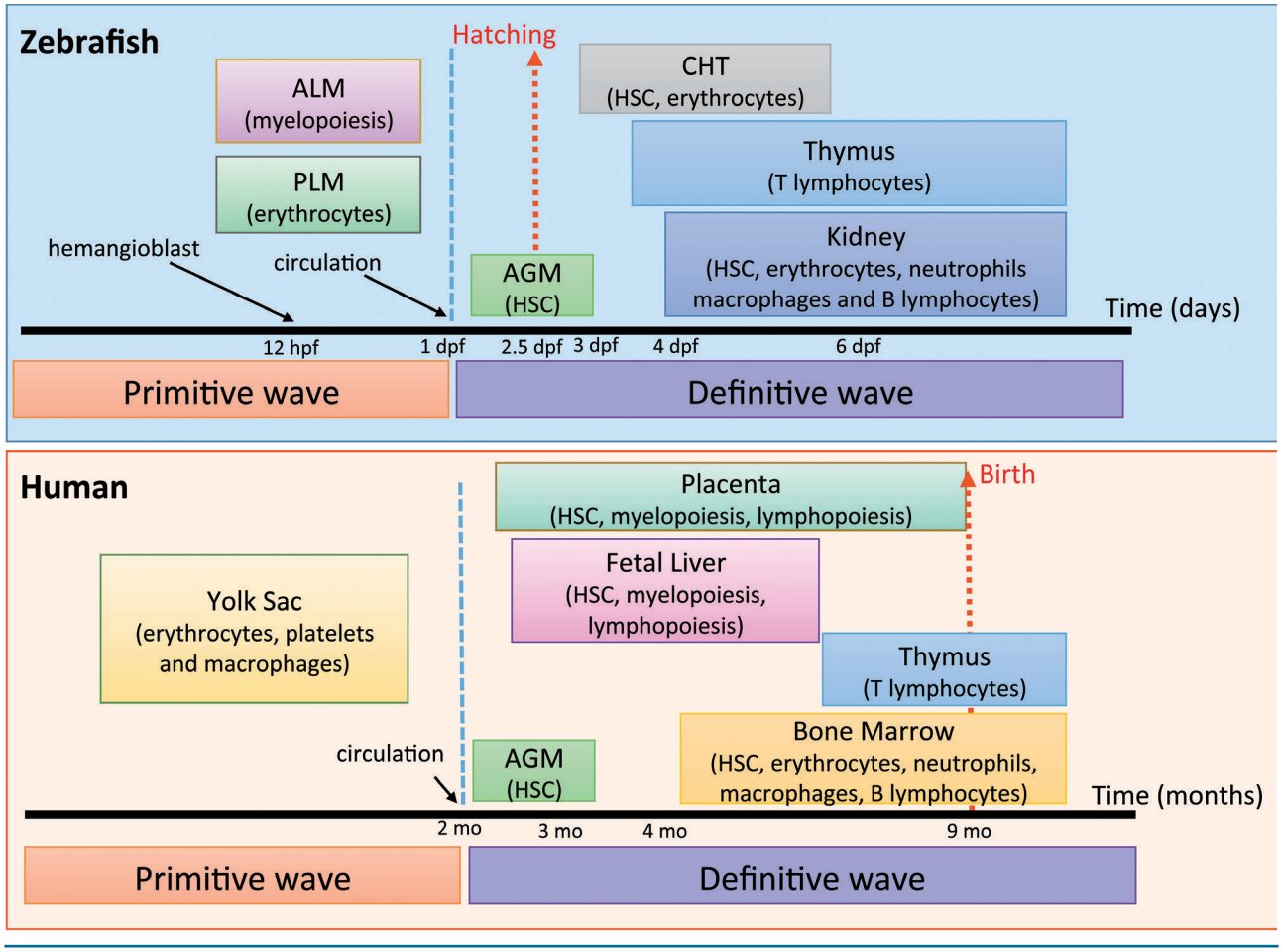



**Figure 1. Comparison of developmental hematopoiesis in humans and zebrafish.** Primitive and definitive hematopoiesis occurs in both species. In human, hematopoietic stem cells (HSC) originate in aorta-gonad-mesonephros (AGM) and placenta, from where they colonize fetal liver and finally the bone marrow. In zebrafish, primitive hematopoiesis starts after hemangioblast formation around 12 hpf in the anterior lateral mesoderm (ALM) and posterior lateral mesoderm (PLM). Later, HSCs originate in the AGM and then mobilize to caudal hematopoietic tissue (CHT) prior to their final destination of the kidney (modified from Teittinen *et al.*[12] and de Jong and Zon[14]).

addition to severe anemia, individuals with DBA may display physical anomalies that include thumb, upper limb, craniofacial, cardiovascular and kidney malformations, and short stature. DBA patients have a 25% higher risk of developing myelodysplastic syndromes (MDS), acute myeloid leukemia (AML), and osteosarcoma.

Diamond-Blackfan anemia is an autosomal dominant disorder with a disease incidence of 5-7 per million live births, equally distributed between genders.[28] DBA patients have mutations in approximately 20 genes encoding ribosomal proteins; the most common (25%) is RPS19.[29] Frameshift, splice defects, intragenic deletions and insertions, nonsense, as well as missense mutations have all been identified. Mutations involve other ribosomal genes: *RPL5* (7%), *RPL26* (6.6%), *RPL11* (5%), *RPS10* (3%), *RPS26* (3%), *RPL35A* (3%), *RPS24* (2.4%), *RPS17* (1%), *RPL15, RPS28, RPS29, RPS7, RPS15, RPS27a, RPS, RPL9, RPL18, RPL26, RPL27, RPL31*.[3] These findings support DBA as a disorder of ribosomal biogenesis and/or function. Mutations in three non-ribosomal proteins, GATA1, TSR2, and EPO, are also associated to DBA.[3,29] It is hypothesized that DBA results from apoptosis due to aberrant activation of TP53 that induces cell cycle arrest or apoptosis in response to ribosomal stress.[30] Some of the reports implicating TP53, as reviewed below, are based on MO-mediated effects.

In two different studies, *rps19*-deficient zebrafish were created using MO. Knockdown of *rps19* in zebrafish recapitulates the hematopoietic and developmental phenotypes of DBA, including erythropoietic failure with severe anemia, with cell cycle arrest and increased apoptosis, with p53 upregulation. The *rps19*-deficient phenotype was rescued by injection of zebrafish *rps19* mRNA.[31-33] Moreover, these phenotypes were not rescued by expressing *rps19* mRNAs with a missense or nonsense mutation found in DBA patients.[32] Co-injection of MOs against *rps19* and *p53*, showed a complete rescue of the morphological abnormalities, but did not rescue the hematologic defects. These results suggest that there is an erythroid specificity in *Rps19* deficiency in zebrafish, independently of Tp53 activity. (See below for further discussion on Tp53 in DBA pathogenesis).[34]

Chakraborty *et al.* analyzed the effect of MO-mediated loss of *rpl11* in zebrafish. Knockdown of *rpl11* led to morphological defects in the developing brain, head, and eyes, and pericardial edema. These phenotypes appear specific as the investigators were able to suppress the morphant by co-injection of MO-resistant *rpl11* mRNA. Similar to the loss of Rsp19 function, knockdown of *rpl11* resulted in an upregulation of *tp53* and *mdm2*. Moreover, co-injection of *rpl11* and *tp53* MO rescued the developmental defects and reduced apoptosis, suggesting that ribosomal dysfunction due to the loss of Rpl11 activates a Tp53-dependent response to prevent faulty embryonic development.

**Table 3.** Comparison of human, mouse and zebrafish blood systems.

| | Human | Mouse | Zebrafish |
|---|---|---|---|
| Adult HSC | Bone marrow | Bone marrow | Kidney marrow |
| Blood cell types | Erythrocytes, granulocytes, lymphocytes and platelets | Erythrocytes, granulocytes, lymphocytes and platelets | Erythrocytes, granulocytes, lymphocytes and thrombocytes |
| Erythrocytes (life span) | Without nucleus (115 days) | Without nucleus (60 days) | With nucleus (at least 10 days) |
| Platelets (life span) | Platelets (8-9 days) | Platelets (4 days) | Thrombocytes (4 days) |
| Neutrophils (life span) | Segmented nuclei with up to four lobes mpo-expressing cells (5.4 days) | Twisted toroid with a central hole mpo expressing cell (12.5 hours) | Segmented nuclei with two or three lobes mpo expressing cells (3.5 days) |
| Primitive myelopoiesis | Yolk sac, AGM, fetal liver | Yolk sac (E7.25-E10), AGM, fetal liver (after E9.5) | ALM (~11 hpf) and CHT (~24 hpf) |
| Definitive myelopoiesis | Fetal liver and bone marrow | Fetal liver (E9.5) Bone marrow | Kidney (~HSC starts seeding at 4 dpf) |
| Primitive erythropoiesis | Yolk sac (3-4 weeks) | Yolk sac (E7.0) | ICM (~12 hpf) |
| Definitive erythopoiesis | Yolk sac (4 weeks) Fetal liver (5-6 weeks) and then bone marrow | Yolk sac (E9.5), fetal liver (E12.5) and then bone marrow | CHT (2-6 dpf) and then kidney marrow (4 dpf) |
| Circulation | Begins at 8 weeks | Begins at E8.5 | Begins at 24 hpf |
| Primitive thrombopoiesis | N/A | N/A | CHT (~48 hpf) |
| Definitive thrombopoiesis | Bone marrow | Bone marrow | Kidney marrow (~5 dpf) |
| Developmental HSC | AGM next fetal liver and finally bone marrow | AGM next fetal liver and finally bone marrow | AGM next CHT and finally kidney marrow |
| B cells | Bone marrow | Bone marrow | Kidney marrow |
| T-cell maturation | Thymus (8-9 weeks) | Thymus (E10-12) | Thymus (7 dpf) |

AGM: aorta-gonadal-mesonephros; ALM: anterior lateral mesoderm; CHT: caudal hematopoietic tissue; dpf: days post fertilization; ICM: intermediate cell mass.

**Table 4.** Comparison of mouse and zebrafish models for inherited bone marrow failure syndromes.

| Disease | Phenotype of mouse model<br>Mouse protein similarity with human protein (%) | Zebrafish protein with similarity with human protein (%) | Phenotype zebrafish morphant |
|---|---|---|---|
| DBA | **RPS19 (99%)** | **Rps19 (86%)** | |
| | *Rps19* KO: embryonic lethal, heterozygous fully compensated<br><br>*Rps19* with ENU-induced missense mutation: embryonic lethality in homozygous. Heterozygous, mild anemia and growth retardation. L-leucine improved the anemia.<br><br>Rps19 deficiency (transgenic line): anemia, leukopenia and bone marrow failure. Loss of p53 rescued the phenotype.[91] | rps19 mutants. Erythroid defects, compensated for the loss of one Rps19 allele developmental defect and tp53 activation, fully compensated in heterozygous. Decreased HSCs.[30,90] | *rps19* morphant. Severe anemia and developmentala bnormalities. Dysregulation of delta Np63 and tp53.[31] |
| | **RPL11 (100%)** | **Rpl11 (96%)** | |
| | *Rpl11* KO embryonic lethal. Heterozygous, haploinsufficiency: anemia, decreased erythroid progenitors.[92] | *rps11* mutants. Erythroid defects, developmental defects and tp53 activation. Decrease HSCs.[30,36,38] | *rpl11* morphant. Morphological defects in the developing brain, small head and eyes and pericardial edema. Upregulation of *tp53* and *mdm2*.[35] |
| | **RPS29 (100%)** | **Rps29 (96%)** | |
| | N/A | *rps29* mutant. Severe anemia and increased apoptosis. P53 mutations near completely rescued rps29 morphological and hematopoietic phenotype.[93] | *rps29* morphant. Defects in red blood cell development and an increase in apoptotic cells.[45] |
| | **RPL5 (98%)** | **Rpl5 (88%)** | |
| | *Rpl5* KO embryonic lethal. Heterozygous fully compensated.[96] | N/A | *rpl5* morphant. Primitive and definitive hematopoiesis affected and morphological abnormalities. |
| | **RPS24 (90%)** | **Rps24 (87%)** | |
| | *Rps24* KO embryonic lethal. Heterozygous fully compensated.[91] | N/A | *rps24* morphant. Morphological defects: aplasia in the brain, a bent tail and reduced size. Severe anemia, in a tp53-independent manner.[42] |
| | **RPL35 (98%)** | **Rpl35 (92%)** | |
| | N/A | *rpl35* mutant very high tumor incidence (100%).[37,43] | *rpl35a* morphants. Morphological defects: aplasia in the brain, a bent tail and reduced size. Severe anemia, in a tp53-independent manner.[42] |
| | **RPL14 (94%)** | **Rpl14 (72%)** | |
| | Conditional deletion of Rps14 (and 8 other genes): anemia, bone marrow apoptosis.[91] | *rpl14* mutant: high number of tumors (74%)[43] | *rpl14* morphant. Severe anemia[45] and morphological abnormalities[98] |
| | **RPS7 (100%)** | **RRps7 (96%)** | |
| | Rps7 mutations ($RPS7^{V156G}$ and $RPS7^{Y177S}$): small size, abnormal skeleton and eye malformation. No anemia.[94] | *rps7* mutant. Hematopoietic and developmental defect. High tumor incidence (47%).[36,37,43] | *rps7* morphant. Impaired hematopoiesis and tp53 activation.[40] |
| | **RPL35A (99%)** | **Rpl35a (90%)** | |
| | N/A | N/A | *rpl35a* morphants. Morphological defects: aplasia in the brain, a bent tail and reduced size. Severe anemia, in a tp53-independent manner.[42,95] |
| | **RPS27 (100%)** | **Rps27 (98%)** | |
| | N/A | N/A | *rps27* morphant. Defective erythropoiesis and morphological abnormalities.[99] |
| | **RPS11 (92%)** | **Rps11 (91%)** | |
| | N/A | *rps11* mutants. Erythroid defects and tp53 activation. | N/A |

*continued on the next page*

*continued from the previous page*

| | | | |
|---|---|---|---|
| DC | **DKC1 (91%)** | | **Dkc1 (80%)** |
| | Hypomorphic Dkc1 mutant recapitulate in the first and second generations (G1 and G2) the clinical features of DC.[96] *Dkc1*$^{\Delta 15}$ mice: growth retardation, increased DNA damage response *via* ATM/p53 pathway.[97] | N/A | *dkc1* morphant. Reduced hematopoiesis, increased tp53 expression, and defective ribosomal biogenesis, no detectable changes in telomerase function.[50] |
| | **NOLA1 (96%)** | | **Nola1 (91%)** |
| | N/A | *nola1* mutant. Reduced hematopoiesis, increased tp53 expression, and defective ribosomal biogenesis, no detectable changes in telomerase function.[50] | N/A |
| | **TERT (62%)** | | **Tert (33%)** |
| | Transgenic line over-expressing TERT: short telomeres and increased DNA damage.[98] | *tert* mutant. Tissue atrophy, premature death, sarcopenia, impaired cell proliferation and accumulation of senescence cells.[55,57] | N/A |
| FA | **FANCD2 (65%)** | | **Fancd2 (53%)** |
| | FancD2 KO: reduced fertility, growth retardation and increased incidence of tumors.[99] | N/A | *fancd2* morphant. Shortened body length, microcephaly and abnormally small eyes, which are due to extensive cellular apoptosis. Upregulation of tp53.[63,64] |
| | **BRCA2 (57%)** | | **Brca2 (41%)** |
| | BRCA2 mutant: embryonic lethality | *brca2* mutants. Genomic instability.[100] | N/A |
| | **RAD51 (98%)** | | **Rad51 (93%)** |
| | *Rad51* mutants. Decreased cell proliferation, embryonic lethal.[101] | *rad51* mutants. Only infertile males, size reduction, hypocellular kidney marrow. Double mutants for Rad51 and P53 rescued HSPC defect but showed higher tumor incidence.[66] | N/A |
| SDS | **SBDS (97%)** | | **Sbds (87%)** |
| | Sbds KO: embryonic lethal.[102] | *sbds* mutant. Size reduction, liver, pancreas and digestive tract atrophy and reduction of neutrophils.[78] | *sbds* morphant. Loss of neutrophils, abnormal skeletal architecture and pancreatic hypoplasia. Sbds knockdown phenotype not rescued by loss of tp53.[76,77] |
| CAMT | **MPL (80%)** | | **Mpl (23%)** |
| | *c-Mpl* KO. Decrease platelets and megakaryocytes | *mpl* mutant. Low number of thrombocytes.[83] | *Mpl* morphant. Low number of thrombocytes.[103] |
| SCN | **CSF3R (73%)** | | **Csf3r (44%)** |
| | *Csf3R* KO. Low number of neutrophils in peripheral blood. Expression of truncated Csf3r confers a strong clonal advantage to HSCs.[104] | *csf3r* mutant. Reduction in neutrophils and myeloid cells in the kidney marrow.[87] | N/A |
| | **SRP54 (99%)** | | **Srp54 (95%)** |
| | N/A | N/A | *srp54* morphant. Loss of neutrophils and chemotaxis, diminished exocrine pancreas.[88] |

*Morphant: an organism that has been treated with a morpholino antisense to temporarily knockdown the expression of a gene.

