

Title: Acute promyelocytic leukemia with torque teno mini virus (TTMV)::*RARA* fusion: an approach to screening and diagnosis

Running title: Acute promyelocytic leukemia with TTMV::*RARA*

Authors: Harrison K. Tsai^{1,2*}, Mark F. Sabbagh^{3*}, Meagan Montesion⁴, Erik A. Williams^{5,4}, Arnaldo Arbini⁶, Daniel R. Boué⁷, Emily M. Harris⁸, Franziska Wachter⁸, Leslie Grimmett¹, Andrew E. Place⁸, Fabienne Lucas², Valentina Nardi³, Annette S. Kim², Carlo Brugnara⁹, Barbara Degar⁸, Jessica Pollard⁸, Marian H. Harris¹, Jacob R. Bledsoe^{1,2}

1. Department of Pathology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA
2. Department of Pathology, Brigham & Women's Hospital, Harvard Medical School, Boston, MA, USA
3. Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA
4. Foundation Medicine Inc., Boston, MA, USA
5. Department of Pathology and Laboratory Medicine, University of Miami, Sylvester Comprehensive Cancer Center, and Jackson Memorial Hospitals, Miami, FL, USA
6. Department of Pathology, NYU Grossman School of Medicine, New York City, NY, USA
7. Department of Pathology & Laboratory Medicine, Nationwide Children's Hospital and The Ohio State University, Columbus, OH, USA
8. Dana Farber/Boston Children's Cancer and Blood Disorders Center, Boston, MA, USA
9. Department of Laboratory Medicine, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

*These authors contributed equally to this work.

Contributing author contact information:

HT: harrison.tsai@childrens.harvard.edu
MS: msabbagh@mgh.harvard.edu
MM: mmontesion@foundationmedicine.com
EW: eaw127@med.miami.edu
AA: arnaldo.arbini@nyulangone.org
DRB: daniel.boue@nationwidechildrens.org
EMH: emily_harris@dfci.harvard.edu
FW: franziska_wachter@dfci.harvard.edu
LG: leslie.grimmet@childrens.harvard.edu
FL: fablucas@uw.edu
VN: vnardi@mgh.harvard.edu
AP: Andrew_place@dfci.harvard.edu
AK: kimannet@med.umich.edu
CB: carlo.brugnara@childrens.harvard.edu
BD: barbara_degar@dfci.harvard.edu
JP: jessica_pollard@dfci.harvard.edu

MH: marian.harris@childrens.harvard.edu

Address correspondence to:

Jacob R. Bledsoe
Boston Children's Hospital
300 Longwood Ave.
Boston, MA 02115
Phone: 617-355-7431
Fax: 617-730-0207
Email: jacob.bledsoe@childrens.harvard.edu

Declarations:

Funding: None

Conflicts of interest: MM: Employee of Foundation Medicine, Inc. and shareholder of Roche Holding AG.

Word count: 6936

Number of tables: 3

Number of figures: 4

Number of supplementary tables: 3

Number of supplementary figures: 4

Abstract

Acute promyelocytic leukemia (APL) with variant *RARA* translocation is linked to over 15 partner genes. Recent publications encompassing six cases have expanded the spectrum of *RARA* partners to torque teno mini virus (TTMV). This entity is likely under-recognized due to lack of clinician and pathologist familiarity, inability to detect the fusion using routine testing modalities, and informatic challenges in its recognition within next-generation sequencing (NGS) data. We describe a clinicopathologic approach and provide necessary tools to screen and diagnose APL with TTMV::*RARA* using existing clinical DNA or RNA-based NGS assays, which led to identification of four cases, all without other known cytogenetic/molecular drivers. One was identified prospectively and three retrospectively, including two from custom automated screening of multiple data sets (50,257 cases of hematopoietic malignancy, including 4809 acute myeloid leukemia (AML)/myeloid sarcoma/APL cases). Two cases presented as myeloid sarcoma, including one with multiple relapses after AML-type chemotherapy and hematopoietic stem cell transplant (HSCT). Two cases presented as leukemia, had a poor response to induction chemotherapy, but achieved remission upon re-induction (including all-trans retinoic acid (ATRA) in one case) and subsequent HSCT. Neoplastic cells demonstrated features of APL including frequent azurophilic granules and dim/absent CD34 and HLA-DR expression. *RARA* rearrangement was not detected by karyotype or FISH. Custom analysis of NGS fusion panel data identified TTMV::*RARA* rearrangements, and in the prospectively identified case, facilitated monitoring in sequential bone marrow samples. APL with TTMV::*RARA* is a rare leukemia with a high rate of treatment failure in described cases. The diagnosis should be considered in leukemias with features of APL that lack detectable *RARA* fusions and other drivers, and may be confirmed by appropriate NGS tests with custom informatics. Incorporation of ATRA may have a role in treatment but requires accurate recognition of the fusion for appropriate classification as APL.

