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 2018

TYPE OF REPORT: Annual Report

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.
Alternative RNA Splicing of CSF3R in Promoting Myelodysplastic Syndromes

Myelodysplastic 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.

Splicing factor, myelodysplastic syndromes, CSF3R
TABLE OF CONTENTS

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>4</td>
</tr>
<tr>
<td>2. Keywords</td>
<td>4</td>
</tr>
<tr>
<td>3. Accomplishments</td>
<td>4-7</td>
</tr>
<tr>
<td>4. Impact</td>
<td>7</td>
</tr>
<tr>
<td>5. Changes/Problems</td>
<td>8</td>
</tr>
<tr>
<td>6. Products</td>
<td>8-9</td>
</tr>
<tr>
<td>7. Participants &amp; Other Collaborating Organizations</td>
<td>9-10</td>
</tr>
<tr>
<td>8. Special Reporting Requirements</td>
<td>10</td>
</tr>
<tr>
<td>9. Appendices</td>
<td>11-15</td>
</tr>
</tbody>
</table>
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

Splicing factor, myelodysplastic syndromes, 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?

  o 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.
  o Specific objectives: See “Detailed Accomplishments” below for further information.
  o 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 factors1-6. MDS are molecularly 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. MDS are molecularly characterized by high frequencies (>50%) of somatic mutations in splicing factors SF3B1, U2AF1, SRSF2, and ZRSR2\textsuperscript{7-10}.

**For aim 1**, we have made remarkable progress in identifying the potential contribution of a tyrosine kinase, not Protein Kinase C, to intron retention (Figure 1). First, we improved the minigene construct, making it more specific for detecting spliced forms. Also shown in Figure 1 are data to suggest a possible role for U2AF1 in intron retention.

In addition, we have obtained the cDNAs for other splicing genes: SRSF2 and Luc7L2. We performed site-directed mutagenesis to create the SRSF2 P95H mutation. The construct was made, sequenced verified, and cloned into a eukaryotic expression vector. The construct was then introduced, along with the plasmid containing the CSF3R minigene, into the HEK293T cells (Figure 2). We have also established stable transductants of HEK293T cells bearing the shRNA to LUC7L2, and we have confirmed that there is silencing of the gene at the protein level. We have transfected the cell lines with the CSF3R minigene (Figure 3).

**For aim 2**, we are proceeding to breed the mice, have prepared the lentiviral particles, have confirmed their function in cell lines, and plan to transduce Csf3r-/- bone marrow cells with the alternatively spliced CSF3R, and perform the more informative RNA-Seq. We have obtained pilot gene expression profiles on Ba/F3 cell lines expressing either the full-length versus truncated CSF3R and have performed gene set enrichment analysis that documents different signatures for JAK-STAT, cell cycle, and cancer signaling.
Figure 2. *SRSF2* P95H increased Class IV *CSF3R* expression upon PMA and sodium orthovanadate treatment. Transient co-expression of *SRSF2* P95H increases the ratio of Class IV to Class I *CSF3R* splicing compared to control and *SRSF2* expression in 293FT cells after treatment with PMA and sodium orthovanadate. PCR using primers specific for both Class I and IV isoforms resulted in 794 bp (Class I) and 354 bp (Class IV) amplicons. A representative PCR gel image is shown. Error bars, s.d.

Figure 3. *LUC7L2* knockdown induces increased Class IV *CSF3R* expression upon sodium orthovanadate treatment. Transient expression of *CSF3R* minigene in 293FT cells significantly increases the ratio of Class IV to Class I *CSF3R* splicing compared to control after treatment with sodium orthovanadate. PCR using primers specific for both Class I and IV isoforms resulted in 794 bp (Class I) and 354 bp (Class IV) amplicons. A representative PCR gel image is shown. Error bars, s.d.
• What opportunities for training and professional development has the project provided?

A doctoral student, Ann Wang, is performing this study under the PI’s supervision. 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?


Presentations: Ann Wang poster presentation at the Massey Cancer Retreat June 2017; 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, and Institute of Molecular Genetics and Biochemistry, Seoul National University

• 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?

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
  None.

- Changes that had a significant impact on expenditures
  We continue to 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

- Books or other non-periodical, one-time publications
  Nothing to report.

- Other publications, conference papers and presentations
o 2017, Jan 6: Columbia Univ Dept of Pediatrics
o 2017, Mar 14: University of Pittsburgh School of Pharmacy
o 2017, Oct 13: University of Virginia Cancer Center Grand Rounds
o 2017, Dec 10,11: American Society Hematology Education Session, Chair and Speaker on “Neutropenias”*
o 2018, March 14: Inst of Molecular Genetics & Biochemistry, Seoul National University
o 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

- **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?**

<table>
<thead>
<tr>
<th>Name:</th>
<th>Seth Corey, MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Role:</td>
<td>Principal Investigator</td>
</tr>
<tr>
<td>Researcher Identifier:</td>
<td>sjcorey (NIH commons)</td>
</tr>
<tr>
<td>Nearest person month worked:</td>
<td>1</td>
</tr>
<tr>
<td>Contribution to Project:</td>
<td>Seth Corey supervises the entire project</td>
</tr>
<tr>
<td>Funding Support:</td>
<td>NIH, LLS, DOD, ALSF, Hyundai</td>
</tr>
</tbody>
</table>

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

  Corey, PI  
  1/1/18 – 12/31/19  
  0.6 calendar Hyundai  
  Hope on Wheels  
  Targeting autophagy in childhood cancer  
  Overlap: none

  Corey, PI  
  10/1/17 – 9/30/19  
  0.48 calendar  
  Alex’s Lemonade Stand Foundation  
  Predictive Modeling of MDS Disease Progression from Inherited Bone Marrow Failure Syndromes Using Novel Branching Process and Next Generation Sequencing to Improve Patient-Specific Survival  
  Overlap: none

- **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.

8. **SPECIAL REPORTING REQUIREMENTS**

- **COLLABORATIVE AWARDS**

  Not applicable.

- **QUAD CHARTS**

  Not applicable.
9. APPENDICES

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

Frances Austin1 | Usua Oyarbide1 | Gita Massey1 | Margaret Grimes2 | Seth J. Corey1,3

1Division of Pediatric Hematology, Oncology, and Stem Cell Transplantation, Children’s Hospital of Richmond at Virginia Commonwealth University, Richmond, Virginia
2Department of Pathology, Virginia Commonwealth University, Richmond, Virginia
3Cancer 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 non-pathogenic 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
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. 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 µg/µl). 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 cancer¹¹,¹² (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


**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](https://doi.org/10.1002/pbc.26584)