An increase in *tp53* expression and its target genes, *cdkn1a* and *mdm2*, was observed in *rpl11* morphants. Genes involved in apoptosis (*bik, bax, puma,* and *noxa*) were also up-regulated.[35] Danilova *et al.* demonstrated that developmental and hematopoietic defects, and lower expression of *α-E1* globin and *hbae1.1* in Rps19-deficient fish were mediated by Tp53 upregulation. Upregulation of *tp53* also occurred in zebrafish mutants for *rps8*, *rps11* and *rps18*.[31]

Danilova *et al.* used a zebrafish *rpl11* mutant to characterize the molecular pathways associated with ribosomal deficiency.[36] This mutant showed anemia, decreased HSCs, and activation of the Tp53 pathway with altered expression in genes involved in cell cycle arrest (*cdkn1a* and *ccng1*) and apoptosis (*bax* and *puma*). Moreover, abnormal regulation of metabolic pathways with a shift from glycolysis to aerobic respiration, upregulation of genes involved in gluconeogenesis and insulin levels, decreased biosynthesis, and increased catabolism were observed. Nucleotide metabolism was affected by upregulation of adenosine deaminase (ada) and xanthine dehydrogenase/oxidase (xdh).[33,37] They showed that treatment of mutant embryos with an exogenous supply of

nucleosides resulted in downregulation of *tp53* and its targets with normalization of *ada* and *xdh* levels. Interestingly, DBA patients show increased erythrocyte adenosine deaminase activity.[38]

Zhang *et al.* generated two zebrafish mutants using TALENs for *rps19* and *rps11*. The knockout of both *rps19* and *rps11* resulted in the erythroid defects similar to DBA, such as lack of mature red blood cells (RBCs) and Tp53 activation. The mutants had significantly reduced production of globin proteins accompanied by either increased or unaffected level of mRNA transcripts. Furthermore, they observed decreased HSCs at 3 dpf in *rps19* mutants and hemoglobin levels by 4 dpf. The authors concluded that this reduction in RBCs may be caused by a decreased cell survival and/or production of definitive HSCs.[30] Similarly, Rowel *et al.* created a 5 bp deletion in *rps19* zebrafish mutant using TALENs. Homozygous *rps19* mutants showed developmental anomalies and anemia, and were dead by 5 days post fertilization (dpf). However, *rps19* heterozygotes showed no difference to their wild-type siblings. Interestingly, exposure to cold stress during the first dpf resulted in a reduced number of RBCs.

To further investigate the biological functions of RPS7, Duan *et al.* used MO to knockdown rps7 in zebrafish.[39] In *rps7*-deficient embryos, *mdm2* and *tp53* were activated, inducing the expression of downstream target genes involved in p53 pathway (*bik*, *bax* and *puma*, *cdkn1a*, and *ccng1*). *rps7* morphants showed severe anemia with reduced expression of *gata1* and the mature erythroid marker αe3 at 24 hours post fertilization (hpf). A marked suppression of hemoglobin at 48 hpf was observed, indicating that the deficiency of Rps7 might cause abnormal proliferation and/or differentiation of erythroid progenitors. There were also severe defects (short body length, tissue necrosis, and curved tail). Furthermore, simultaneous knockdown of the *tp53* by co-injecting a *tp53* MO resulted in partial rescue of morphological abnormalities. The lower levels of *gata1* and *α-E1 globin* were partially rescued in the co-injected embryos, even though *tp53*, *cdk1a*, and *mdm2* were still up-regulated.[39]

The contribution of tp53 to the pathological development of bone marrow failure syndromes may be tissue- and mutation-specific. Antunes *et al.* studied the effect of different *rps7* and *rpl11* mutations in zebrafish. *rps7* mutant showed a stronger phenotype due to less maternal contribution of *rps7* comparing to *rpl11* mutant. Both mutants had severe anemia, morphological abnormalities, and increased apoptosis. Injection of p53 MOs rescued the apoptosis and the morphological phenotypes; however, it was unable to rescue anemia.[40] Taylor *et al.* showed that *rps29* mutants had defects in RBC development and increased apoptosis. Mutant embryos showed upregulation of *tp53* and *cdk1a* expression. Mutation of *tp53* in homozygous *rps29* mutant embryos reversed the apoptotic and hematologic phenotypes. However, mutated *tp53* did not fully rescue the embryonic lethality of *rps29* mutants, suggesting that *tp53*-independent mechanisms were affected by *rps29* knockdown.[41] Yadav *et al.* knocked down five ribosomal protein genes (two DBA-associated, *rpl35a* and *rps24*, and three non-DBA-associated, *rps3*, *rpl35* and *rplp1*), and analyzed these deficiencies on morphology and erythrocyte number in the presence and absence of p53 using MOs. They showed that any ribosomal protein deficiency led to anemia in zebrafish. Elimination of Tp53 function did not significantly affect the anemia, despite improving non-hematopoetic phenotypes.[42] DBA zebrafish models have helped identify MDM2-ribosomal protein interactions, which may interfere with MDM2 inhibition to p53 function. p53 rescue of severe anemia in ribosomal protein deficiency zebrafish models varies (Table 6).[31,35,43,44] Altogether, these findings suggest that there are p53-independent mechanisms

**Table 5.** Comparison between morphants and mutants.[22-24]

| | Morphants *versus* mutants | |
|---|---|---|
| | Morphant | Mutant |
| Effect | Knock down | Permanent changes in DNA |
| Affects | RNA transcripts | Genomic DNA |
| Phenotype | More severe maternal mRNA block by MO | Less severe maternal mRNA |
| Time to create | 1-3 days | 6-8 months |
| Side effects | More off-target effects | Less off target effects |
| Genetic compensation | No | Yes |
| p53 pathway | Affected | Non-affected |

**Table 6.** RP deficiency and p53 rescue in zebrafish models.

| RP | Severe anemia | Developmental malformations | Type of p53 rescue | p53 rescue of anemia | p53 rescue of other phenotypes | Ref |
|---|---|---|---|---|---|---|
| *rps19* morphant | Yes | Yes | p53 MO | No | Yes | 32 |
| *rpl11* morphant | N/A | Yes | p53 MO | N/A | Yes | 35 |
| *rps7* morphant | Yes | Yes | p53 MO | Partial | Partial | 40 |
| Rps7 mutant | Yes | Yes | p53 MO | No | Yes | 36 |
| Rps29 mutant | Yes | Yes | p53M214K | Yes | Yes | 93 |
| *rps24* & *rpl35a* morphants | Yes | Yes | p53M214K | No | Yes | 42 |

involved in bone marrow failure. One p53-independent effect may be translational dysfunction. Zebrafish can provide a model organism to identify Tp53-independent pathways that contribute to marrow failure mice or humans.

Zebrafish may also be a valuable model organism for drug development for DBA treatment. Several groups have tested the hypothesis that L-leucine and L-arginine can stimulate translation *via* the mTOR pathway and rescue affected DBA fish. Treatment of *rpl19* and *rpl14* zebrafish morphants with L-leucine improved developmental defects and hemoglobin levels.[45] Yadav *et al.* rescued the morphological defects of Rpl35a-deficient embryos and were able to improve erythroid cell number.[42] They concluded that translation deficit, not Tp53 activation, is the primary defect perturbing erythropoiesis.[42] While there have been anecdotal reports of leucine stimulation of erythropoiesis in DBA patients,[46] definitive clinical trial results are still pending. Another study found that RAP-011, an activin receptor ligand trap, partially restored erythropoiesis in *rpl11* morphants as well as *rpl11* and *rpl19* mutants.[47] Zebrafish also provided an *in vivo* model for further drug development of SMER28, a small molecule inducer of ATG5-dependent autophagy.[48] Given these results, we await clinical translation of SMER28 as a potential treatment for DBA.

## Dyskeratosis congenita

Dyskeratosis congenita is associated with abnormal skin pigmentation, nail dystrophy, and oral leukoplakia. DC patients may have other organ involvement, including the pulmonary, gastrointestinal, skeletal, neurological, immunological, and ophthalmological systems. Eighty-five percent of DC patients experience bone marrow failure, which accounts for much of the DC-related mortality. Other causes of mortality include infections, pulmonary complications, and hematologic and non-hematologic malignancies.[49-51]

Dyskeratosis congenita is a genetically heterogeneous disorder, showing autosomal recessive, autosomal dominant, and X-linked inheritance. So far, at least 21 mutated genes have been identified that can cause DC: *DKC1*, *TERC*, *TERT*, *NHP2*, *NOP10*, *CTC1*, *WRD79*, *TR*, *NOLA2*, *NOLA3*, *PARN*, *TPP1*, *POT1*, *CTC1*, *USB1*, *TCAB1*, *RTEL1*, *ACD*, *PARN*, *WRAP53* and *TINF2* (*http://telomerase.asu.edu/diseases.html*).[51-53] The X-linked *DKC1* has a more severe phenotype compared with the autosomal dominant forms. Although there is a broad consensus that DC results from stem cell renewal failure due to defective telomere maintenance, some mutated genes (e.g. *TERT*, *TERC*, and *DKC1*) are required for pre-rRNA processing.[2,49,50,54] How telomerase activity and impaired ribosomal biogenesis contribute to the pathophysiology of DC is still not known. Telomeres are complex DNA-protein structures at the end of chromosomes, and they shorten with each cell division. When telomeres become critically short, a DNA damage response is activated, causing cell cycle arrest or death. In humans, telomerase-based telomere elongation is the major mechanism that counteracts this process of continuous telomere shortening. In peripheral white blood cells, rapid telomere shortening occurs within the first year of life, followed by a more gradual decline over time.[49] Genetic diseases that cause telomerase deficiency are associated with premature aging and cancer susceptibility. As in humans, zebrafish chromosomes possess telomeres that progressively decline with age, reaching lengths in old age comparable to those observed when telomerase is mutated.[55] Several studies have helped to characterize its well-conserved molecular and cellular physiology. Different zebrafish mutants and morphants for telomere and telomerase research showed shorter lifespan, shorter telomeres, and different affected tissues (mainly brain, blood, gut and testes). These results make zebrafish an excellent model to unravel the connection between telomere shortening, tissue regeneration, aging and disease.[55,56]

Amsterdam *et al.* isolated the *nop10*$^{hi2578}$ mutant allele where a viral insertion within the first intron resulted in *nop10* decreased expression. This mutation is homozygously lethal by 5 dpf.[21] *nop10* encodes for a protein involved in 18S rRNA processing and is also part of the telomerase complex. Pereboom *et al.* observed that *nop10* loss in this mutant line resulted in a failure of the 18S rRNA to be properly processed, which led to the instability of the 40S ribosomal subunit. Due to the loss of 18S RNA, ribosomal proteins cannot be incorporated into a ribosome subunit and interact with other proteins, including the E3 ubiquitin ligase Mdm2. Mdm2 regulates Tp53 by promoting its ubiquitination and degradation. By binding to Mdm2, Rps7 enhances the E3 ubiquitin ligase activity of Mdm2 that promotes the degradation of Rps7. Furthermore, they observed that an increase in Tp53-specific apoptosis is coupled to the increased binding of Mdm2 to the Rps7. They observed that *nop10* mutants failed to form HSCs, a phenotype that is rescued by introducing a loss-of-function *tp53* mutation. However, they detected no changes in telomere length in *nop10* mutants.[53] They concluded that the cytopenia(s) of DC could be the result of ribosome biogenesis defects. This would lead to Tp53-mediated apoptosis of HSCs during early development, caused partially by the association of Rps7 with Mdm2.[53]

Two different approaches were used by Zhang *et al.* to study DC in zebrafish. First, MO-mediated knockdown was used to study the mechanisms whereby *dkc1* morphants result in HSC failure. Second, they performed retroviral-insertional mutagenesis of *nola1*. *NOLA1* encodes for GAR1, involved in rRNA maturation, and is also a key component telomerase complex. No mutations in *NOLA1* have been described in DC patients so far, but suspicion should be aroused in individuals with unexplained marrow failure or fibrosis. Both zebrafish models resulted in reduced hematopoiesis with reduction in *runx1* and *c-myb*, increased *tp53* expression, and defective ribosomal biogenesis without detectable changes in telomerase function. Their findings suggest that a telomerase-independent, Tp53-dependent mechanism contribute to hematopoietic failure in DC.[50]

Henriques *et al.* and Anchelin *et al.* studied the zebrafish telomerase reverse transcriptase *tert* mutant. These mutants develop normally for the first six months, but progressively develop tissue degeneration (gastrointestinal atrophy, loss of body mass, inflammation, a decrease in total blood cells and cell proliferation), and die prematurely. They also observed a Tp53-dependent response with increased transcripts of *puma*, *cdkn1a*, and *ccng1a*. Upregulation of cell cycle arrest inhibitors led to a G1 arrest and senescence. To study the effect of Tp53 in *tert* mutants, they created a double mutant *tert*$^{-/-}$, *tp53*$^{-/-}$ and

observed rescue of cell proliferation, which partially suppressed the degenerative phenotypes.[55,57] In another study, Kishi *et al.* studied the effect of ablation of *terfa*; they found multiple malformations mainly in brain, spinal cord, and eye.[58]

Recently, 2 patients with a phenotype overlapping with DBA and DC (pure red cell aplasia, hypogammaglobulinemia, growth retardation, and microcephaly) harbored a *de novo TP53* germline mutation that caused a C-terminal truncation in the last exon. This resulted in enhanced p53-mediated transcriptional activity. Using an MO that targets the 3′ splice site of intron 10, Toki *et al.* developed a zebrafish that displayed reduced number of erythrocytes, severe developmental defects, and died at 96 hpf.[7]

## Fanconi anemia

Fanconi anemia is mostly an autosomal recessive condition characterized by congenital abnormalities, progressive bone marrow failure, chromosome fragility, and an early onset of cancers such as myelodysplastic syndromes (MDS) /acute myeloid leukemia (AML) and epithelial malignancies. FA is characterized by non-hematologic phenotype, including short stature, microcephaly, microphthalmia, hypogonadism, and infertility. The mechanisms by which FA leads to developmental anomalies in blood, skeleton, eyes, and gonads are poorly understood; however, genotoxic stress by chemicals, mutagens, and viruses may contribute.[59,60]

Mutations in at least 20 genes can cause FA. However, since some cases of FA cannot be assigned to any of these genes, additional genes still have to be identified.[1,59] Proteins encoded by these genes constitute the FA pathway required for the efficient repair of damaged DNA. The FA core complex consists of at least 8 proteins: FANCA, FANCB, FANCC, FANCE, FANCF, FANCG FANCL, and FANCM. These proteins function as an E3 ligase and mediate the activation of the FANCD2 and FANCI (ID) complex. Once monoubiquitinated, the ID complex interacts with a third group of FANC proteins, including BRCA2 (FANCD1), FANCJ (BRIP1), FANCN (PALB2), FANCO (RAD51C), FANCP (SLX4), BRCA1, FAN1, histone H2AX, and RAD51, thereby contributing to DNA repair *via* homologous recombination.[1,59,61,62] Until now, 20 genes have been associated with causing FA: *FANCA*, *FANCB*, *FANCC*, *BRCA2*, *FANCD2*, *FANCE*, *FANCF*, *FANCG*, *FANCI*, *FANCJ*, *FANCL*, *FANCN*, *FANCP*, *FANCQ*, *RAD51*, *BRCA1*, *FANCT*, *FANCU*, *FANCV* and *FANCW*. Information about all these genes is available on the public Fanconi Anemia Mutation Database (*http://www.rockefeller.edu/fanconi/*).