Introduction

Acute promyelocytic leukemia (APL) is a distinct type of childhood and adult leukemia that is characterized by the presence of a *PML::RARA* gene fusion, most commonly in the form of the diagnostic t(15;17)(q24;q21) translocation. APL with variant *RARA* translocation to genes other than *PML* occurs rarely and is acknowledged as a subtype of APL by the WHO classification of hematolymphoid tumors.^{1,2} The diagnosis of APL depends on recognition of characteristic clinical and pathologic features including the tendency to present with coagulopathy, promyelocytic morphology, and a typical flow cytometry immunophenotype that includes lack of HLA-DR and CD34 expression in most cases. These features prompt cytogenetic and molecular testing to confirm the presence of *PML::RARA* or alternate *RARA* fusion. Rapid recognition of APL is necessary to identify patients who will benefit from immediate initiation of all-trans retinoic acid (ATRA) to limit their risk of potentially life-threatening bleeding related to disseminated intravascular coagulation. Accurate diagnosis is critically important as the treatment strategy for APL differs from that of other types of AML and centers on therapy with ATRA and arsenic trioxide (ATO).³ Among pediatric patients with APL, the use of ATRA and ATO has been shown to safely allow elimination or substantial reduction of exposure to cytotoxic chemotherapy with non-inferior outcomes.⁴

Torque teno mini virus (TTMV) is a non-enveloped, circular, single-stranded DNA virus in the Anellovirus family.⁵ Anelloviruses are the most abundant eukaryotic viruses in the human virome, and it has been estimated that 90-100% of humans are infected at an early age.^{6,7} Anelloviruses have been detected in almost all human body fluids suggesting that the virus infects many different cell types.⁶ Prior studies have identified particular TTMV types in the serum or plasma of patients with various hematologic malignancies including lymphomas and multiple myeloma, suggesting a possible etiologic link.^{8,9}

Recent reports have documented the fascinating occurrence of fusions between *RARA* and TTMV as a cause of APL with clinicopathologic features similar to typical APL, including response to APL-type therapy.^{10,11} To our knowledge only six prior cases have been documented in the peer-reviewed literature,^{10,12-15} with variable clinical and molecular characterization. While the occurrence of APL with TTMV::*RARA* is thought to be uncommon, lack of awareness by clinicians and pathologists, and limitations of current clinical testing modalities may result in under-recognition of this entity.

We report four cases of APL with TTMV::*RARA* fusions that were not detected by routine cytogenetic and molecular testing, but were identified by custom analysis of fusion data from two commercial clinical platforms. We highlight clinicopathologic features useful to recognize these cases and describe an approach to screening and identification. Using an automated method, we screen several large cohorts of acute leukemia and other hematopoietic malignancies for TTMV and Anellovirus integration, confirming the rarity of this entity and allowing initial estimations of its frequency in pediatric and adult populations.

Materials and Methods

Case identification and screening of historic cases

The study was conducted in accordance with the Helsinki Declaration and with approval of Institutional Review Boards. The first stage of the study assembled a cohort of leukemias both retrospectively (review of pathology files) and prospectively (during routine clinical practice) with clinicopathologic features suggestive of APL but with negative *PML::RARA* and/or *RARA* break-apart fluorescent in situ hybridization (FISH) testing. These candidates were then tested by a clinical fusion panel supplemented by custom informatic methods designed to screen for *TTMV::RARA*.

Features suggestive of APL included:

- a. Promyelocyte-like or leukemic cell morphology with heavy cytoplasmic granulation.
- b. APL-like flow cytometry immunophenotype (negative for HLA-DR and negative or partial CD34).
- c. Cases in which a diagnosis of APL was initially considered by the pathologist but subsequently excluded by *PML::RARA* and/or *RARA* break-apart FISH.

The second stage of the study screened broad cohorts of hematologic malignancy samples with existing clinical fusion panel data through application of our custom informatics methods.

NGS panel testing for fusions

RNA-based fusion panels: Total nucleic acid extracted from peripheral blood, bone marrow, or extramedullary disease sites was tested at the Clinical Laboratory Improvement Amendments (CLIA)-certified Laboratory of Molecular Pediatric Pathology at Boston Children's Hospital (BCH) between 6/11/2018 and 9/29/2023 or at the CLIA-certified Center for Integrated Diagnostics at Massachusetts General Hospital (MGH) between 5/18/2017 and 8/3/2023 using clinically validated targeted RNA next generation sequencing (NGS) panels based on anchored multiplex PCR (AMP) designed principally for fusion detection.¹⁶ The assays include *RARA* gene specific primers targeting the 5' ends of exons 1-9 as 3' fusion partners and the 3' ends of exons 6-7 as 5' fusion partners, where exon annotations are relative to the *RARA* transcript NM_000964.4. The BCH assay utilized the commercially available ArcherDx FusionPlex Heme v2 kit (Integrated DNA Technologies, Coralville, Iowa, USA; Product Number AB0063), with library preparation performed according to recommended assay protocols including the use of unique molecular identifiers (UMI). The MGH assay also utilized primers from ArcherDx within an internally developed protocol for library preparation. Samples underwent 2 x 151 bp paired-end Illumina sequencing followed by bioinformatics processing consisting of Archer Analysis v6.2.7 (AA) under default parameters at BCH and an internally developed computational pipeline at MGH. NGS samples not passing clinical quality control metrics were excluded.

DNA-based panel: Comprehensive genomic profiling (CGP) was performed by Foundation Medicine Inc (FMI) between 12/28/2013 and 3/7/2023 using the FoundationOne@Heme assay as part of routine clinical care in a CLIA-certified, College of American Pathologists-accredited, New York State-approved Laboratory, as previously described.¹⁷ DNA extracted from peripheral

blood, bone marrow, or formalin-fixed, paraffin-embedded (FFPE) tissue underwent a hybridization-capture, adapter ligation-based library preparation, followed by 2 x 49 bp paired-end Illumina sequencing to a median exon coverage depth of at least 500x for all exons and select introns of over 400 cancer-related genes including *RARA*. Analysis for genomic alterations, including short variant alterations (base substitutions as well as small insertions and deletions), copy number alterations (amplifications and homozygous deletions), and gene rearrangements, was performed as previously described.¹⁸ Approval for inclusion of these samples in this study, including a waiver of informed consent and the Health Insurance Portability and Accountability Act (HIPAA) waiver of authorization, was obtained from the Western Institutional Review Board (Protocol No. 20152817).