Although zebrafish contain the full complement of FA family members found in humans,[63] loss-of-function models have been described for only a few. Liu *et al.* analyzed the zebrafish ortholog of the human *FANCD2* gene using MO.[64] They demonstrated developmental defects that arose during embryogenesis after *fancd2* knockdown, phenocopying the reduction in body length, and smaller head and eyes, which are frequently observed among FA patients. This suggests that the FA pathway plays a similar role in zebrafish and humans. They showed that the defects in *fancd2*-deficient embryos were the result of inappropriate and selective activation of Tp53-mediated apoptotic pathways in highly proliferative cells.[64]

Titus *et al.* characterized the developmental and tissue-specific expression of FA pathway genes in zebrafish.[63] They found maternal deposition of mRNA *fanc* genes can provide Fanc proteins to repair DNA damage encountered in rapid cleavage divisions. Zebrafish *fancl* mutants develop only as sterile males but without hematopoietic defects. The sex reversal was due to abnormal increase of germ cell apoptosis that compromises survival of developing oocytes and masculinizes the gonads. Interestingly, when the *tp53* mutation was introduced, the sex reversal phenotype could be rescued.[65] Botthoff *et al.* created a *rad51* knockout zebrafish mutant. In this model, zebrafish lacking rad51 survived to adulthood, but they were all infertile males with fewer HSPCs in the kidney. In earlier stages (2 and 4 dpf), they found that *rad51*$^{-/-}$ embryos also had a lower number, increased apoptosis, and reduced proliferation of HSPCs compared with their wild-type siblings. To study the role of p53 in the *rad51* mutants, they generated a zebrafish with mutations in both genes. After four months post fertilization, HSPCs were the same in wild-type and double mutants. The sex reversal was also corrected, but neither females nor male double mutants were fertile.[66]

## Shwachman-Diamond syndrome

Shwachman-Diamond syndrome is an autosomal recessive disorder characterized by exocrine pancreatic insufficiency, bone marrow dysfunction, and skeletal abnormalities. Hematologic abnormalities are a major cause of morbidity and mortality, and include cytopenia(s), MDS, and AML. Neutropenia occurs in approximately 90% of patients and occurs as early as the neonatal period. Skeletal abnormalities, such as metaphyseal chondrodysplasia, thoracic dystrophy, and short stature are common in SDS. In 2003, mutations in the Shwachman–Bodian-Diamond syndrome (*SBDS*) gene were identified.[70] In approximately 90% of cases, SDS is caused by two common mutations in exon 2 of *SBDS*: 183-184TA→CT introduces an in-frame stop codon (K62X) and 258+2T>C (C84Cfs) disrupts the donor splice site of intron 2, allowing a hypomorph to be produced.[67] Fifty percent of cases are compound heterozygotes with respect to these two mutations. Boocock *et al.* found that both changes correspond to sequences that occur normally in the pseudogene. Both mutations can also occur in the same allele.[67] Studies have identified additional changes in the coding sequence of *SBDS* that led to frameshift and missense mutations.

In 2007, Menne *et al.* characterized the function of the yeast SBDS ortholog Sdo1 in 60S maturation and translational activation of ribosomes.[68] SBDS is a protein with a well-documented role in the later steps of ribosome biogenesis. SBDS interacts with the GTPase EFL1 to trigger release of eIF6 from the 60S ribosomal subunit. EIF6 is critical for biogenesis and nuclear export of pre-60S subunits and prevents ribosomal subunit association. Removal of eIF6 is a prerequisite for the association of the 60S with the 40S subunit, and thus for the formation of an actively functioning ribosome.[4] Recently, mutations in *DNAJC21*[69,70] and *EFL1*[71] have been identified in individuals with SDS-like conditions. All of the SDS-associated mutant genes affect ribosome maturation. These important discoveries advance the concept of SDS as a ribosomopathy, and beg the question as to how ribosomopathies like DBA, SDS, or del (5q) can result in different defects in hematopoietic and non-hematopoietic tissues.

There have been no reports of homozygosity for *SBDS* null alleles, suggesting that human *SBDS* is essential and

that SDS patients carry at least one hypomorphic *SBDS* allele.[67,72-75] This is consistent with the finding that mice homozygous for null alleles of *sbds* exhibit early embryonic lethality, indicating that SBDS function is an essential for life.[30] While conditional knock-outs for *sbds* have been made, this approach is limited, costly, and time-consuming to generate. Thus, we and others have turned to the zebrafish also to study SDS. Venkatasubramani and Mayer used MO to knockdown sbds in zebrafish embryos, and study the effect in pancreas and myeloid development (Table 4). They observed an alteration in the spatial relationship between endocrine and exocrine pancreas. They also documented abnormal neutrophil distribution in the knockdown zebrafish model.[76] In a subsequent study, also using MO, Provost *et al.* observed that their model fully recapitulated the spectrum of developmental abnormalities observed in SDS patients: loss of neutrophils, skeletal defects, and pancreatic hypoplasia, as well as changes in the ribosomal subunit ratio. In this case, loss of Tp53 did not rescue the developmental defects associated with loss of sbds in zebrafish morphants.[77] Our recent work showed that sbds mutants obtained by CRISPR/Cas9 editing phenocopied SDS and displayed neutropenia, growth retardation, and atrophy of the pancreas.[78]

### Congenital amegakaryocytic thrombocytopenia

Congenital amegakaryocytic thrombocytopenia is a rare autosomal recessive condition characterized by thrombocytopenia, absence of megakaryocytes, and occasional evolution to aplastic anemia or leukemia.[79,80] Mutations in *MPL* have been described as the cause of CAMT.[81] *MPL* gene encodes for myeloproliferative leukemia protein (CD110), the receptor for thrombopoietin. Mice with genetic ablation of *Mpl* showed normal development but a deficiency in megakaryocytes and severe thrombocytopenia.[82] In zebrafish, disruption of *mpl* caused a severe reduction in thrombocytes (platelet equivalents), bleeding, and a decrease in HSCs. By phenocopying the human disease, affected zebrafish provide an accurate model to study this disease and for drug screening.[83] Reduction in HSCs and repopulation defects in affected zebrafish demonstrate that c-Mpl function in hematopoiesis is highly conserved. Moreover, the partial rescue of thrombocyte number by IL-11 provides a model to finely dissect JAK/STAT signaling in thrombopoiesis.

### Severe congenital neutropenia

Severe congenital neutropenia is a group of heterogeneous genetic disorders characterized by a maturation arrest at the promyelocyte stage of granulopoiesis and a high propensity to develop MDS/AML.[84] Over the past eighteen years, the following mutations have been identified as causing SCN: *ELANE, GFI1, HAX1, VPS45, JAGN, CSF3R,* and *WAS*. *ELANE* is the most commonly mutated gene in SCN, but there is no zebrafish ortholog. However, zebrafish has proven to be a powerful model to validate and characterize the function of newly described gene candidates for SCN. Vacuolar Protein Sorting 45 Homolog (*VPS45*) encodes a protein associated with protein trafficking into distinct organelles. Biallelic mutations in this gene are the cause of SCN5. A zebrafish model of *vps45* knockdown also showed a large decrease in neutrophils.[85] Mutations in CSF3R cause SCN7.[86] Pazhakh *et al.* mutated *csf3r* in zebrafish to study the effect on neutrophil production. They found that csf3r zebrafish mutants survive until adulthood with a 50% reduction in neutrophils and a substantial reduction in myeloid cells in the kidney marrow.[87] Recently, SRP54 mutations have been identified as the second most common cause of SCN (with some features of SDS).[88,89] Knockdown of SRP54 in zebrafish recapitulated the human phenotype of neutropenia, chemotaxis defect, and pancreatic exocrine insufficiency.[88]

## Conclusions

Despite the identification of specific gene mutations and pathway involvement for the great majority of patients with IBMFS, little is known about how they result in single or multiple lineage cytopenias. Furthermore, very little is known about co-operating mutations that effect transformation to MDS, AML, or solid tumors. Patient-based studies are problematic owing to the rarity of these disorders and to the long latency before bone marrow failure or malignancy. Zebrafish provide a relatively inexpensive, rapidly developing, vertebrate model organism. Despite some differences in their respective hematopoietic organs, mutations or silencing of relevant zebrafish genes phenocopies human IBMFS. Studies on gene mutations or suppression in zebrafish have validated the role of ribosome biogenesis, and advanced the hypothesis that the TP53 pathway plays a major role in the pathophysiology of some of the IBMFS. Zebrafish modeling may also contribute to drug development, as suggested by studies on L-leucine and SMER28 for DBA.

***Acknowledgments***

*SJC is supported by funding from NIH R01 HL128173, NIH R21 CA159203, Department of Defense Bone Marrow Failure Idea Development Award BM140102, Shwachman-Diamond Syndrome Foundation, Connor's Heroes, and the CURE Childhood Cancer Foundation. Due to space restrictions, the authors deeply regret not being able to cite all of our colleagues' publications.*

## References

1. Rosenberg PS, Alter BP, Ebell W. Cancer risks in Fanconi anemia: findings from the German Fanconi Anemia Registry. Haematologica. 2008;93(4):511-517.
2. Alter BP, Giri N, Savage SA, Rosenberg PS. Cancer in dyskeratosis congenita. Blood. 2009;113(26):6549-6557.
3. Da Costa L, O'Donohue MF, van Dooijeweert B, et al. Molecular approaches to diagnose Diamond-Blackfan anemia: The EuroDBA experience. Eur J Med Genet. 2017 Oct 26. [Epub ahead of print]
4. Finch AJ, Hilcenko C, Basse N, et al. Uncoupling of GTP hydrolysis from eIF6 release on the ribosome causes Shwachman-Diamond syndrome. Genes Dev. 2011;25(9):917-929.
5. Carrillo J, Gonzalez A, Manguan-Garcia C, Pintado-Berninches L, Perona R. p53 pathway activation by telomere attrition in X-DC primary fibroblasts occurs in the absence of ribosome biogenesis failure and as a consequence of DNA damage. Clin Transl Oncol. 2014;16(6):529-538.
6. Ceccaldi R, Parmar K, Mouly E, et al. Bone marrow failure in Fanconi anemia is triggered by an exacerbated p53/p21 DNA damage response that impairs hematopoietic stem and progenitor cells. Cell Stem Cell. 2012;11(1):36-49.

7. Toki T, Yoshida K, Wang R, et al. De Novo Mutations Activating Germline TP53 in an Inherited Bone-Marrow-Failure Syndrome. Am J Hum Genet 2018;103(3):440-447.
8. Glaubach T, Minella AC, Corey SJ. Cellular stress pathways in pediatric bone marrow failure syndromes: many roads lead to neutropenia. Pediatr Res. 2014;75(1-2):189-195.
9. Garaycoechea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. Nature. 2012;489(7417):571-575.
10. Howe K, Clark MD, Torroja CF, et al. The zebrafish reference genome sequence and its relationship to the human genome. Nature. 2013;496(7446):498-503.
11. Veldman MB, Lin S. Zebrafish as a Developmental Model Organism for Pediatric Research. Pediatr Res. 2008;64(5):470-476.
12. Teittinen KJ, Grönroos T, Parikka M, Rämet M, Lohi O. The zebrafish as a tool in leukemia research. Leuk Res. 2012;36(9): 1082-1088.
13. Davidson AJ, Zon LI. The 'definitive' (and 'primitive') guide to zebrafish hematopoiesis. Oncogene. 2004;23(43): 7233-7246.
14. de Jong JLO, Zon LI. Use of the Zebrafish System to Study Primitive and Definitive Hematopoiesis. Annu Rev Genet. 2005;39(1):481-501.
15. Palis J. Primitive and definitive erythropoiesis in mammals. Front Physiol. 2014;5:3.
16. Carradice D, Lieschke GJ. Zebrafish in hematology: sushi or science? Blood. 2008;111(7):3331-3342.
17. Chen AT, Zon LI. Zebrafish blood stem cells. J Cell Biochem. 2009;108(1):35-42.
18. Ellett F, Lieschke GJ. Zebrafish as a model for vertebrate hematopoiesis. Curr Opin Pharmacol. 2010;10(5):563-570.
19. Bowman TV, Zon LI. Lessons from the Niche for Generation and Expansion of Hematopoietic Stem Cells. Drug Discov Today Ther Strateg 2009;6(4):135-140.
20. Solnica-Krezel L, Schier AF, Driever W. Efficient Recovery of Enu-Induced Mutations from the Zebrafish Germline. Genetics. 1994;136(4):1401-1420.
21. Amsterdam A, Burgess S, Golling G, et al. A large-scale insertional mutagenesis screen in zebrafish. Genes Dev. 1999;13(20):2713-2724.
22. Stainier DYR, Raz E, Lawson ND, et al. Guidelines for morpholino use in zebrafish. PLoS Genet. 2017;13(10):e1007000.
23. Robu ME, Larson JD, Nasevicius A, et al. p53 Activation by Knockdown Technologies. PLoS Genet. 2007;3(5):e78.
24. Kok FO, Shin M, Ni CW, et al. Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. Dev Cell. 2015;32(1):97-108.
25. Rossi A, Kontarakis Z, Gerri C, et al. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature. 2015;524(7564):230-233.
26. Chiabrando D, Tolosano E. Diamond Blackfan Anemia at the crossroad between ribosome biogenesis and heme metabolism. Adv Hematol. 2010;2010:790632.
27. Giri N, Kang E, Tisdale JF, et al. Clinical and laboratory evidence for a trilineage haematopoietic defect in patients with refractory Diamond-Blackfan anaemia. Br J Haematol. 2000;108(1):167-175.
28. Landowski M, O'Donohue MF, Buros C, et al. Novel deletion of RPL15 identified by array-comparative genomic hybridization in Diamond-Blackfan anemia. Hum Genet. 2013;132(11):1265-1274.
29. Da Costa L, Narla A, Mohandas N. An update on the pathogenesis and diagnosis of Diamond–Blackfan anemia. F1000Res. 2018;7.
30. Zhang Y, Ear J, Yang Z, Morimoto K, Zhang B, Lin S. Defects of protein production in erythroid cells revealed in a zebrafish Diamond–Blackfan anemia model for mutation in RPS19. Cell Death Dis. 2014;5(7):e1352.
31. Danilova N, Sakamoto KM, Lin S. Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. Blood. 2008;112(13):5228-5237.
32. Uechi T, Nakajima Y, Chakraborty A, Torihara H, Higa S, Kenmochi N. Deficiency of ribosomal protein S19 during early embryogenesis leads to reduction of erythrocytes in a zebrafish model of Diamond-Blackfan anemia. Hum Mol Genet. 2008;17(20):3204-3211.
33. Danilova N, Bibikova E, Covey TM, et al. The role of the DNA damage response in zebrafish and cellular models of Diamond Blackfan anemia. Dis Model Mech. 2014;7(7):895-905.
34. Torihara H, Uechi T, Chakraborty A, Shinya M, Sakai N, Kenmochi N. Erythropoiesis failure due to RPS19 deficiency is independent of an activated Tp53 response in a zebrafish model of Diamond-Blackfan anaemia. Br J Haematol. 2011;152(5):648-654.
35. Chakraborty A, Uechi T, Higa S, Torihara H, Kenmochi N. Loss of Ribosomal Protein L11 Affects Zebrafish Embryonic Development through a p53-Dependent Apoptotic Response. PLoS One. 2009;4(1):e4152.
36. Amsterdam A, Nissen RM, Sun Z, Swindell EC, Farrington S, Hopkins N. Identification of 315 genes essential for early zebrafish development. Proc Natl Acad Sci USA. 2004;101(35):12792-12797.
37. Danilova N, Sakamoto KM, Lin S. Ribosomal protein L11 mutation in zebrafish leads to haematopoietic and metabolic defects. Br J Haematol. 2011;152(2): 217-228.
38. Danilova N, Gazda HT. Ribosomopathies: how a common root can cause a tree of pathologies. Dis Model Mech. 2015;8 (9):1013-1026.
39. Duan J, Ba Q, Wang Z, et al. Knockdown of ribosomal protein S7 causes developmental abnormalities via p53 dependent and independent pathways in zebrafish. Int J Biochem Cell Biol. 2011;43(8):1218-1227.
40. Antunes AT, Goos YJ, Pereboom TC, et al. Ribosomal Protein Mutations Result in Constitutive p53 Protein Degradation through Impairment of the AKT Pathway. PLoS Genet. 2015;11(7):e1005326.
41. Taylor AM, Humphries JM, White RM, Murphey RD, Burns CE, Zon LI. Hematopoietic defects in rps29 mutant zebrafish depend upon p53 activation. Exp Hematol. 2012;40(3):228-237.e5.
42. Yadav GV, Chakraborty A, Uechi T, Kenmochi N. Ribosomal protein deficiency causes Tp53-independent erythropoiesis failure in zebrafish. Int J Biochem Cell Biol. 2014;49:1-7.
43. Amsterdam A, Sadler KC, Lai K, et al. Many Ribosomal Protein Genes Are Cancer Genes in Zebrafish. PLoS Biol. 2004;2(5):e139.
44. Zhang Y, Lu H. Signaling to p53: ribosomal proteins find their way. Cancer Cell. 2009;16(5):369-377.
45. Payne EM, Virgilio M, Narla A, et al. L-Leucine improves the anemia and developmental defects associated with Diamond-Blackfan anemia and del(5q) MDS by activating the mTOR pathway. Blood. 2012;120(11):2214-2224.
46. Pospisilova D, Cmejlova J, Hak J, Adam T, Cmejla R. Successful treatment of a Diamond-Blackfan anemia patient with amino acid leucine. Haematologica. 2007;92(5):e66-e67.
47. Ear J, Huang H, Wilson T, et al. RAP-011 improves erythropoiesis in zebrafish model of Diamond-Blackfan anemia through antagonizing lefty1. Blood. 2015;126(7):880-890.
48. Doulatov S, Vo LT, Macari ER, et al. Drug discovery for Diamond-Blackfan anemia using reprogrammed hematopoietic progenitors. Sci Transl Med. 2017;9(376).
49. Du HY, Pumbo E, Ivanovich J, et al. TERC and TERT gene mutations in patients with bone marrow failure and the significance of telomere length measurements. Blood. 2008;113(2):309-316.
50. Zhang Y, Morimoto K, Danilova N, Zhang B, Lin S. Zebrafish Models for Dyskeratosis Congenita Reveal Critical Roles of p53 Activation Contributing to Hematopoietic Defects through RNA Processing. PLoS One. 2012;7(1):e30188.
51. Ruggero D, Shimamura A. Marrow failure: a window into ribosome biology. Blood. 2014;124(18):2784-2792.
52. Ballew BJ, Yeager M, Jacobs K, et al. Germline mutations of regulator of telomere elongation helicase 1, RTEL1, in Dyskeratosis congenita. Hum Genet. 2013;132(4):473-480.
53. Pereboom TC, van Weele LJ, Bondt A, MacInnes AW. A zebrafish model of dyskeratosis congenita reveals hematopoietic stem cell formation failure resulting from ribosomal protein-mediated p53 stabilization. Blood. 2011;118(20):5458-5465.
54. Freed EF, Bleichert F, Dutca LM, Baserga SJ. When ribosomes go bad: diseases of ribosome biogenesis. Mol Biosyst. 2010;6(3): 481-493.
55. Henriques CM, Carneiro MC, Tenente IM, Jacinto A, Ferreira MG. Telomerase Is Required for Zebrafish Lifespan. PLoS Genet. 2013;9(1):e1003214.
56. Carneiro MC, de Castro IP, Ferreira MG. Telomeres in aging and disease: lessons from zebrafish. Dis Model Mech. 2016;9(7):737-748.
57. Anchelin M, Alcaraz-Perez F, Martinez CM, Bernabe-Garcia M, Mulero V, Cayuela ML. Premature aging in telomerase-deficient zebrafish. Dis Model Mech. 2013;6(5):1101-1112.
58. Kishi S, Bayliss PE, Uchiyama J, et al. The Identification of Zebrafish Mutants Showing Alterations in Senescence-Associated Biomarkers. PLoS Genet. 2008;4(8):e1000152.
59. Rodríguez-Marí A, Postlethwait JH. The Role of Fanconi Anemia/BRCA Genes in Zebrafish Sex Determination. The Zebrafish: Disease Models and Chemical Screens: Elsevier BV; 2011:461-490.
60. Soulier J. Detection of somatic mosaicism and classification of Fanconi anemia patients by analysis of the FA/BRCA pathway. Blood. 2004;105(3):1329-1336.
61. Geiselhart A, Lier A, Walter D, Milsom MD. Disrupted Signaling through the Fanconi Anemia Pathway Leads to Dysfunctional

Hematopoietic Stem Cell Biology: Underlying Mechanisms and Potential Therapeutic Strategies. Anemia. 2012; 2012:1-18.
62. Dong H, Nebert DW, Bruford EA, Thompson DC, Joenje H, Vasiliou V. Update of the human and mouse Fanconi anemia genes. Hum Genomics. 2015;9:32.
63. Titus TA, Selvig DR, Qin B, et al. The Fanconi anemia gene network is conserved from zebrafish to human. Gene. 2006;371 (2):211-223.
64. Liu TX, Howlett NG, Deng M, et al. Knockdown of zebrafish Fancd2 causes developmental abnormalities via p53-dependent apoptosis. Dev Cell. 2003;5(6):903-914.
65. Rodríguez-Marí A, Cañestro C, BreMiller RA, et al. Sex Reversal in Zebrafish fancl Mutants Is Caused by Tp53-Mediated Germ Cell Apoptosis. PLoS Genet. 2010;6(7): e1001034.
66. Botthof JG, Bielczyk-Maczynska E, Ferreira L, Cvejic A. Loss of the homologous recombination gene rad51 leads to Fanconi anemia-like symptoms in zebrafish. Proc Natl Acad Sci USA. 2017;114(22):E4452-e4461.
67. Boocock GR, Morrison JA, Popovic M, et al. Mutations in SBDS are associated with Shwachman-Diamond syndrome. Nat Genet. 2003;33(1):97-101.
68. Menne TF, Goyenechea B, Sánchez-Puig N, et al. The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. Nature Genet. 2007;39(4):486-495.
69. Dhanraj S, Matveev A, Li H, et al. Biallelic mutations in DNAJC21 cause Shwachman-Diamond syndrome. Blood. 2017;129(11): 1557-1562.
70. Tummala H, Walne AJ, Williams M, et al. DNAJC21 Mutations Link a Cancer-Prone Bone Marrow Failure Syndrome to Corruption in 60S Ribosome Subunit Maturation. Am J Hum Genet. 2016;99 (1):115-124.
71. Stepensky P, Chacon-Flores M, Kim KH, et al. Mutations in EFL1, an SBDS partner, are associated with infantile pancytopenia, exocrine pancreatic insufficiency and skeletal anomalies in aShwachman-Diamond like syndrome. J Med Genet. 2017;54(8):558-566.
72. Erdos M, Alapi K, Balogh I, et al. Severe Shwachman-Diamond syndrome phenotype caused by compound heterozygous missense mutations in the SBDS gene. Exp Hematol. 2006;34(11):1517-1521.
73. Kuijpers TW, Alders M, Tool AT, Mellink C, Roos D, Hennekam RC. Hematologic abnormalities in Shwachman Diamond syndrome: lack of genotype-phenotype relationship. Blood. 2005;106(1):356-361.
74. Makitie O, Ellis L, Durie PR, et al. Skeletal phenotype in patients with Shwachman-Diamond syndrome and mutations in SBDS. Clin Genet. 2004;65(2):101-112.
75. Dror Y, Donadieu J, Koglmeier J, et al. Draft consensus guidelines for diagnosis and treatment of Shwachman-Diamond syndrome. Ann N Y Acad Sci. 2011;1242:40-55.
76. Venkatasubramani N, Mayer AN. A Zebrafish Model for the Shwachman-Diamond Syndrome (SDS). Pediatr Res. 2008;63(4):348-352.
77. Provost E, Wehner KA, Zhong X, et al. Ribosomal biogenesis genes play an essential and p53-independent role in zebrafish pancreas development. Development. 2012;139(17):3232-3241.
78. Oyarbide U, Kell MJ, Farinas J, Topczewski J, Corey S. Gene disruption of zebrafish Sbds phenocopies human Shwachman-Diamond Syndrome but suggests more global and lineage defects. Blood. 2016;128(22):336.
79. Ballmaier M. c-mpl mutations are the cause of congenital amegakaryocytic thrombocytopenia. Blood. 2001;97(1):139-146.
80. Rose MJ, Nicol KK, Skeens MA, Gross TG, Kerlin BA. Congenital amegakaryocytic thrombocytopenia: the diagnostic importance of combining pathology with molecular genetics. Pediatr Blood Cancer. 2008;50(6):1263-1265.
81. Ihara K, Ishii E, Eguchi M, et al. Identification of mutations in the c-mpl gene in congenital amegakaryocytic thrombocytopenia. Proc Natl Acad Sci USA. 1999;96(6):3132-3136.
82. Alexander WS, Roberts AW, Nicola NA, Li R, Metcalf D. Deficiencies in progenitor cells of multiple hematopoietic lineages and defective megakaryocytopoiesis in mice lacking the thrombopoietic receptor c-Mpl. Blood. 1996;87(6):2162-2170.
83. Lin Q, Zhang Y, Zhou R, et al. Establishment of a congenital amegakaryocytic thrombocytopenia model and a thrombocyte-specific reporter line in zebrafish. Leukemia. 2017;31(5):1206-1216.
84. Skokowa J, Germeshausen M, Zeidler C, Welte K. Severe congenital neutropenia: inheritance and pathophysiology. Curr Opin Hematol. 2007;14(1):22-28.
85. Vilboux T, Lev A, Malicdan MCV, et al. A Congenital Neutrophil Defect Syndrome Associated with Mutations in VPS45. N Engl J Med. 2013;369(1):54-65.
86. Ward AC, van Aesch YM, Gits J, et al. Novel point mutation in the extracellular domain of the granulocyte colony-stimulating factor (G-CSF) receptor in a case of severe congenital neutropenia hyporesponsive to G-CSF treatment. J Exp Med. 1999;190(4):497-507.
87. Pazhakh V, Clark S, Keightley MC, Lieschke GJ. A GCSFR/CSF3R zebrafish mutant models the persistent basal neutrophil deficiency of severe congenital neutropenia. Sci Rep. 2017;7:44455.
88. Carapito R, Konantz M, Paillard C, et al. Mutations in signal recognition particle SRP54 cause syndromic neutropenia with Shwachman-Diamond-like features. J Clin Invest. 2017;127(11):4090-4103.
89. Bellanné-Chantelot C, Schmaltz-Panneau B, Marty C, et al. Mutations in SRP54 gene cause severe congenital neutropenia as well as Shwachman-Diamond-like syndrome. Blood. 201;132(12):1318-1331.
90. Rowell J, Pietka G, Virgilio M, Pena O, Hockings C, Payne E. A Zebrafish Model of Diamond-Blackfan Anemia Results in Bone Marrow Failure and Demonstrates Defective Translation in Erythroid Cells By Ribosome Footprinting. Blood. 2017;130 (Suppl 1):871.
91. McGowan KA, Mason PJ. Animal models of Diamond Blackfan Anemia. Semin Hematol. 2011;48(2):106-116.
92. Morgado-Palacin L, Varetti G, Llanos S, Gomez-Lopez G, Martinez D, Serrano M. Partial Loss of Rpl11 in Adult Mice Recapitulates Diamond-Blackfan Anemia and Promotes Lymphomagenesis. Cell Rep. 2015;13(4):712-722.
93. Taylor AM, Humphries JM, White RM, Murphey RD, Burns CE, Zon LI. Hematopoietic defects in rps29 mutant zebrafish depend upon p53 activation. Exp Hematol. 2012;40(3):228-237.e5.
94. Watkins-Chow DE, Cooke J, Pidsley R, et al. Mutation of the diamond-blackfan anemia gene Rps7 in mouse results in morphological and neuroanatomical phenotypes. PLoS Genet. 2013;9(1):e1003094.
95. Uechi T, Nakajima Y, Nakao A, et al. Ribosomal protein gene knockdown causes developmental defects in zebrafish. PLoS One. 2006;1:e37.
96. Ruggero D, Grisendi S, Piazza F, et al. Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. Science. 2003;299(5604):259-262.
97. Gu BW, Bessler M, Mason PJ. A pathogenic dyskerin mutation impairs proliferation and activates a DNA damage response independent of telomere length in mice. Proc Natl Acad Sci USA. 2008;105(29):10173-10178.
98. Jaskelioff M, Muller FL, Paik J-H, et al. Telomerase reactivation reverses tissue degeneration in aged telomerase deficient mice. Nature. 2011;469(7328):102-106.
99. Parmar K, D'Andrea A, Niedernhofer LJ. Mouse models of Fanconi anemia. Mutat Res. 2009;668(1-2):133-140.
100. Rodriguez-Mari A, Postlethwait JH. The role of Fanconi anemia/BRCA genes in zebrafish sex determination. Methods Cell Biol. 2011;105:461-490.
101. Lim DS, Hasty P. A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. Mol Cell Biol. 1996;16(12):7133-7143.
102. Zhang S, Shi M, Hui CC, Rommens JM. Loss of the mouse ortholog of the Shwachman-Diamond syndrome gene (Sbds) results in early embryonic lethality. Mol Cell Biol. 2006;26(17):6656-6663.
103. Lim KH, Chang YC, Chiang YH, et al. Expression of CALR mutants causes mpl-dependent thrombocytosis in zebrafish. Blood Cancer J. 2016;6(10):e481.
104. Liu F, Kunter G, Krem MM, et al. Csf3r mutations in mice confer a strong clonal HSC advantage via activation of Stat5. J Clin Invest. 2008;118(3):946-955.

PLOS COMPUTATIONAL BIOLOGY

# Mutation, drift and selection in single-driver hematologic malignancy: Example of secondary myelodysplastic syndrome following treatment of inherited neutropenia

Tomasz Wojdyla, Hrishikesh Mehta, Taly Glaubach, Roberto Bertolusso, Marta Iwanaszko, Rosemary Braun, Seth J. Corey , Marek Kimmel

Published: January 7, 2019 • https://doi.org/10.1371/journal.pcbi.1006664

This is an uncorrected proof.

## Abstract

Cancer development is driven by series of events involving mutations, which may become fixed in a tumor via genetic drift and selection. This process usually includes a limited number of driver (advantageous) mutations and a greater number of passenger (neutral or mildly deleterious) mutations. We focus on a real-world leukemia model evolving on the background of a germline mutation. Severe congenital neutropenia (SCN) evolves to secondary myelodysplastic syndrome (sMDS) and/or secondary acute myeloid leukemia (sAML) in 30–40%. The majority of SCN cases are due to a germline *ELANE* mutation. Acquired mutations in *CSF3R* occur in >70% sMDS/sAML associated with SCN. Hypotheses underlying our model are: an *ELANE* mutation causes SCN; *CSF3R* mutations occur spontaneously at a low rate; in fetal life, hematopoietic stem and progenitor cells expands quickly, resulting in a high probability of several tens to several hundreds of cells with *CSF3R* truncation mutations; therapeutic granulocyte colony-stimulating factor (G-CSF) administration early in life exerts a strong selective pressure, providing mutants with a growth advantage. Applying population genetics theory, we propose a novel two-phase model of disease development from SCN to sMDS. In Phase 1, hematopoietic tissues expand and produce tens to hundreds of stem cells with the *CSF3R* truncation mutation. Phase 2 occurs postnatally through adult stages with bone marrow production of granulocyte precursors and positive selection of mutants due to chronic G-CSF therapy to reverse the severe neutropenia. We predict the existence of the pool of cells with the mutated truncated receptor *before* G-CSF treatment begins. The model does not require increase in mutation rate under G-CSF treatment and agrees with age distribution of sMDS onset and clinical sequencing data.