The cohorts tested by NGS panels for rearrangements consisted of:

1. BCH cohort: n=510 pediatric hematologic samples comprised of 110 AML, 256 B-ALL, 40 T-ALL, 12 lymphoma, 8 MPAL, 9 MDS, 5 CML, 4 MPN, 30 remission, and 36 other/unknown.
2. MGH cohort: n=2174 hematologic samples including 862 adult (age >18 years) AML and 10 pediatric AML.
3. FMI cohort: n=47,573 total cases of hematopoietic malignancies including 3112 adult and 650 pediatric AML/myeloid sarcoma, 53 adult and 12 pediatric APL, and 2065 adult and 1183 pediatric other leukemias, respectively.

TTMV::*RARA* custom analysis

Manual review for TTMV::*RARA* was performed by visually screening binary alignment map (BAM) files in Integrative Genomics Viewer (IGV) for unusual reads that either (i) aligned partially to *RARA* intron 2 with 5' soft clipping not found in a selection of other samples with established non-APL diagnoses, or (ii) were spliced to *RARA* exon 3 but with 5' soft clipping patterns not seen in the selected other samples. The 5' soft clips were then further evaluated using the web-based interface for nucleotide blast to the nr/nt nucleotide database (<https://blast.ncbi.nlm.nih.gov>; last accessed 9/28/2023). This approach was effective for evaluating a small number of samples, as long as the underlying BAM files were generated by a modern local aligner to the human genome and employed relatively relaxed alignment thresholds for fusion detection, but was impractical for screening a large cohort of samples or BAM files from suboptimal aligners in this context.

We therefore developed a focused automated algorithmic approach to evaluate NGS data specifically for integration events of TTMV within the *RARA* breakpoint cluster region around intron 2.^{19,20} The approach employs standard fusion detection principles based on chimeric split-reads and discordant reads, but also integrates alignments from different aligners to both the human genome and TTMV genomes. The code is available at github.com/ht50/ttmv_rara_SR (last accessed 2/21/2024) and addresses two main objectives:

1. The first objective was to efficiently screen for potential mutant TTMV::*RARA* reads, by using samtools (version 1.17) to extract a select subset of reads from an input NGS bam file and then applying blastn (BLAST+ version 2.13.0) to search for alignments to TTMV (NCBI taxonomy id 93678 in the nt_viruses database). If the input BAM file was generated by a

modern local aligner and included supplementary and unmapped alignments to the human genome, then it was sufficient to extract the subset of paired-end reads harboring an alignment of either read of the pair to an approximate *RARA* breakpoint cluster region (chr17:38486109-38505716 in hg19 or chr17:40329857-40349464 in hg38, spanning *RARA* exon 2, intron 2, and exon 3 buffered by an additional 1 kb on each side). Otherwise, unmapped reads and their paired reads were also extracted from either the BAM file or the original FASTQ files. In pseudocode:

- Input: [bam] file including unmapped reads
- samtools view [bam] [region] | cut -f 1 | sort -u > [screen_reads]
- samtools view -f 4 [bam] | cut -f 1 | sort -u >> [screen_reads] (if needed)
- sort -u [screen_reads] -o [screen_reads] (if needed)
- samtools view -b -N [screen_reads] [bam] > [screen_bam]
- samtools fasta [screen_bam] > [screen_fa]
- blastn -query [screen_fa] -db [nt_viruses] -taxids 93678 -outfmt 6 -out [ttmv.out]
- if [! -s ttmv_out]; then rm [screen_files]

2. The second objective was to evaluate reads found to have blastn alignments to TTMV, and to establish strong evidence of TTMV::*RARA* through identifying mutant junctional reads with chimeric split-read alignments including both (i) a blastn alignment of one part of the read to a TTMV accession number and (ii) a local alignment of another part of the read to *RARA*. The *RARA* local alignments were obtained from the intermediate BAM ([screen_bam]) when a modern local aligner was used to generate the input BAM, otherwise a realignment step was performed with bwa mem (version 0.7.17). For each chimeric alignment, mutant breakpoints were calculated and aggregated into counts according to TTMV accession number and by breakpoints. Non-chimeric discordant alignments, which were characterized by paired-end reads with only blastn alignments to TTMV for one read of the pair and only local alignments to *RARA* for the other read of the pair, were also counted but were not associated with mutant breakpoints. The TTMV accession number with the highest overall count from both chimeric and non-chimeric alignments was likely to be the best TTMV match.

A final objective was to assemble and characterize a TTMV::*RARA* contig or contigs. To this end, various assemblers including TRINITY²¹ (version 2.5.1) for RNA data and VELVET²² (version 1.2.10 using k-mer lengths of 15-29) for DNA data were applied to paired-end reads with a blastn alignment of either read of the pair to TTMV, supplemented by paired-end reads where at least one read of the pair had a partial local alignment to the approximate *RARA* breakpoint cluster region but without a supplementary, secondary, or primary alignment resolving the 5' or 3' part of the read. The resulting contigs were then aligned to TTMV by blastn and locally aligned to the human genome by bwa mem or blat in order to maximally characterize the fusion event. The contig with the highest blastn alignment score was also evaluated further by blastx to TTMV.

The algorithm was used to screen the BCH, MGH, and FMI cohorts. As an additional proof of concept, public RNA-sequencing data from the initial publication describing APL with TTMV::*RARA* was also re-analyzed by downloading FASTQ files from the NCBI sequence read

archive (SRR14060870, SRR14060871), aligning to hg19 using bwa mem, and applying the algorithm.¹⁰ Software used in this project included applications installed and configured by BioGrids.²³

Anellovirus::*RARA* screening

To screen more widely across the anelloviridae family of viruses, the TTMV::*RARA* script included an option to replace the taxonomy id 93678 for TTMV with a fixed list of taxonomy ids from anelloviruses (github.com/ht50/ttmv_rara_SR/blob/main/anellovirus.txids). The list consisted of 645 taxonomy ids obtained by executing the blast script “get_species_taxids.sh –t 687329”, where 687329 is the taxonomy id for the family anelloviridae.

Expressed TTMV::*RARA* quantification

In RNA-based fusion panel cases, analysis of *RARA* isoform junctions was performed as previously described to enable RNA quantification of TTMV::*RARA* in terms of an expressed variant allele frequency (VAF).²⁴ In cases where TTMV was directly spliced to *RARA* exon 3, an expressed VAF of the fusion was calculated by counting TTMV::*RARA* split-reads spliced to exon 3 and dividing by split-reads of all junctions spliced to exon 3 derived from gene-specific primers targeting exon 3 or downstream exons. In cases where TTMV::*RARA* fusion transcripts connected TTMV to *RARA* intron 2 and also harbored aberrant cryptic *RARA* splicing connecting a segment of intron 2 to exon 3, empirical analysis was performed to confirm that the splice event did not occur in a historical cohort, and the expressed VAF of the cryptic isoform junction then served as a proxy quantification for the fusion. For serial samples that were subsequently negative for TTMV::*RARA* (based on the anellovirus::*RARA* algorithm, isoform split-read analysis, and manual review), an additional step of grep analysis of (i) the 30 bp junctional sequence spanning the diagnostic TTMV::*RARA* mutant junction and (ii) its reverse complement was also applied to the raw FASTQ files prior to UMI reduction, to further support an absence of the fusion.

Results

Case identification

We searched several clinical cohorts for APL with TTMV::*RARA*, resulting in detection of four cases. Both methods of custom analysis (manual review and automated algorithm) were capable of identifying TTMV::*RARA* fusions in all four cases, whereas the clinically validated pipelines of the RNA and DNA-based panels were consistently negative since these pipelines were not intended to detect fusions from viral genomes. The cases were discovered as follows. First, retrospective analysis of BCH pathology records identified five candidate acute leukemias from 2012-2022 with clinicopathologic features of APL but without *PML*::*RARA* or other driver fusions by conventional testing (karyotype, FISH, and/or RT-PCR). These five candidates were then tested by the fusion panel, and one of the five was confirmed to harbor TTMV::*RARA* fusion (Case 1) upon custom analysis. Using this experience, Case 2 was identified at BCH at the time of initial leukemia diagnosis, where clinical ordering of the fusion panel and custom analysis for TTMV::*RARA* were prompted by pathologist recognition of pathologic features

suspicious for APL in the absence of detectable *RARA* fusion by conventional testing. Thus, identification of Case 1 and Case 2 required both (i) the decision to submit for fusion panel testing and (ii) custom analysis of the data. Manual review was the initial method of detection in Cases 1-2 (**Figure 1A-B**), followed by algorithmic analysis for enhanced characterization. The remainder of the BCH cohort and the entire MGH cohort was negative for TTMV::*RARA* by the algorithmic approach. Case 2 also underwent DNA panel sequencing within the FMI cohort, where the TTMV::*RARA* fusion on the DNA level was identified retrospectively by custom analysis. Cases 3 and 4 were then identified by application of the algorithmic approach to the remainder of the FMI cohort.

As an additional proof of principle, the algorithmic approach was applied to whole transcriptome RNA sequencing data from the inaugural TTMV::*RARA* publication, yielding *RARA*::TTMV::*RARA* contigs that recapitulated and refined previously described findings (**Supplementary Figures 1A-B**).¹⁰ Namely, the contigs demonstrated out-of-frame spliced insertions between *RARA* exons 2 and 3 of chimeric material (274 bp and 383 bp in length) composed of 5' ends with high sequence identity to TTMV isolates and 3' ends derived from short segments of *RARA* intron 2, thereby disrupting the standard frame of *RARA* but predicted to instead utilize a TTMV open reading frame (ORF) containing a putative conserved domain from the TT viral ORF2 superfamily and sharing the frame of *RARA* exon 3 (**Supplementary Figure 2A-B**). The fusion transcripts were thus inferred to contain *RARA* exons 1-2 but the encoded proteins were predicted to lack the portion of the N-terminal activation function domain corresponding to *RARA* exon 2. Of note, SRR14060871 also indicated cryptic splicing out of ~1.7 kb within the TTMV component of its *RARA*::TTMV::*RARA* contig, thereby joining its ORF2 segment to a short, out-of-frame segment within ORF1 (**Supplementary Figure 2B**). Although TTMV::*RARA* breakpoints have usually been characterized within TTMV ORF2, this finding suggests that they may occur more widely throughout the TTMV genome, where splicing within TTMV and/or *RARA* consistently leads to joining of TTMV ORF2 to *RARA* exon 3 in the same frame.

TTMV characteristics

For three of the four cases (Cases 2-4), the blastn alignments were predominantly to a collection of six highly homologous TTMV isolates in the viral database (**Table 1**), featuring pairwise identities of 95.50-99.82% relative to one another (**Supplementary Table 1**). By contrast, Case 1 had blastn alignments to only a single TTMV isolate that was considerably different, with pairwise identities of 74.35-74.65% relative to the six isolates from Cases 2-4. Similarly for the two cases from the original publication, the blastn alignments were predominantly to a collection of five highly homologous TTMV isolates, with pairwise identities of 96-100% relative to each other, 78.44-84.15% relative to the six isolates of Cases 2-4, and 77.32-78.12% relative to the isolate from Case 1.