## Author summary

Cancer develops by multistep acquisition of mutations in a progenitor cell and its daughter cells. Severe congenital neutropenia (SCN) manifests itself through an inability to produce enough granulocytes to prevent infections. SCN commonly results from a germline *ELANE* mutation. Large doses of the blood growth factor granulocyte colony-stimulating factor (G-CSF) rescues granulocyte production. However, SCN frequently transforms to a myeloid malignancy, commonly associated with a somatic mutation in *CSF3R*, the gene encoding the G-CSF Receptor. We built a mathematical model of evolution for *CSF3R* mutation starting with bone marrow expansion at the fetal development stage and continuing with postnatal competition between normal and malignant bone marrow cells. We employ tools of probability theory such as multitype branching process and Moran models modified to account for expansion of hematopoiesis during human development. With realistic coefficients, we obtain agreement with the age range at which malignancy arises in patients. In addition, our model predicts the existence of a pool of cells with mutated *CSF3R* before G-CSF treatment begins. Our findings may be clinically applied to intervene more effectively and selectively in SCN patients.

**Citation:** Wojdyla T, Mehta H, Glaubach T, Bertolusso R, Iwanaszko M, Braun R, et al. (2019) Mutation, drift and selection in single-driver hematologic malignancy: Example of secondary myelodysplastic syndrome following treatment of inherited neutropenia. PLoS Comput Biol 15(1): e1006664. https://doi.org/10.1371/journal.pcbi.1006664

**Editor:** James Gallo, University at Buffalo - The State University of New York, UNITED STATES

**Received:** August 29, 2018; **Accepted:** November 19, 2018; **Published:** January 7, 2019

**Copyright:** © 2019 Wojdyla et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** TW and MK were supported by the grant DEC-2012/04/A/ST7/00353 from the National Science Center (Poland). MK, MI, RB, HM, and SJC were supported by NIH R01HL128173. HM was supported by a Leukemia Research Foundation award. TG was supported by an MDS International Foundation Young Investigator Award. SJC was supported by grants from

the NIH RO1CA108992, American Society for Hematology Bridge Award, Leukemia and Lymphoma Translational Award, and the Department of Defense Bone Marrow Failure IDEA grant. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

Cancer development is driven by series of mutational events, which may become fixed in a hematologic or non-hematologic tumor via genetic drift. This process usually includes a limited number of driver (advantageous) mutations, and a greater number of passenger (neutral or mildly deleterious) mutations. Driver mutations for several hundred different cancers have been identified by sequencing and functional assays. The relationship between driver and passenger mutations has been investigated using mathematical models representing carcinogenesis in terms of a "tug-of war" between the former and the latter [1, 2]. Another related problem is whether carcinogenesis is driven by acquisition of single point mutations or by saltatory changes amounting to major genome rearrangement events [3, 4]. Mathematical modeling of interactions among multiple drivers has been described by Nowak and Durrett and their colleagues [5–7]. These frequently involve branching processes and related mathematical models [8]. Among stochastic models in hematology, an example is [9]. Hematopoiesis provide the best-characterized system for cell fate decision-making in both health and disease [10], as well as connections between stimuli such as inflammation and cancer [11].

Here, we model a disease evolving on the background of a germline mutation. The acquired driver mutation recurs during tissue expansion phase in fetal life and becomes selectively advantageous in early childhood, leading to development of malignancy. As a prominent example of such disease, we model the important hematologic disorder Severe Congenital Neutropenia (SCN), a monogenic inherited disorder, that acquires new mutations and evolves to secondary myelodysplastic syndrome (sMDS) or secondary acute myeloid leukemia (sAML). This model is similar to the "fish" graph of Tomasetti and Vogelstein [12]; however the latter is more comprehensive and involves multiple driver case. Here, we use tools of population genetics and population dynamics to model progression from SCN to sMDS and dissect the contributions of mutation, drift and selection at different stages of an individual's life. More specifically, we consider:

1. In an individual primed by an inherited genotype, the driver mutation occurs recurrently in the embryonic expansion stage, although these mutations do not necessarily confer selective advantage.
2. At birth, due to environmental and behavioral factors or treatment, the driver mutation acquires a selective advantage in a tissue or organ, while the driver mutation may or may not recur as frequently any more.
3. The mutant driver variant increases in frequency due to selection, and eventually it dominates the stem cells of the tissue or the organ, contributing to development of disease.

Accordingly, SCN is most commonly due to germline mutations in *ELANE*, which encodes the neutrophil elastase [13]. SCN is characterized by the near absence of circulating neutrophils, which renders the child, typically an infant, susceptible to recurrent life-threatening infections. The introduction in the 1990s of recombinant granulocyte colony-stimulating factor (G-CSF) to increase circulating neutrophils, markedly improved survival and quality of life for SCN patients [14].

SCN often transforms into sMDS or sAML [15, 16]. Clinical studies have demonstrated a strong association between exposure to G-CSF and sMDS/AML [17–21]. Mutations in the distal domain of the Granulocyte Colony-Stimulating Factor Receptor (*CSF3R*) have been isolated from almost all SCN patients who developed sMDS/AML [22, 23]. Clonal evolution over approximately 20 years was documented using next generation sequencing and quantification of *CSF3R* allele frequency variation in an SCN patient who developed sMDS/sAML [24]. Strikingly, out of four different mutations in *CSF3R*, one persisted into the leukemic clone but the other three were lost, supporting the assumption of different selective values in the presence of G-CSF that underlies our model. As clonal evolution is a central feature in cancer [25–28] and next generation sequencing has revealed complexed genomic landscapes, SCN may provide a simpler real-world example to study cancer development.

Two opposing paradigms have been proposed for cell fate decision making in blood cells: stochastic hematopoiesis (based on variability observed in cultured bone marrow cells as first suggested by McCulloch and Till [29]) and deterministic, or instructive, hematopoiesis (growth factor-driven production of specific blood cell types) [30, 31]. In spite of substantial experimental findings, particularly recent single-cell measurements [32], the two opposing theories await a grand synthesis. Disease-accompanying dynamics have been variously modeled over the years as deterministic or stochastic[10]. SCN may also provide a simpler real-world example to study cell fate determination.

Little is known about the molecular mechanism(s) by which SCN leads to myeloid malignancy and how important are the truncating mutations such as *CSF3R D715* in this process. Notwithstanding the exact molecular mechanism by which the *CSF3R* truncation mutants lead to sMDS/AML, two pivotal questions concerning the population dynamics and population genetics of the mutant clones are: (i) whether the *CSF3R* truncation mutants are present before application of G-CSF, and (ii) whether G-CSF administration increases mutation rate in hematopoietic stem cells with *ELANE* mutation. If the answer to the first question is affirmative, then the presence of a small subpopulation of *CSF3R* truncation mutants among infants with SCN might be of prognostic value and a preventive therapy might be sought. Concerning the second question, determination whether G-CSF is mutagenic or provides a selective pressure may influence the degree G-CSF therapy is conducted, e.g. should it be more or less aggressive.

To provide insight into the course of this disease and its clinical management, we propose a novel model of the emergence and fixation of *CSF3R* truncating mutations, which also follows the paradigm outlined earlier on. We note that the most common of these mutations associated with transition to sMDS is the *CSF3R D715*. However, at the resolution level of our model, we are not able to make more specific distinctions nor to consider coexistence or competition of more than one truncation mutant. The model assumes that the answer to question (i) is affirmative, but it is negative to question (ii). The model's hypotheses are:

1. An inherited mutation in *ELANE* causes SCN;

2. *ELANE* mutations coexist with the non-mutated allele in an intracellular environment in which *CSF3R* truncating mutations occur spontaneously at a low rate;
3. During fetal life, the number of stem cells and committed cells expands quickly, which results in high probability of approximately $10^1 - 10^2$ *CSF3R* truncation mutant cells; these would be of no consequence had their number not expanded further;
4. Administration of pharmacologic G-CSF early in life exerts selective pressure, providing *CSF3R* mutants with selective advantage. This assumption is supported by experiments, in which growth differential of wild-type and truncation mutant depends on the G-CSF concentration (**Fig 1**).

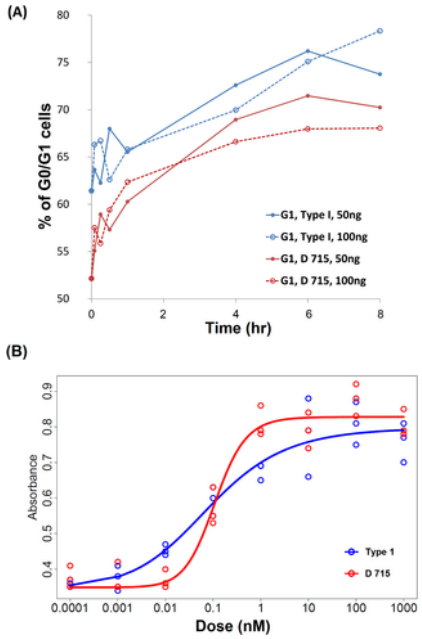



**Fig 1.**
**(A). Effect of CSF3R (wild type or D715) stimulation on cell cycle**. Ba/F3 cells expressing either wild type (Type I, blue) or mutant (D715, red) CSF3R were stimulated with two different doses of G-CSF (50 ng/ml and 100 ng/ml) and were analyzed using flow cytometry to determine the effect of G-CSF dose on receptor subtype activation on cell cycle progression. The cells were treated as detailed in the **Methods**. Flow cytometry data for each time point were analyzed using FlowJo employing the built-in cell cycle module. The data are represented as percentage of G1/G0 cells obtained from the FlowJo analysis plotted against time. Following an initial transient, from the 1 hr time point the G1/G0 percentage differences between cells with the Type I and D715 receptors have been on the average equal to 6.05% for 100 ng of G-CSF and 4.28% for 50 ng of G-CSF, respectively.
https://doi.org/10.1371/journal.pcbi.1006664.g001

1. Onset of sMDS is equivalent with replacement of normal hematopoietic bone marrow by mutant hematopoietic stem cells, their progeny, and release of factors that suppress normal hematopoiesis. At the level of resolution of our model, we are not able to make more specific distinctions.

**(B) Effect of CSF3R (wild type vs. D715) on cell proliferation**. Ba/F3 cells expressing either wild type (Type I) or mutant (D715) *CSF3R* were treated with increasing doses of recombinant human G-CSF (ng/ml) and proliferation was measured by the MTT assay performed in triplicates in a 96 well plate. The data are raw absorbance values at 600 nm and represent the three replicates plotted against increasing dose of G-CSF, fitted using least squares by Hill-type curves (Type I, blue, D715 red). For details of experimental procedures see the **S1 Appendix**. Fitting and statistical procedures are explained in the **Methods**.

We show that hypotheses 2 and 3 are needed, by first building a proof-of-principle simple Moran process (a stochastic model used in population genetics) with no expansion, which fits the data only if it is started by a cell population including $\sim 10^1 - 10^2$ cells expressing a *CSF3R* truncation mutation. Then we show that this number of *CSF3R* truncation mutant cells can be produced in the late fetal period expansion of hematopoiesis in bone marrow. Existence of the pool of *CSF3R* truncation mutant cells before exposure to G-CSF can be discovered only by deep targeted sequencing. Then we follow up with a full-fledged comprehensive model, which accounts for the important detail of time change of the size of the hematopoietic system, but which confirms the conclusions of the proof-of-principle model.

## Methods

Measuring the selective advantage of D 715 mutant cells

Experiment.

As shown in **Fig 1A**, the fraction of Ba/F3 cells expressing either wild type (Type I, blue) or mutant (D715, red) CSF3R that reside in the G1/G0 phase has been measured, employing Technical details are found in the **S1 Appendix**. Briefly, following release from the starvation cell cycle arrest, cells were stimulated with two different doses of G-CSF (50 ng/ml and 100 ng/ml) and were analyzed using flow cytometry to determine the effect of G-CSF dose on receptor subtype activation in terms of cell cycle progression. Flow cytometry data for each time point were analyzed using FlowJo software [33] employing the built-in cell cycle module. The data are represented as percentage of G1/G0 cells obtained from the FlowJo analysis plotted against time. After an initial transient, from the 1 hr time point the G1/G0 percentage differences between cells with the Type I and D715 receptors have been on the average equal to 6.05% for 100 ng of G-CSF and 4.28% for 50 ng of G-CSF, respectively. The data points are derived from flow cytometry, and because it measures thousands of cells, statistical fluctuations play a minor role and therefore the difference are statistically "highly significant".

G1/G0 cell fraction as a measure of selective advantage of cycling cells. We employ a simple cell cycle model to relate the G1/G0 cell fraction to the cell growth rate and selective advantage. Mathematical foundations can be found for example in ref. [34, 35]. Briefly, suppose that the interdivision time $T$ of cell in a population is constant and that the residence time $T_1$ in the G1/G0 is also constant (relaxing these assumptions is possible, as implicit from [34, 35], but it does not affect the first-order approximation we need here). Let us also suppose that the efficiency of divisions $\alpha$ (probability that a progeny cell enters the cell cycle) is constant. Under these hypotheses, the expected cell growth after initial transients becomes exponential with rate $\lambda = \ln(2\alpha)/T$, i.e. the cell count at time $t$ is equal to

$$N(t) = N(0)e^{\lambda t}$$

The fraction $\varphi$ of cells in G1/G0, tends to the following value

$$\varphi = \frac{1 - e^{-\lambda T_1}}{1 - e^{-\lambda T}}$$

This latter expression can be solved for the ratio $\psi = T_1/T$

$$\psi = \frac{\ln\{1 - \varphi[1 - (2\alpha)^{-1}]\}}{-\ln(2\alpha)}$$

Suppose now that we consider another ("asterisked") population of cycling cells, which has respective parameters denote with superscript $(^*)$.Let us also assume that the difference in cell division times between the two populations is due only to shortening of the G1/G0 phase

$$T - T^* = T_1 - T_1^* = \Delta$$

Let us note that in the experiment described earlier on, we measure the ratios $\varphi$ (G1/G0 fraction in Type I receptor cells) and $\varphi^*$ (G1/G0 fraction in D 715 receptor cells). Following some algebra, we obtain an estimate of $\Delta$ in the terms of $\varphi$ and $\varphi^*$ given we assume $T$ and $\alpha$ and $\alpha^*$

$$\Delta = \frac{(\psi - \psi^*)}{1 - \psi^*} T$$

where $\psi$ and $\psi^*$ have been already expressed in the terms of $\varphi$ and $\varphi^*$.

Selective coefficient in the Moran model is defined as $s = r - 1$, where $r$ is the ratio of probabilities that the replacent for a withdrawn (divided) self-renewing cell is a mutant (in this case an "asterisked" cell). This interpretation is consistent with the confined environment of the bone marrow, in which on the average one progeny cell dies of differentiates, leaving the self-renewing cell pool. This leads to expression

$$s = r - 1 = \frac{\lambda^* - \lambda}{\lambda}$$

This interpretation allows understanding the bounds on the selection coefficient provided by the experiment. We accepted $T = 4$ days, and we may start from assuming $\alpha = \alpha^* = 1$ (perfect division efficiency). Suppose that in approximate agreement with the experiment, we consider $\varphi = 0.7$ and $\varphi^* = 0.66$, which leads to $\psi = 0.621$ and $\psi^* = 0.567$, and to $\Delta = 0.503$. In turn, this provides $\lambda = 0.173\ \mathrm{d}^{-1}$ and $\lambda^* = 0.198\ \mathrm{d}^{-1}$ which leads to $s = 0.144$. Lowering equally the division efficiencies of both cell types changes the selection coefficient only slightly. However, if the "asterisked" (mutant) efficiency is lowered to $\alpha^* = 0.93$, the selective advantage of faster proliferating cells shrinks to almost 0. Consistently with this, we use the range of $s$-values from 0.002 to 0.1 (Table 1) and from 0.02 to 0.008 (**Fig 2**), which stay below the upper bound provided by $s = 0.144$.