Split-read analysis of RNA fusion panel data demonstrated TTMV spliced directly to *RARA* exon 3 (Case 1) or fused to a short segment (26 bp) within *RARA* intron 2 (chr17:38494366-38494391) followed by cryptic splicing to *RARA* exon 3 (Case 2) (**Supplementary Table 2**). The TTMV::*RARA* fusion junction of Case 2 (TTMV to chr17:38494366) represented expression of the underlying DNA breakpoint as confirmed later by the DNA panel, whereas in Case 1, the

DNA breakpoint was spliced out on the RNA level but was predicted to occur similarly within *RARA* intron 2. RNA-based assembly generated chimeric contigs of lengths 413 bp (Case 1) and 574 bp (Case 2) whose first 204 bp and 337 bp aligned respectively to TTMV isolates MN768572.1 (98% identity) and MN770942.1 (96% identity) by blastn, and whose remaining 3' ends aligned to *RARA* (100% identity) (**Figure 2A-B**). The blastn alignments indicated splicing out of short segments (87-90 bp) from TTMV in both Cases, followed immediately by potential start codons (ATG) whose open reading frames (ORF) were in the frame of *RARA* exon 3 and predicted by blastx to encode putative conserved domains from the TT viral ORF2 superfamily (pfam02957 amino acids 3-47; E-value < 1e-5) (**Figure 2C-D**). This splicing pattern was also observed within a predicted minor fusion isoform of SRR14060871 (**Supplementary Figure 2C**). The blastn alignments of Case 1 ended immediately before a GT dinucleotide sequence from TTMV (MN768572.1), consistent with a splice event to *RARA* exon 3.

Split-read analysis of DNA data demonstrated both *RARA::TTMV* and *TTMV::RARA* junctions in Case 2 and Case 3, where companion breakpoints were separated by only 16 bp (Case 2) and 12 bp (Case 3) within *RARA* intron 2 but were much further apart (estimated over 1 kb in both Cases) within TTMV, overall consistent with viral integration of linearized segments of TTMV within *RARA* intron 2 (**Supplementary Figure 3; Supplementary Table 2**). DNA-based assembly generated contigs of 441-454 bp (Case 2) and 341-402 bp (Case 3) surrounding the integration sites, however the DNA panel did not sequence across the complete viral insertion (**Figure 2E-F**). The predicted sizes of the viral insertions based on TTMV breakpoints were approximately 1257 bp (Case 2) and 1108 bp (Case 3). In Case 4, only a *TTMV::RARA* breakpoint (at chr17:38503070 within *RARA* intron 2) was detected and generated a 386 bp contig by assembly (**Figure 2G**), where the short read length size (49 bp) of the DNA panel may have reduced the sensitivity for detecting a companion *RARA::TTMV* breakpoint. Blastx of the assembled contig from Case 3 detected the same putative conserved domain from the TT viral ORF2 superfamily (pfam02957 amino acids 3-47; E-value < 1e-5) as Cases 1-2, whereas no conserved domain was detected in Case 4, however blastx yielded an alignment to TT viral ORF1 (pfam02956) of UGV36713.1 with 100% (68/68) identity on the protein level (**Supplementary Figure 4**). The DNA breakpoints connecting TTMV to *RARA* intron 2 thus occurred within ORF2 of TTMV in Cases 2-3, similar to most previously reported *TTMV::RARA* cases, and within ORF1 in Case 4, similar to our re-analysis of SRR14060871. RNA sequencing data was not available to characterize the transcription consequences on the RNA level for Cases 3-4, however the fusion transcripts were predicted to connect TTMV ORF2 in-frame to *RARA* exon 3 either directly or via cryptic splicing within TTMV and/or *RARA* intron 2, similar to prior cases characterized on the RNA level.

We also screened the BCH and MGH cohorts for anellovirus::*RARA* reads by running the *TTMV::RARA* script with the anellovirus option. No additional cases with split-reads or discordant reads between an anellovirus and *RARA* were found.

Case descriptions (clinical course and pathologic features)

Clinical and pathologic features of the cases are summarized in **Tables 2 and 3**, respectively.

Case 1

An adolescent male (11-20 years old) presented with right-sided hearing loss and ear fullness. Imaging demonstrated destructive lesions involving the temporal bone, maxilla, and sphenoid wing. Complete blood count (CBC) and differential were normal. There was no evidence of coagulopathy. Specimens from surgical debridement demonstrated myeloid sarcoma (**Figure 3A**). The neoplastic cells had eosinophilic granular cytoplasm and ovoid to folded nuclei. Necrosis was present. By immunohistochemistry the neoplastic cells expressed CD43, MPO (**Figure 3B**), and lysozyme (subset), and were negative for CD34 (**Figure 3B inset**), CD117, CD56, and other markers. Dual-color/dual-fusion FISH was negative for *PML::RARA* rearrangement. Malignant cells were detected in the cerebrospinal fluid (CSF). The bone marrow was not involved.

The patient underwent treatment with standard dose cytarabine, daunorubicin, and etoposide with weekly intrathecal therapy resulting in clearance of CSF leukemic cells. Three months after initial diagnosis, a paraspinal mass was incidentally noted on chest CT done for infection surveillance. PET/CT showed FDG-avid lesions involving the extra-pleural soft tissue, right lung base, and left iliac bone. Biopsy of the chest mass demonstrated relapsed myeloid sarcoma (relapse #1). Chromosome analysis and MDS/AML FISH panels were negative. The bone marrow was not involved.

The patient received high dose cytarabine and mitoxantrone and underwent radiation therapy to the right posterior mid-thorax, followed by conditioning with busulfan and cyclophosphamide and then HSCT. Seven months after HSCT the patient developed lymphadenopathy, cutaneous and subcutaneous soft tissue lesions involving the scalp. Biopsy of the scalp demonstrated relapsed myeloid sarcoma (relapse #2). Bone marrow biopsy demonstrated involvement by acute leukemia with promyelocytic features (**Figure 3C-D**). The leukemic cells were large with ovoid to indented nuclei, fine chromatin, and abundant cytoplasm with azurophilic granules. No definitive Auer rods were identified. Flow cytometry demonstrated leukemic cells with high SSC and expression of CD33, CD13 (bright), and MPO, with dim to negative CD34 and CD117, and negative for HLA-DR (**Figure 3E**). Chromosome analysis revealed an abnormal karyotype: 46,XY,add(7)(q22)[5]/46,XY[15]. An AML FISH panel including *PML::RARA* dual-color/dual-fusion FISH was negative. Given the absence of detectable *RARA* rearrangement or other driver, the leukemia was considered AML with maturation.

The patient underwent chemotherapy with liposomal daunorubicin and FLAG (fludarabine, cytarabine, and G-CSF) resulting in remission. He received a second HSCT following conditioning with cyclophosphamide/total body irradiation. Twenty-four months after initial diagnosis (7 months after 2nd HSCT; 19 months after 1st HSCT) the patient was found to have supraclavicular neck swelling. Biopsy demonstrated recurrent myeloid sarcoma (relapse #3) and the patient developed leukocytosis with peripheral leukemic cells and disseminated intravascular coagulation. He was transitioned to palliative measures and succumbed to progressive disease.

Fusion panel testing was performed retrospectively on the myeloid sarcoma tissue and bone marrow specimens from the relapse #2 time point, with results described earlier.

Case 2

A previously healthy girl (0-5 years old) presented with hip pain and decreased activity. CBC demonstrated a WBC of 17 K cells/uL, anemia, and thrombocytopenia. PT and INR were slightly elevated. Leukemic cells with irregular nuclei, fine chromatin, prominent nucleoli, and a moderate amount of cytoplasm containing azurophilic granules were present in the blood (**Figure 4A**) and bone marrow (**Figure 4B-D**). No Auer rods were identified. Flow cytometry demonstrated leukemic cells with increased SSC and expression of CD45(dim), MPO, CD13(bright), CD33, CD38, CD64, and CD117 (partial), with dim CD34 expression, and no expression of HLA-DR (**Table 3; Figure 4E**). A diagnosis of AML was rendered and suspicion for APL was raised based on morphology and immunophenotype. The CSF was involved. The patient received three doses of ATRA during which time she developed high fever, progressive leukocytosis and transient coagulopathy.

PML::RARA dual-color/dual-fusion and *RARA* break-apart FISH analyses showed three non-rearranged copies of *RARA*. Chromosome analysis demonstrated an abnormal karyotype with two distinct clones, one with an unbalanced translocation between the short arms of chromosomes 11 and 17 resulting in deletion 11p and 17p, and the other with an isodicentric chromosome 17 resulting in trisomy 17q and monosomy 17p. Given the absence of detectable *RARA* translocation by cytogenetic studies, ATRA was discontinued and chemotherapy was initiated with standard dose cytarabine, daunorubicin with dexrazoxane, gemtuzumab ozogamicin, and intrathecal methotrexate, cytarabine, and hydrocortisone.

A routine DNA-based NGS assay (without *RARA*) demonstrated a pathogenic mutation in *WT1* c.1142C>A (p.S381*) at 2% variant allele frequency (VAF). Copy number analysis demonstrated one copy deletions of *WT1* on 11p and *PRPF8* and *TP53* on 17p, and partial gain of 17q, consistent with the karyotype.

In-house fusion panel testing demonstrated the cryptic *TTMV::RARA* fusion, as previously described. The leukemia was classified as APL with *TTMV::RARA*. Additional DNA-based NGS panel testing performed at FMI detected a pathogenic hotspot *TRRAP* p.S722F variant.

Bone marrow biopsy done after one course of induction chemotherapy showed residual acute leukemia at 3% by flow cytometry. Chromosome analysis showed a normal female karyotype. FISH demonstrated an extra non-rearranged copy of *RARA* in 2% of cells. Custom fusion analysis demonstrated persistent *TTMV::RARA* fusion at low level (see below). Given the suboptimal response, the patient proceeded to IDA-FLAG (idarubicin, fludarabine, cytarabine, and G-CSF) chemotherapy with the addition of ATRA on days 8-21 of the treatment cycle. Bone marrow biopsy done after the second treatment course demonstrated a flow cytometry negative remission. FISH and DNA-based NGS was negative for previously reported findings. Fusion panel testing showed a decreasing presence of the *TTMV::RARA* fusion to a very low level (see below). She then went on to receive high-dose cytarabine and etoposide plus ATRA

for consolidation therapy. Again, she was flow, FISH and DNA NGS panel negative but continued to have the *TTMV::RARA* fusion identified at an extremely low level (see below).

The patient then received HSCT conditioned with busulfan, fludarabine, and clofarabine for disease consolidation. Bone marrow examinations done 35, 101, and 233 days later confirmed remission with no detectable *TTMV::RARA* fusion.

Fusion panel testing was used to track the *TTMV::RARA* fusion. At the time of initial diagnosis, the aberrant cryptic *RARA* splice junction (chr17:38,484,391 to chr17:38504568) connecting the 26 bp intron 2 segment to exon 3 within *TTMV::RARA* fusion transcripts had an expressed VAF of 80.5%. Subsequent fusion panel testing after treatment demonstrated expressed VAFs of 10.2% (after chemotherapy course one), 0.8% (after course two), 0.1% (after course three, pre-HSCT), and 0% (post-HSCT day +35, day +101, day +233).