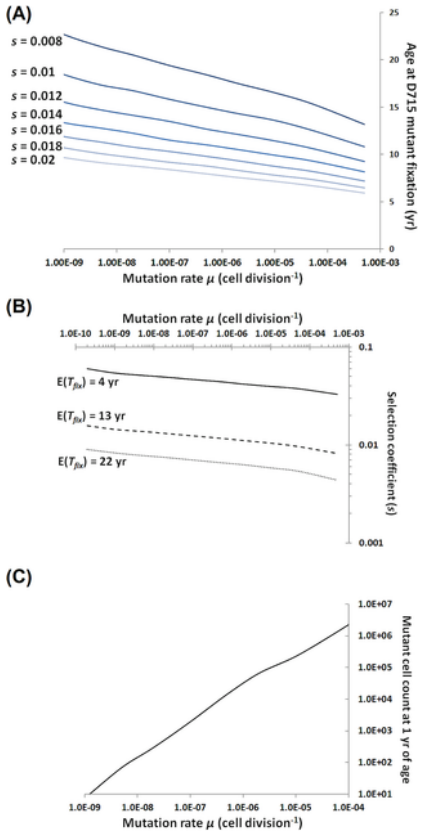



**Fig 2.**
**(A)** Age at which the d715 mutants replace the normal cells in the CMP compartment, as predicted by the modified Moran model with varying cell population size, for a range of mutation rate and selection coefficient values. **(B)** Selection coefficient values corresponding to the ages at replacement of 4 years old (solid line), 13 years old (dashed line), and 22 years old (dotted line) for a range of mutation rates values. **(C)** Mutant cell count at age of 1 year for a range of mutation rate values.
https://doi.org/10.1371/journal.pcbi.1006664.g002

| $s$ | $i$ | $\mu$ (cell division$^{-1}$) | $E[T_i \mid T_N < T_i]$ (yr.) |
|---|---|---|---|
| 0.005 | 240 | $2.01\times10^{-08}$ | 90.57 |
| 0.01 | 120 | $1.00\times10^{-08}$ | 46.05 |
| 0.02 | 60 | $4.99\times10^{-09}$ | 23.41 |
| 0.03 | 40 | $3.31\times10^{-09}$ | 15.75 |
| 0.04 | 29 | $2.47\times10^{-09}$ | 11.90 |
| 0.05 | 23 | $1.97\times10^{-09}$ | 9.57 |
| 0.1 | 11 | $9.57\times10^{-10}$ | 4.86 |

Results of the computations of the expected time to fixation of a *CSF3R* truncation mutation with probability of fixation kept at $P[T_N < T_i] = 0.7$, assuming the Moran process with directional selection. Notation: $s$, selection coefficient; $i$, mutant count at birth; $\mu$, mutation rate (cell division$^{-1}$); $E[T_i \mid T_N < T_i]$ is the age at fixation of the *CSF3R* truncation mutant (yr). For detailed hypotheses and derivations, see the present section.

https://doi.org/10.1371/journal.pcbi.1006664.t001

**Table 1. Results of the computations of the expected time to fixation.**
https://doi.org/10.1371/journal.pcbi.1006664.t001

Effect of CSF3R (wild type vs. D715) on cell proliferation.

Ba/F3 cells expressing either wild type (Type I) or mutant (D715) *CSF3R* were treated with increasing doses of G-CSF (ng/ml) and proliferation was measured using the MTT assay performed in triplicates in a 96 well plate. Full account of experimental procedures is found in the **S1 Appendix**. The data are raw absorbance values at 600 nm and represent the three replicates plotted against increasing dose of G-CSF.

Concerning recalculation from the pharmacological dose to serum concentration of the G-CSF, ref.[36] is providing relevant information. In this publication, see their **Fig 1**, therapeutic doses from $5 - 15\ \mu g/kg$ body weight resulted in serum concentrations of the order from $1 - 100\ ng/ml$, with a maximum for the lowest dose of $5\ \mu g/kg$ reached after 4 hours and equal to over $10\ ng/ml$, and remaining over $1\ ng/ml$ for 10 days.

To better understand the data, we performed parametric least-square fitting of the Type 1 and D715 data using Hill-like sigmoid curves, which results in clarified visualization of trends with the two curves crossing approximately at G-CSF concentration of $0.1ng$. Hill equation is a prototypical sigmoidal curve widely used in systems biology [37]. It has the form

$$y = y_0 + a\frac{(\ln x - c)^n}{(\ln x - c)^n + b^n}$$

The best least-square fit to the data (separately for Type I and D715 data). Values of estimated coefficients: Common for Type I and D715, $y_0 = 0.349$; Type 1, $a = 0.449, b = 20.802, c = -23.309, n = 11.149$; D715, $a = 0.479, b = 21.075, c = -23.278, n = 28.255$.

In addition, we carried out rigorous testing using one-sided Wilcoxon two-sample rank test of the difference between the Type 1 and D715 data points separately for concentrations $> 0.1$ ng/ml (resulting in highly significant difference at $p = 0.0034$), and for concentrations $\leq 0.1$ng/ml (resulting in borderline significant difference in opposite direction at $p = 0.089$). This testing justifies the assertion of higher growth rate of D715 cell in the range of G-CSF concentrations above $0.1$ng/ml.

Proof-of-principle model

Moran process with directional selection.

For a simplified proof-of-principle model of competition between mutant and wild type cells in adult bone marrow, we use the standard Moran process with selection [38]. In this process (**Fig 3**), the population of granulocyte precursors is assumed to be constant, with the proportion of mutants varying in time and time running in discrete units (such as days or cell-division times). We consider a population of $N$ biological cells, which at time 0 contains $i$ mutant cells. The mutant has a selective advantage defined as the relative fitness $r$, which is frequently expressed as $r = 1 + s$, where $s$ is called the selection coefficient. For an advantageous mutant, $r > 1$, or $s > 0$. Further details are found in the corresponding section of the **S1 Appendix**.

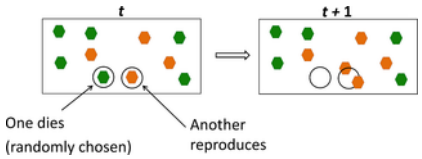



**Fig 3. Moran Process with discrete time and directional selection.**
One of the $N$ cells ($i$ mutants, in orange, and $N - i$ wildtype, in green), present at time $t$, dies (no preference for type). Its replacement at time $t + 1$ is selected from all cells present at $t$, with odds (relative fitness) *r* in favor of the mutant. Relative fitness is frequently expressed as $r = 1 + s$, where $s$ is called the selection coefficient. For an advantageous mutant, $r > 1$, or $s > 0$.
https://doi.org/10.1371/journal.pcbi.1006664.g003

Selective advantage in cycling cells.

CSF3R is a member of the hematopoietin/cytokine receptor family and functions as a homodimer. The cytoplasmic region consists of a proximal domain essential for proliferation and a distal domain critical for differentiation. Acquired CSF3R mutations have been observed to cluster between nucleotides 2384 to 2522 (residues 715 to 750), resulting in the loss of the distal domain [23]. Epidemiological studies demonstrated that the risk of sMDS/AML increased with the dose of G-CSF [16, 39].

Mutant clones may divide more frequently and/or be less apoptotic. In the simplest case of no cell death, selective advantage can be related to shortening of the interdivision times in mutant cells, relative to normal cells. A proxy for shorter interdivision time is a lower proportion of cells in the G1 phase, as this phase is usually most variable. **Fig 1A** presents a summary of the dynamics of cell cycle distribution of Ba/F3 cells expressing either the full-length wild type CSF3R or the CSF3R D715 mutant following their release from a starvation block. The fraction of cells in G1 in D715 mutants is lower by about 0.05 compared to that in the wild type CSF3R expressing cells, which translates into a growth rate advantage of the CSF3R D715 mutant. The difference is highly statistically significant, as flow cytometry measures thousands of cells per condition.

As an independent check, **Fig 1B** shows a direct comparison of growth curves of Ba/F3 cells expressing either the full-length CSF3R or the truncated CSF3R D715. The dose dependence shows that mutants have a selective advantage over a range of high G-CSF concentrations, whereas for low (normal) concentrations, they are at best neutral or possibly disadvantageous. Relevant laboratory techniques are found in **S1 Appendix**. Similar results were found depending on the particular CSF3R cytoplasmic mutant [40].

CSF3R truncation mutations at the fetal-life expansion phase of bone marrow.

Hematopoiesis in the human fetus moves from the liver into bone marrow about 90 days before the end of the pregnancy [41, 42]. The requirement of more than one mutant cell present at time *t* = 0 of the Moran process can be satisfied as follows. Suppose that *CSF3R* truncation mutations occur during the embryonic bone marrow expansion stage. In this time interval, because of the rapid expansion on the bone marrow, cell proliferation and mutation can be described using the time-continuous Markov branching process model [8], originally developed by Coldman and Goldie in a different context [43, 44]. The assumptions are as follows (**Fig 4**):

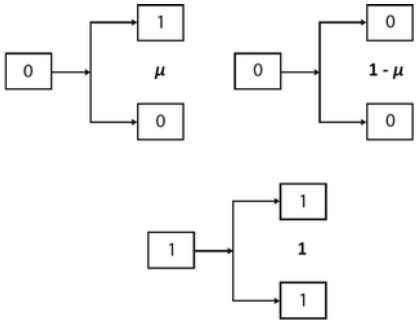



**Fig 4. The branching process of cell proliferation with irreversible mutation.**
Briefly, the cancer cell population is initiated by a single "wild type" (WT) cell, denoted as "0". At each division, with probability $\mu$, one of the WT progeny cells mutates. Mutants, denoted as "1" produce a pair of mutant progeny. Mutants have the same distribution of the interdivision time as the WT cells.
https://doi.org/10.1371/journal.pcbi.1006664.g004

1. The stem cell population (here, pooled Hematopoietic Stem Cells or HSC, and Common Myeloid Progenitors or CMP) is initiated by a small number $N_0$ of "wild type" (WT) cell that already acquired the *ELANE* mutation.
2. Interdivision time of WT cells is a random variable from an exponential distribution with parameter $\lambda$. Accordingly, mean interdivision time of WT cells is equal to $1/\lambda$.
3. At each division of a WT cell, with probability $\mu$, one of the progeny cells acquires the *CSF3R D715* mutation. *CSF3R* truncation mutants always produce *CSF3R* truncation mutants when dividing.
4. At the expansion phase, mutants are assumed selectively neutral (have the same parameter $\lambda$ as WT cells).

Mathematical details are found in the corresponding section of the **S1 Appendix**.

Comprehensive model of fixation of CSF3R truncation mutants

Modeling the age-dependent changes of sizes of the bone marrow cell compartments.

For more precise simulations, we build a model of dynamics of the hematopoietic stem cells (HSC) these latter defined as HSC and long term culture-initiating cells (LTC-IC) [45], and of the common myeloid progenitors (CMP), which give rise to neutrophils and/or monocytes. The model schematic is depicted in **Fig 5**. Dynamics of these cell populations are represented in the compartmental model by a system of three differential equations following the model of Arino and Kimmel [46]. The equations of the model are explained in the corresponding section of the **S1 Appendix** and in the legends to **Figs 5** and **6.**

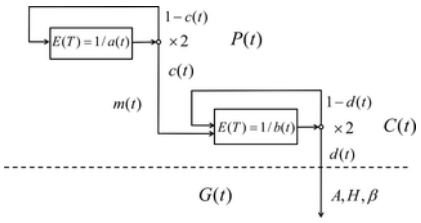



**Fig 5. Deterministic model of age-related dynamics of pluripotent and committed cells and granulocytes.**
$P(t)$, $C(t)$, and $G(t)$ are the numbers of pluripotent cells (HSC compartment), committed cells (CMP compartment), and peripheral granulocytes at time $t$, respectively; $m$ is the ratio of the committed cells in the hematopoietic cell lineage associated with the granulocyte line (assumed to be $m = 1/4$); $a$and $c$ are the proliferation rate and self-renewal probability of the $P$ cells; and, similarly, $b$ and $d$ are the proliferation rate and self-renewal probability of the $C$ cells. $E(T)$ denotes expected interdivision time of a cell.
https://doi.org/10.1371/journal.pcbi.1006664.g005

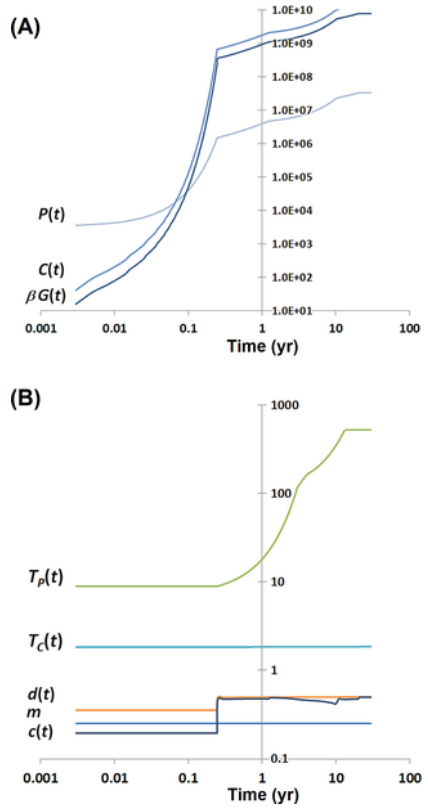



**Fig 6. Estimated dynamics of human hematopoiesis in the bone marrow, starting from its onset in the final weeks of fetal life.**
(For additional details see the twin sections titled *Estimates of the parameters of the model of age-dependent dynamics of the granulocyte arm of the hematopoietic system*, in the Results and in the **S1 Appendix**). (A) Counts of the HSCs ($P(t)$) and CMPs ($C(t)$), and the rate of granulocyte production per day ($\beta G(t)$), as a function of time (individual's age), calculated for human body weight following the Theron's formula. (B) Time (individual's age) dependent parameters of the mathematical model of the human hematopoietic system: expected time (days) between HSC divisions $T_P(t) = 1/a(t)$, expected time (days) between CMP divisions $T_C(t) = 1/b(t)$, maturation probability of the HSC $c(t)$, and differentiation probability of the CMP, $d(t)$. Ratio of the committed cells in the hematopoietic cell lineage associated with the granulocyte lineage is assumed equal to $m = 1/4$ for all ages. Please note the log-log scale of the graphs.
https://doi.org/10.1371/journal.pcbi.1006664.g006

Our model is consistent with the lifetime changes in the bone marrow volume and the HSC interdivision times [47, 48], distinguishes between the periods before and after birth, and accounts for the body mass and HSC proliferation changes during childhood and adulthood. This requires that some of the model parameters change with time $t$ (individual's age). We found that it is sufficient to vary (i) the mean interdivision time of the HSC equal to $T_p = 1/a(t)$, (ii) the maturation probability of the HSC equal to $c(t)$, and (iii) the differentiation probability of the CMP equal to $d(t)$. See **S1 Appendix** for explanation of mathematical symbols. The time (age) patterns of these and other coefficients that fit the data in refs. [47, 48], are depicted in **Fig 6B**.

The model accommodates two apparently contradictory observations in the data viz. that (i) the interdivision times of the HSC dramatically increase over the lifetime, and (ii) the granulocyte volume remains proportional to the body weight [47, 48]. Surprisingly, this can be achieved by relatively small changes in the cell maturation and differentiation coefficients. In the real-life system, this is accomplished by nonlinear regulatory feedbacks, with possible configurations similar as in ref. [45]. However, for our purposes, it is sufficient to assume that differentiation coefficients $c(t)$ and $d(t)$ vary with time (**Fig 6B**). Mathematical details and parameter estimates are found in the corresponding sections of the **S1 Appendix**.

Model of expansion of the CSF3R truncation mutant in the bone marrow in the form of the Moran process with variable population size, directional selection, and recurrent mutation.

In contrast to the standard Moran process, this model assumes that the population size (cell count) *N(t)* varies in time. The population consists of cells of two types: wild type (WT; cells that already acquired the *ELANE* mutation) and *CSF3R* truncation mutants (M). For the WT-cells and M-cells equally, life lengths depend on individual's age. Upon death of a cell, another randomly chosen cell proliferates. Selective advantage is represented by a bias in choice of proliferating cell (as in the standard Moran model. WT-cell may irreversibly mutate into a M-cell at rate $\mu$. Mathematical details are found in the corresponding section of the **S1 Appendix**.