Case 3

A previously healthy boy (0-5 years old) presented with recurrent fevers, weight loss, bilateral leg and hip pain, and recent viral infection. CBC demonstrated anemia and mild thrombocytopenia. LDH was elevated. Coagulopathy was present with elevated D-dimer and PT. MRI demonstrated marrow signal abnormality within the pelvis and proximal femur.

Bone marrow examination demonstrated a hypercellular myeloid-predominant marrow with 5% blasts and 19% promyelocytic cells along with maturing myeloid and erythroid elements. A differential diagnosis of early myeloid neoplasia versus regeneration after marrow insult was considered. Two weeks later thrombocytopenia worsened and 2% circulating blasts were reported. Repeat bone marrow examination demonstrated AML with features suggestive of APL including hypergranular promyelocytic cells with occasional Auer rods. The blasts expressed MPO, CD33, lysozyme, partial CD117, CD13(dim), and were negative for HLA-DR and CD34.

Cytogenetic analysis demonstrated an abnormal karyotype 46,XY,der(7)ins(7;8)(q22;q13q22) and FISH demonstrated an extra copy of *RUNX1T1* on 8q. NGS testing performed at FMI demonstrated no pathogenic or likely pathogenic variants. Fusion testing through the standard pipeline did not identify any fusions including *RARA*. However, retrospective custom analysis identified a *TTMV::RARA* fusion as described earlier.

The patient underwent AML induction chemotherapy with cytarabine, daunorubicin, and etoposide. Post-induction bone marrow showed persistent leukemia with 15% atypical promyelocytic cells. The patient underwent second induction with cytarabine and mitoxantrone and the subsequent bone marrow and ancillary studies confirmed remission. The patient underwent HSCT following conditioning with busulfan and cyclophosphamide, and all subsequent bone marrow examinations demonstrated continued remission. The patient is alive and disease free for >8 years. The CSF was negative for leukemic cells at all time points.

Case 4

A young adult woman (21-30 years old) presented with leg pain and imaging showed a large destructive lesion of the distal femur. CBC showed mild anemia only, with no circulating leukemic cells. There was no coagulopathy.

Several biopsies of the femur mass yielded only necrotic material. The patient underwent partial resection and reconstruction of the distal femur mass. The bone specimen and popliteal lymph node demonstrated myeloid sarcoma comprised of primitive large cells with open chromatin, indented nuclei, and eosinophilic granular cytoplasm, suggestive of promyelocytes or granular myeloblasts. Bone marrow iliac crest and CSF were negative for leukemia. Flow cytometry of the myeloid sarcoma demonstrated immature myeloid cells expressing CD13 (slightly bright), CD33 (bright), partial CD117, and CD38, but negative for CD34, HLA-DR, CD15, CD14, CD56, and CD10 (**Table 3**). By immunohistochemistry the lesional cells expressed CD45 (dim), MPO, CD43, and CD68.

FISH and cytogenetics were attempted but failed. NGS testing performed at FMI demonstrated pathogenic *KMT2C* p.Q2635* and *WT1* p.R369* variants, and a likely pathogenic *ATR* splice site variant c.6320-2A>C. Fusion testing through the standard pipeline did not identify any fusions including *RARA*. However, retrospective custom analysis identified a *TTMV::RARA* fusion, as described earlier.

PET/CT scan performed after 3 months demonstrated bilateral renal masses that were biopsied and confirmed to be myeloid sarcoma. The patient received AML induction chemotherapy with cytarabine and daunorubicin (7+ 3 regimen) with complete response, followed by consolidation with high dose intermittent cytarabine (HiDAC). The patient underwent HSCT and was known to be free of disease two years later.

Discussion

We describe four cases of APL with *TTMV::RARA* fusion and provide methods to accurately identify and screen for the fusion, including clinicopathologic considerations and novel bioinformatics software that we have made publicly available. Like previously described cases, these share clinical, morphologic and immunophenotypic features of APL but without demonstrable *PML::RARA* rearrangement.¹¹ This entity is relatively unknown among both clinicians and pathologists, but its recognition is important as it appears to be an aggressive disease and patients might benefit from incorporation of APL-directed treatment.¹⁰

The frequency of APL with *TTMV::RARA* is unknown. Given its only recent recognition and the few reports in the literature, this entity can be assumed to be rare, but is likely underrecognized.^{10,12,13} We retrospectively identified and tested five cases of AML with features of APL but without *PML::RARA*, of which one was found to harbor a *TTMV::RARA* fusion. Astolfi and colleagues retrospectively examined whole transcriptome sequencing data from 22 pediatric cytogenetically-normal AML cases and identified one with *TTMV::RARA* fusion.¹⁰ We screened fusion panel data from 50,257 additional cases of hematopoietic neoplasia including 782 pediatric and 4,027 adult AMLs and identified two cases. While selection bias cannot be excluded, these findings confirm that APL with *TTMV::RARA* is uncommon but appears to be

enriched amongst leukemias with promyelocytic features, particularly within the pediatric and young adult age groups.