Technically, the model is limited to the CMP compartment of the bone marrow, since (1) cells in this compartment respond to G-CSF signaling as opposed to the HSC, which most likely do not, and (2) CMP compartment is much larger than the HSC one. Further stages of granulocyte precursors are assumed to only transmit the descendants of the mutated cells into peripheral blood

and tissue and not to have even limited self-renewal properties. Before the GCS-F treatment is initiated, the mutants do not have selective advantage (i.e. $s = 0$). Advantage appears at the time G-CSF treatment is administered, assumed to be six months after birth.

## Results

### Proof-of-principle model

#### Mutation rate, selection and age at sMDS onset in the proof-of-principle model.

In this section, we consider a simple Moran model with the mutants being cells carrying the *CSF3R* truncation mutation. Since around 70% of sMDS associated with SCN carry this mutation, this means that fixation of the mutant (i.e., elimination of the wild-type *CSF3R* in the SCN population) occurs with probability 0.7 (ref. [23]). Assuming this and a given selection coefficient of the mutant over the wild type, we calculate (see **S1 Appendix**) the expected time to fixation and the required number of mutant cells at time 0 of the model (corresponding to birth of the individual). **Table 1** presents results of the computations of the expected time to fixation of a *CSF3R* truncation mutation with probability of fixation kept at $P = P[fixation] = 0.7$, assuming the Moran process with directional selection. The parameter values are consistent with the adolescent phase values in the accurate model of bone marrow expansion, as well as with the experiment-based estimates of the selective advantage of cells harboring the *CSF3R* truncation mutation, as outlined further on.

#### Determination of the expected time to fixation of the CSF3R truncation mutant and the required count of mutant-harboring cells at the end of fetal bone marrow expansion.

Expected times to fixation of the *CSF3R* truncation mutant were computed by first solving Eq. (1) in the **S1 Appendix** to find the initial count $i$ of mutants required for $P = 0.7$, given the summary number of HSC and CMP cells $N$ = 1.98×10$^8$ cells/kg ×75 kg (the average adult body weight) and selection coefficient $s$ varying in a wide range.

In mathematical terms, if the probability of fixation of the mutant is provided by the Eq. (1) in the **S1 Appendix**

$$P[T_N < T_0] = \frac{1-(1-s)^i}{1-(1-s)^N},$$

and for large $N$ and small $i$, the expected time to fixation (given that fixation occurs), is asymptotically equivalent to [38]

$$E[T_N | T_0 > T_N] = \frac{2\ln N - \ln i}{s},$$

then we can solve the first of these two equations for $i$, assuming $P[T_N < T_0] = 0.7$ to obtain

$$i = \log_{1-s}\{1 - 0.7[1-(1-s)^N]\}$$

and then substitute this into the second equation to obtain

$$E[T_N | T_0 > T_N] = \{2\ln N - \ln\log_{1-s}\{1 - 0.7[1-(1-s)^N]\} - \ln(1-s)\}/s$$

The latter has to be divided to the number of cell divisions per year (assumed to be equal to 90. as in Stiehl et al. [45]) to obtain time in years.

The resulting mutant cell counts and the expected age at fixation of the *CSF3R* truncation mutation are collected in the second and fourth column of **Table 1**, respectively.

For $s$ from the interval (0.02, 0.1), the estimates of the expected time to fixation of the mutant (time when only mutant allele remains) belong to the interval [4.86, 23.41] (yr), and are approximately consistent with the timing of the sMDS onset. From the European SCN Registry data, age at diagnosis of SCN with sMDS and *CSF3R* mutation is 13 ± 9 years. However, it is necessary that at the time G-CSF treatment begins, $i$ = 11–60 mutant cells are already present in the cell population (second column of **Table 1**).

#### Determination of the mutation rates required to obtain the mutant cell counts at birth.

To shed light at the possibility of this number of mutants being present approximately at the birth time, we solved Equation (3) in the **S1 Appendix** to obtain the estimates of mutation rate $\mu$ which makes it possible, absent selection by G-CSF, to obtain the corresponding initial mutant count $i(t) = i$ in the range 11–60 in the fetal hematopoietic population of HSC and CMP expanding to the size of about $N(t) = N$ = 1.98×10$^8$ cells/kg ×5 kg (the approximate infant body weight).

Mathematically, under the branching process model (**S1 Appendix**), we obtain the equations for the expected (average) number $N(t)$ of normal and $i(t)$of mutant cells [8]

$$N(t) = N_0\exp[(1-\mu)\lambda t], i(t) = N_0\{\exp(\lambda t) - \exp[(1-\mu)\lambda t]\}$$

Eliminating time from the relationship between $N(t)$ and $i(t)$, we obtain that the expected number of mutant cells is equal to

$$i(t) = N(t) - N_0^{\mu} N(t)^{1-\mu} \cong \mathrm{N(t)} - \mathrm{N(t)}^{1-\mu} \cong \mu \mathrm{N(t)} \ln \mathrm{N(t)}$$

where $N(t)$ (resp. $N_0$) is the number of cells after (resp. before) expansion. The approximation on the right-hand side is valid for moderate $N_0$ and small $\mu$. Inverting this expression, we obtain the mutation rates for a given expected number of mutant cells.

The resulting mutation rates per cell division, ranging from $9.57\times10^{-10}$ to $4.99\times10^{-09}$, with cell cycle time of the CMP (which dominate in the pooled HSC and CMP population; see further on) assumed equal to 4 days (as in Stiehl et al. [45]), do not exceed the values considered normal for human cells (**Table 1**).

As for **Table 2**, we select the mutation rates and expected times to fixation to be a subset of those in **Table 1** and compare the respective selection coefficients stemming from the simplified model and the comprehensive model. A more complete review of possible parameter combination is possible using **Fig 2** as a nomogram; see further on.

| Age at CSF3R truncation fixation (yr) | Mutation rate $\mu$ (cell division$^{-1}$) | Selection coefficient ($s$), based on proof-of-principle model | Selection coefficient ($s$), based on comprehensive mo |
|---|---|---|---|
| 4 | $1 \times 10^{-9}$ | 0.10 | 0.054 |
| 13 | $2.5 \times 10^{-9}$ | 0.04 | 0.014 |
| 22 | $5 \times 10^{-9}$ | 0.02 | 0.008 |

Comparison of estimates of selection coefficients (rounded) needed for fixation of the CSF3R truncation mutant at a given age, based on the proof-of-principle and comprehensive models, with the corresponding mutation rates (rounded) corresponding to simulations underlying estimates in Table 1.

https://doi.org/10.1371/journal.pcbi.1006664.t002

**Table 2. Comparison of estimates of selection coefficients needed for fixation of the *CSF3R mutant*.**
https://doi.org/10.1371/journal.pcbi.1006664.t002

Comprehensive model of fixation of CSF3R truncation mutants

The results of the proof-of-principle modeling summarized in **Table 1** suggest that it is feasible to build a more comprehensive model consistent with normal hematopoiesis as well as mutation and selection mechanisms modified by the age-dependent cellularity of the bone marrow and administration of pharmacological G-CSF.

Estimates of the parameters of the model of age-dependent dynamics of the granulocyte arm of the hematopoietic system.

We assume that the total number of HSC, CMP, and granulocytes is proportional to body weight, which increases according to Theron's formula [49] from 3.4 kg at the birth time to 75 kg in the adult life. The exact numbers of these cells at time are calculated based on the estimates given by Stiehl et al. [45] (in that paper's Online Supplement, Scenario 2). Further details are available in the corresponding section of the **S1 Appendix**. **Fig 6** presents the age-trajectories of model coefficients as well as those of the cell numbers in the two compartments and the flux rate of mature granulocytes into blood.

Mutation rate, selection and age at sMDS onset in the detailed model including recurrent mutation.

The main results of the paper are obtained by application of the comprehensive model of expansion of the *CSF3R* truncation mutants in the bone marrow. As described in Methods section, the model hypotheses are consistent with the age-dependence of the cell number in the CMP compartment. We disregard the influx of mutants arising in the HSC compartment, since this is a relatively minor influence compared to their number in the CMP, the increase of which is driven by their selective advantage following the G-CSF treatment. Because the Moran model with recurrent mutation has mutant fixation probability equal to 1, it applies strictly speaking only to the 70% of cases with *CSF3R* truncation mutation fixed at the sMDS diagnosis.

**Fig 2A** depicts the age at which the *CSF3R* truncation mutants replace the normal cells in the CMP compartment. The age at replacement reaches the mean value of 13 years observed in clinical data for a range of parameter values, including mutation rate expected for human genome (ca. $10^{-9}$) and modest selection coefficient values such as 0.014 consistent with small selective advantage expected based on **Fig 3** though not explicitly quantifiable. **Fig 2B** depicts the values of the selection coefficient , which for a given mutation rate lead to sMDS onset at 4, 13 and 22 years, respectively.

Comparison of simulations in **Fig 2** (comprehensive model) and computations in **Table 1** (simplified model), leads to the conclusions summarized in **Table 2**. Estimates of the selection coefficients needed for fixation of the *CSF3R* truncation mutant at ages 4, 13 and 22 years obtained from the comprehensive model are about 2–3 times higher than those based on the simplified model, with the corresponding mutation rates adjusted to those required in the simplified model. This difference stems from two facts: (1) the simplified model does not account for the change of the hematopoietic system performance with age, and (2) the comprehensive model includes recurrent mutations of *CSF3R* over the lifetime not only in the fetal period. In both models, the mutation rates and selection coefficients required seem to be within acceptable ranges.

Another graph, in **Fig 2C**, depicts the corresponding mutant count at 1 year of age. We see that at mutation rates ranging over $10^{-9}$–$10^{-7}$, the mutant count at birth remains in the $10^{1}$–$10^{3}$ range. These results confirm the proof-of-concept analysis and show that in the range of "normal" human mutation rates, the number of mutants at birth is of the order of $10^{2}$—practically undetectable even by very deep sequencing.

## Discussion

Here we presented a model of fixation of a *CSF3R* truncation mutant in the transition from an inherited neutropenia to sMDS: from the expansion phase in the prenatal hematopoietic tissues, to initiation of the G-CSF treatment, to expansion of the mutant, and to replacement of the normal bone marrow by the pre-leukemic mutants. By modifying the simple Moran model of population genetics, we provided an explanation for the evolution of sMDS in about 70% of cases in which *CSF3R* truncation mutant acts as an oncogenic driver. We first used a proof-of-concept two-stage model including the initial creation of the mutant clone before the

selective agent G-CSF has been applied, followed by the period of selective pressure after initiation of treatment. We followed up with a more comprehensive model, which used the estimates of age-dependent productivity changes in hematopoietic stem cells, obtained based on telomere shortening estimates by the Abkowitz and Aviv groups [47, 48].

Our model provides a real-world setting that may further illuminate principles of clonal hematopoiesis of indeterminate potential, first described as age-related clonal hematopoiesis. As recently summarized by [50], HSC clonality and association with malignancy begins with somatic genetic lesions in adult stem cells that accumulate and persist and that "given a large enough population (of HSC), every base pair in the genome will be mutated within at least one HSC". Further, "these mutations provide the substrate for clonal selection". The original and distinctive feature of our present model is to show that mutations occurring during the bone marrow expansion in the fetal period are likely to play a major role in creating this substrate.

The expected times to fixation of the *CSF3R* truncation mutant (4–22 years) are consistent with the timing of the sMDS onset. According to data published from the European SCN Registry data, the average age at diagnosis of SCN with sMDS and *CSF3R* mutation is 13 ± 9 years [51]. The 70% fixation probability requires 11–60 "initial" cells harboring the mutation. We experimentally validated our mathematical model by measuring the growth advantage of the CSF3R D715-expressing cells and found a significant growth advantage (**Fig 1A**). Further validation will require next generation sequencing of specimens from these rare patients. Qiu et al. [52] recently reported that this truncation mutation also permits granulocytic precursors to avoid apoptosis.

Our comprehensive model is based on the hypothesis that the rate of cell division after birth, when the rapid expansion of bone marrow slows down, is still very high. Hence, acquisition of new mutants during that phase is still substantial. However, selection is the force that leads the mutant-receptor cells to dominate. This also means that supply of new mutants in the expansion phase might not be necessary for the disease to emerge. However, it is likely that proliferation slows down by one or two orders of magnitude, depending on exact characteristics of subtypes of stem cells. Then in order to fit the data, somewhat higher selection coefficients are needed. In that case, the comprehensive model will behave approximately as the "proof of the concept" model, i.e. most of the mutant are supplied in the marrow expansion stage.

An alternative hypothesis states that an inherited neutropenia induces a maladaptive increase in replicative stress and higher mutation rate in HSC that contributes to transformation to sMDS/AML [53]. However, measurements of the mutation burden in individual hematopoietic stem/progenitor cells (HSPCs) from SCN patients failed to support that. $CD34^{+}CD38^{-}$ cells were sorted from blood or bone marrow samples and cultured for 3–4 weeks on irradiated stromal feeder cells. The exomes of the expanded HSPC clones were sequenced with unsorted hematopoietic cells from the same patient served as a normal control. The average number of somatic mutations per exome was 3.6 ± 1.2 for SCN, compared to 3.9 ± 0.4 for the healthy controls. Those patient-derived findings support our model. Our conclusions require that the mutation rate per site per cell division equals about $10^{-9}$, which is consistent with normal mutation rate in human genome.

This latter issue warrants discussion since the somatic mutation rate in humans is about two orders of magnitude higher than the germline mutation rate, as suggested by [54]. However, a recent paper by Milholland et al. [55], argues that this former (somatic rate) is of the order of $10^{-9}$ per base per mitosis, while the former (germline rate) is of the order of $10^{-11}$ per base per mitosis (**Fig 1B** in that paper). Moreover, as seen in our **Fig 2**, using the $10^{-7}$ mutation rate

[54] would only slightly change our conclusions.

Two other mechanisms drive the expansion of the CSF3R truncation mutants, (i) the initial CSF3R truncation mutant cell clones arising in the expansion phase of fetal hematopoietic bone marrow and (ii) competitive advantage of the CSF3R truncation mutant harboring cells at later ages, hypothetically due to increased G-CSF pressure. "Mutator phenotype" does not need to be invoked in the SCN progression to sMDS.

A characteristic feature of human cancers is their wide heterogeneity with respect to extent of involvement, genotype, and rate of progression and spread [56]. This variability contrasts markedly to induced animal tumors, which grow at a relatively uniform rate. sMDS/AML secondary to SCN is not an exception, with onset varying from 1 to 38 years of age. Previously, we constructed a stochastic model of the SNC→sMDS→sAML transition based on stochastic events [9]. It considered each new mutation to provide more selective advantage to the arising clone. This linear structure of mutation conferred desirable simplicity to modeling but was not necessarily realistic. In the framework of multitype branching processes and special processes such as Griffiths and Pakes branching infinite allele model [57, 58], more complicated scenarios might be contemplated. Interestingly, the model of ref. [9] suggests that the spread in the age of onset of sAML is not due solely to stochastic nature of clone transitions, but requires a large variability in proliferative potential from one affected individual to another.

Similar effect can be predicted in the Moran process. According to [38], the time course of the Moran process under mutant selective advantage can be split into three periods: (1) relatively long period from small number of mutant cells to a threshold, followed by (2) a much shorter period from the threshold to near-fixation of the mutant, and (3) a relatively long period to complete fixation. Accordingly, once the mutant count exceeds certain threshold, the process accelerates. This results in the spread of times to fixation depending at least as strongly on the selection coefficient as on the "intrinsic" randomness. This justifies the approach we took in this study, to concentrate on the effects of the selection coefficient. In addition, determination of the threshold may help establish a target for monitoring the progress of the disease. Gaining more insight will require a further study.

While our model advances the understanding of multistep progression to cancer with a real-world condition and application to the clinic, other factors could be incorporated. These include: a correlation between G-CSF dosage for neutrophil recovery in SCN patients and the risk of malignant transformation and acquisition of an additional mutation, such as *RUNX1*, in the evolution to sAML.

In the current report, we focus on a single aspect of the SCN-related leukemogenesis: expansion of *CSF3R* truncation mutant cells leading to the sMDS transformation. The model we present here provides potentially testable hypotheses (i) the *CSF3R* truncation mutants are present in [illegible] cells before G-CSF treatment is applied and (ii) a slight selective advantage of the *CSF3R* truncation mutant-harboring cells under G-CSF pressure is sufficient to lead to their expansion. The second hypothesis seems to be

consistent with findings in ref [53]. Current dogma holds that clonal dynamics in relation to the development of sMDS/AML are highly heterogeneous and unpredictable. Our model supports the clinical value of more accurate disease surveillance with next generation sequencing and better timing of therapeutic interventions, such as stem cell transplantation.

## Supporting information

**S1 Appendix. Supporting methods.**

https://doi.org/10.1371/journal.pcbi.1006664.s001
(DOCX)

## References


1. McFarland CD, Mirny LA, Korolev KS. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. Proc Natl Acad Sci U S A. 2014;111(42):15138–43. Epub 2014/10/04. pmid:25277973; PubMed Central PMCID: PMCPMC4210325.
View Article • PubMed/NCBI • Google Scholar

2. McFarland CD, Yaglom JA, Wojtkowiak JW, Scott JG, Morse DL, Sherman MY, et al. The Damaging Effect of Passenger Mutations on Cancer Progression. Cancer Res. 2017;77(18):4763–72. Epub 2017/05/26. pmid:28536279; PubMed Central PMCID: PMCPMC5639691.
View Article • PubMed/NCBI • Google Scholar

3. Davis A, Gao R, Navin N. Tumor evolution: Linear, branching, neutral or punctuated? Biochim Biophys Acta. 2017;1867(2):151–61. Epub 2017/01/23. pmid:28110020; PubMed Central PMCID: PMCPMC5558210.
View Article • PubMed/NCBI • Google Scholar

4. Gao R, Davis A, McDonald TO, Sei E, Shi X, Wang Y, et al. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. Nat Genet. 2016;48(10):1119–30. Epub 2016/08/16. pmid:27526321; PubMed Central PMCID: PMCPMC5042845.
View Article • PubMed/NCBI • Google Scholar

5. Durrett R. Branching process models of cancer. Golubitsky M, Reed M., editor: Springer, Cham; 2015. 63 p.

6. Tomasetti C, Durrett R, Kimmel M, Lambert A, Parmigiani G, Zauber A, et al. Role of stem-cell divisions in cancer risk. Nature. 2017;548(7666):E13–E4. Epub 2017/08/11. pmid:28796214.
View Article • PubMed/NCBI • Google Scholar

7. Nowak MA. Evolutionary dynamics. Exploring the Equations of Life: Harvard University Press; 2006.

8. Kimmel M, Axelrod DE. Branching Processes in Biology: Springer 2015.

9. Kimmel M, Corey S. Stochastic Hypothesis of Transition from Inborn Neutropenia to AML: Interactions of Cell Population Dynamics and Population Genetics. Frontiers in oncology. 2013;3:89. pmid:23641360; PubMed Central PMCID: PMC3638131.
View Article • PubMed/NCBI • Google Scholar

10. Whichard ZL, Sarkar CA, Kimmel M, Corey SJ. Hematopoiesis and its disorders: a systems biology approach. Blood. 2010;115(12):2339–47. pmid:20103779; PubMed Central PMCID: PMCPMC2845894.
View Article • PubMed/NCBI • Google Scholar

11. Matatall KA, Jeong M, Chen S, Sun D, Chen F, Mo Q, et al. Chronic Infection Depletes Hematopoietic Stem Cells through Stress-Induced Terminal Differentiation. Cell Rep. 2016;17(10):2584–95. Epub 2016/12/08. pmid:27926863; PubMed Central PMCID: PMCPMC5161248.
View Article • PubMed/NCBI • Google Scholar

12. Tomasetti C, Vogelstein B, Parmigiani G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. Proc Natl Acad Sci U S A. 2013;110(6):1999–2004. Epub 2013/01/25. pmid:23345422; PubMed Central PMCID: PMCPMC3568331.
View Article • PubMed/NCBI • Google Scholar

13. Horwitz MS, Corey SJ, Grimes HL, Tidwell T. ELANE mutations in cyclic and severe congenital neutropenia: genetics and pathophysiology. Hematology/oncology clinics of North America. 2013;27(1):19–41, vii. pmid:23351986; PubMed Central PMCID: PMC3559001.
View Article • PubMed/NCBI • Google Scholar

14. Mehta HM, Malandra M, Corey SJ. G-CSF and GM-CSF in Neutropenia. J Immunol. 2015;195(4):1341–9. pmid:26254266.
View Article • PubMed/NCBI • Google Scholar

15. Donadieu J, Leblanc T, Bader Meunier B, Barkaoui M, Fenneteau O, Bertrand Y, et al. Analysis of risk factors for myelodysplasias, leukemias and death from infection among patients with congenital neutropenia. Experience of the French Severe Chronic Neutropenia Study Group. Haematologica. 2005;90(1):45–53. Epub 2005/01/12. pmid:15642668.
View Article • PubMed/NCBI • Google Scholar

16. Rosenberg PS, Alter BP, Bolyard AA, Bonilla MA, Boxer LA, Cham B, et al. The incidence of leukemia and mortality from sepsis in patients with severe congenital neutropenia receiving long-term G-CSF therapy. Blood. 2006;107(12):4628–35. pmid:16497969; PubMed Central PMCID: PMCPMC1895804.
View Article • PubMed/NCBI • Google Scholar

17. Kojima S, Tsuchida M, Matsuyama T. Myelodysplasia and leukemia after treatment of aplastic anemia with G-CSF. N Engl J Med. 1992;326(19):1294–5. pmid:1373226.
View Article • PubMed/NCBI • Google Scholar

18. Hershman D, Neugut AI, Jacobson JS, Wang J, Tsai WY, McBride R, et al. Acute myeloid leukemia or myelodysplastic syndrome following use of granulocyte colony-stimulating factors during breast cancer adjuvant chemotherapy. J Natl Cancer Inst. 2007;99(3):196–205. pmid:17284714.
View Article • PubMed/NCBI • Google Scholar

19. Beekman R, Touw IP. G-CSF and its receptor in myeloid malignancy. Blood. 2010;115(25):5131–6. Epub 2010/03/20. pmid:20237318.
View Article • PubMed/NCBI • Google Scholar

20. Ehlers S, Herbst C, Zimmermann M, Scharn N, Germeshausen M, von Neuhoff N, et al. Granulocyte colony-stimulating factor (G-CSF) treatment of childhood acute myeloid leukemias that overexpress the differentiation-defective G-CSF receptor isoform IV is associated with a higher incidence of relapse. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2010;28(15):2591–7. Epub 2010/04/22. pmid:20406937.
View Article • PubMed/NCBI • Google Scholar

21. Touw IP, Beekman R. Severe congenital neutropenia and chronic neutrophilic leukemia: an intriguing molecular connection unveiled by oncogenic mutations in CSF3R. Haematologica. 2013;98(10):1490–2. Epub 2013/10/05. pmid:24091926; PubMed Central PMCID: PMCPMC3789450.
View Article • PubMed/NCBI • Google Scholar

22. Dong F, Brynes RK, Tidow N, Welte K, Lowenberg B, Touw IP. Mutations in the gene for the granulocyte colony-stimulating-factor receptor in patients with acute myeloid leukemia preceded by severe congenital neutropenia. N Engl J Med. 1995;333(8):487–93. Epub 1995/08/24. pmid:7542747.
View Article • PubMed/NCBI • Google Scholar

23. Germeshausen M, Skokowa J, Ballmaier M, Zeidler C, Welte K. G-CSF receptor mutations in patients with congenital neutropenia. Curr Opin Hematol. 2008;15(4):332–7. pmid:18536571.
View Article • PubMed/NCBI • Google Scholar

24. Beekman R, Valkhof MG, Sanders MA, van Strien PMH, Haanstra JR, Broeders L, et al. Sequential gain of mutations in severe congenital neutropenia progressing to acute myeloid leukemia2012 2012-05-31 00:00:00. 5071–7 p.
View Article • Google Scholar

25. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, et al. Genetic alterations during colorectal-tumor development. N Engl J Med. 1988;319(9):525–32. pmid:2841597.
View Article • PubMed/NCBI • Google Scholar

26. Yates LR, Campbell PJ. Evolution of the cancer genome. Nat Rev Genet. 2012;13(11):795–806. pmid:23044827; PubMed Central PMCID: PMCPMC3666082.
View Article • PubMed/NCBI • Google Scholar

27. Ortmann CA, Kent DG, Nangalia J, Silber Y, Wedge DC, Grinfeld J, et al. Effect of mutation order on myeloproliferative neoplasms. N Engl J Med. 2015;372(7):601–12. pmid:25671252.
View Article • PubMed/NCBI • Google Scholar

28. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science. 2015;348(6237):880–6. pmid:25999502; PubMed Central PMCID: PMCPMC4471149.
View Article • PubMed/NCBI • Google Scholar

29. McCulloch EA, Till JE. The radiation sensitivity of normal mouse bone marrow cells, determined by quantitative marrow transplantation into irradiated mice. Radiation research. 1960;13:115–25. pmid:13858509.
View Article • PubMed/NCBI • Google Scholar

30. Groopman JE, Molina JM, Scadden DT. Hematopoietic growth factors. Biology and clinical applications. N Engl J Med. 1989;321(21):1449–59. pmid:2682244.
View Article • PubMed/NCBI • Google Scholar

31. Kaushansky K. Lineage-specific hematopoietic growth factors. N Engl J Med. 2006;354(19):2034–45. pmid:16687716.
View Article • PubMed/NCBI • Google Scholar

32. Rieger MA, Hoppe PS, Smejkal BM, Eitelhuber AC, Schroeder T. Hematopoietic cytokines can instruct lineage choice. Science. 2009;325(5937):217–8. pmid:19590005.

View Article • PubMed/NCBI • Google Scholar

33. FlowJo L. FlowJo, LLC, 2013–2018; [cited 2018]. Available from: https://www.flowjo.com/solutions/flowjo.

34. Kimmel M. Cellular population dynamics. I. Model construction and reformulation. Mathematical Biosciences. 1980;48(3/4):211–24.
View Article • Google Scholar

35. Kimmel M. Cellular population dynamics. II. Investigation of solutions. Mathematical Biosciences. 1980;48(3/4):225–39.
View Article • Google Scholar

36. Stute N, Santana VM, Rodman JH, Schell MJ, Ihle JN, Evans WE. Pharmacokinetics of subcutaneous recombinant human granulocyte colony-stimulating factor in children. Blood. 1992;79(11):2849–54. Epub 1992/06/01. pmid:1375115.
View Article • PubMed/NCBI • Google Scholar

37. Alon U. An Introduction to Systems Biology: Design Principles of Biological Circuits Boca Raton, FL: Chapman & Hall; 2007.

38. Durrett R. Probability Models for DNA Sequence Evolution: Springer; 2008.

39. Donadieu J, Beaupain B, Rety-Jacob F, Nove-Josserand R. Respiratory distress and sudden death of a patient with GSDIb chronic neutropenia: possible role of pegfilgrastim. Haematologica. 2009;94(8):1175–7. pmid:19644144; PubMed Central PMCID: PMCPMC2719042.
View Article • PubMed/NCBI • Google Scholar

40. Dong F, van Buitenen C, Pouwels K, Hoefsloot LH, Lowenberg B, Touw IP. Distinct cytoplasmic regions of the human granulocyte colony-stimulating factor receptor involved in induction of proliferation and maturation. Mol Cell Biol. 1993;13(12):7774–81. Epub 1993/12/01. pmid:8246993; PubMed Central PMCID: PMCPMC364849.
View Article • PubMed/NCBI • Google Scholar

41. Dowling DJ, Levy O. Ontogeny of early life immunity. Trends Immunol. 2014;35(7):299–310. Epub 2014/06/02. pmid:24880460; PubMed Central PMCID: PMCPMC4109609.
View Article • PubMed/NCBI • Google Scholar

42. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. Cell. 2008;132(4):631–44. Epub 2008/02/26. pmid:18295580; PubMed Central PMCID: PMCPMC2628169.
View Article • PubMed/NCBI • Google Scholar

43. Goldie JH, Coldman AJ. A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate. Cancer Treat Rep. 1979;63(11–12):1727–33. pmid:526911.
View Article • PubMed/NCBI • Google Scholar

44. Goldie JH, Coldman AJ. Genetic instability in the development of drug resistance. Semin Oncol. 1985;12(3):222–30. pmid:4048965.
View Article • PubMed/NCBI • Google Scholar

45. Stiehl T, Ho AD, Marciniak-Czochra A. The impact of CD34+ cell dose on engraftment after SCTs: personalized estimates based on mathematical modeling. Bone Marrow Transplant. 2014;49(1):30–7. pmid:24056742.
View Article • PubMed/NCBI • Google Scholar

46. Arino O, Kimmel M. Stability Analysis of Models of Cell Production Systems. Math Modelling. 1986;7(9–12):1269–300. PubMed PMID: WOS:A1986F488000009.
View Article • Google Scholar

47. Shepherd BE, Guttorp P, Lansdorp PM, Abkowitz JL. Estimating human hematopoietic stem cell kinetics using granulocyte telomere lengths. Exp Hematol. 2004;32(11):1040–50. pmid:15539081.
View Article • PubMed/NCBI • Google Scholar

48. Sidorov I, Kimura M, Yashin A, Aviv A. Leukocyte telomere dynamics and human hematopoietic stem cell kinetics during somatic growth. Exp Hematol. 2009;37(4):514–24. pmid:19216021.
View Article • PubMed/NCBI • Google Scholar

49. So TY, Farrington E, Absher RK. Evaluation of the accuracy of different methods used to estimate weights in the pediatric population. Pediatrics. 2009;123(6):e1045–51. pmid:19482737.
View Article • PubMed/NCBI • Google Scholar

50. Jan M, Ebert BL, Jaiswal S. Clonal hematopoiesis. Semin Hematol. 2017;54(1):43–50. Epub 2017/01/17. pmid:28088988.
View Article • PubMed/NCBI • Google Scholar

**51.** Skokowa J, Steinemann D, Katsman-Kuipers JE, Zeidler C, Klimenkova O, Klimiankou M, et al. Cooperativity of RUNX1 and CSF3R mutations in the development of leukemia in severe congenital neutropenia: a unique pathway in myeloid leukemogenesis. Blood. 2014. Epub 2014/02/14. pmid:24523240.
View Article • PubMed/NCBI • Google Scholar

**52.** Qiu Y, Zhang Y, Hu N, Dong F. A Truncated Granulocyte Colony-stimulating Factor Receptor (G-CSFR) Inhibits Apoptosis Induced by Neutrophil Elastase G185R Mutant: implication for understanding CSF3R gene mutatiopns in severe congenital neutropenia. J Biol Chem. 2017;292(8):3496–505. pmid:28073911; PubMed Central PMCID: PMCPMC5336180.
View Article • PubMed/NCBI • Google Scholar

**53.** Xia J, Miller CA, Baty J, Ramesh A, Jotte MRM, Fulton RS, et al. Somatic mutations and clonal hematopoiesis in congenital neutropenia. Blood. 2017. pmid:29092827.
View Article • PubMed/NCBI • Google Scholar

**54.** Araten DJ, Golde DW, Zhang RH, Thaler HT, Gargiulo L, Notaro R, et al. A quantitative measurement of the human somatic mutation rate. Cancer Res. 2005;65(18):8111–7. Epub 2005/09/17. pmid:16166284.
View Article • PubMed/NCBI • Google Scholar

**55.** Milholland B, Dong X, Zhang L, Hao X, Suh Y, Vijg J. Differences between germline and somatic mutation rates in humans and mice. Nat Commun. 2017;8:15183. Epub 2017/05/10. pmid:28485371; PubMed Central PMCID: PMCPMC5436103.
View Article • PubMed/NCBI • Google Scholar

**56.** Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74. pmid:21376230.
View Article • PubMed/NCBI • Google Scholar

**57.** Griffiths RC, Pakes AG. An infinite alleles version of the simple branching process. Adv Appl Prob. 1988;20:489–524.
View Article • Google Scholar

**58.** McDonald T, Kimmel M. A multitype infinite-allele branching process with applications to cancer evolution. Journal of Applied Probability 2015;52(3):864–76.
View Article • Google Scholar