The cryptic nature of the TTMV::*RARA* fusion complicates diagnosis. Characteristics of TTMV insertions and TTMV::*RARA* fusions for reported cases are summarized in **Supplementary Table 3**. The inserted TTMV sequence size has been reported to range from an estimated 1045-2872 base pairs, which is below the level of resolution of standard cytogenetic techniques and thus undetectable by karyotype and FISH.^{10,12,13} Moreover, since the fusion involves a virus external to the human genome, molecular NGS fusion platforms may not detect TTMV::*RARA* fusions in their standard pipelines, which typically align sequencing data only to the human genome or occasionally to a limited set of viral genomes from established oncoviruses unlikely to include TTMV. Therefore, a high degree of suspicion based on clinicopathologic features is needed to prompt both testing that can robustly capture the TTMV::*RARA* fusion and custom analysis to recognize the fusion within the raw or intermediate data. Screening strategies include fusion/transcriptomic analysis of cases of AML with promyelocytic features or of any AML without a demonstrable molecular driver, followed by manual review of exon 3/intron 2 of *RARA*, or implementation of automated methods to identify such cases.

Knowledge of characteristic clinicopathologic features may prompt recognition of APL with TTMV::*RARA*. Eight of ten reported cases (including this study) have been children and adolescents, along with two adults (both under 40 years old).^{10,12,13,25} Coagulopathy was present in several cases, similar to APL with *PML*::*RARA*.^{10,12,13} CSF involvement, which is uncommon in APL with *PML*::*RARA*, was present in two of four cases presented here. After treatment with ATRA, two cases demonstrated a marked increase in WBC along with coagulopathy in one case.¹³ Myeloid sarcoma is frequent, occurring in four of the nine cases with clinical data, and was the initial presentation in three. In patients presenting with myeloid sarcoma, routine bone marrow examination was frequently negative for leukemia at the time of initial diagnosis but became positive after relapse.^{13,15} Characteristic leukemic cell morphology with frequent azurophilic granules and/or cells resembling promyelocytes were present in all described cases.^{10,12-15,25} Auer rods were present in only some cases.¹³ The leukemic immunophenotype was similar to APL with *PML*::*RARA* including absent or only partial/dim expression of HLA-DR and CD34 in all described cases, along with expression of MPO and other myeloid markers including CD33 and CD13.^{10,12,13} By flow cytometry, CD64 and bright CD13 expression were frequently present and may be characteristic.¹³

Isochromosome 17q and isodicentric 17p11 are uncommon cytogenetic abnormalities in AML.²³ In adults, myeloid neoplasms with isolated iso(17q) typically lack *TP53* mutations and are characterized by mutations in *ASXL1*, *SRSF2*, and *SETBP1*, suggesting different pressures and mechanisms in formation of iso(17q).²⁶ APL infrequently demonstrates iso(17q) or idic(17p11), where it usually co-occurs with t(15;17) resulting in an additional copy of *PML*::*RARA* but can also be an isolated finding associated with a cryptic *PML*::*RARA*.²⁷⁻²⁹ Interestingly, two of the described cases of APL with TTMV::*RARA* fusion demonstrate iso(17q) or idic(17p), suggesting that this may represent a characteristic chromosomal abnormality. Indeed, several cases of AML with promyelocytic features, iso(17q), and no detectable *PML*::*RARA* have been described, possibly representing APL with unidentified TTMV::*RARA* fusions.^{30,31}

Three cases of APL with TTMV::*RARA* harbored *WT1* mutations, which is known to be enriched in APL with *PML>::RARA* and variant fusion partners.^{15,32,33} *FLT3*-ITD and mutations in *NRAS*, *ARID1B*, and *KMT2C* have also been documented. *KMT2C* mutations appear to be relatively common in APL with variant fusions, whereas *FLT3*-ITDs and to a lesser extent *NRAS* mutations are known to be relatively frequent in APL with *PML>::RARA*. Interestingly, one case reported by Chen et al. developed a mutation in the ligand-binding domain of *RARA* after therapy with ATRA suggesting leukemic clonal evolution to bypass ATRA efficacy.¹⁴ The same group reported another case in which germline frameshift *SAMD9* mutation was identified suggesting the possibility of germline predisposition.¹⁵ Gene expression profiling of several prior cases showed clustering of TTMV::*RARA* cases with the group of APL with *PML>::RARA* and away from other AML.^{14,15,25}

Based on the limited number of cases published to date, APL with TTMV::*RARA* appears to have an aggressive clinical course and poor response to standard AML chemotherapy. Of the nine cases with available clinical data, relapse occurred in five and induction failure or poor response to induction in three others. The appropriate treatment of APL with TTMV::*RARA* remains uncertain, but the presence of *RARA* translocation provides rationale for the use of APL-type therapy. Indeed, use of ATRA along with other agents resulted in complete remission in several cases including case 2 of this report.^{10,14,14} ATRA has dual therapeutic effects in APL, it activates transcription of differentiation genes and leads to degradation of the PML-*RARA* oncoprotein.^{34,35} APL with certain *RARA* fusion partners are not sensitive to ATRA therapy.³² The emergence of synthetic retinoids with enhanced differentiation activity, like tamibarotene, has shown promise in rare fusions resistant to ATRA. ATO hypothetically has no therapeutic potential in APL with TTMV::*RARA* because its effect is mediated by its binding to PML of PML-*RARA*. This underscores the ongoing evolution of molecular testing on therapeutic strategies in APL.^{36,37}

TTMV insertion characteristics are similar between all described cases. The TTMV DNA insertion site is within intron 2 and results in either direct splicing of TTMV ORF2 in-frame to *RARA* exon 3 (e.g. Case 1) or retention of a small segment of intron 2 separating the TTMV ORF2 sequence in-frame from exon 3 (e.g. Case 2). The breakpoints within TTMV usually occur within ORF2 but can also occur within ORF1 (Case 4), where reanalysis of public data (SRR14060870) showed how cryptic splicing within both TTMV and *RARA* connects ORF2 in-frame to *RARA* exon 3 with retention of small intervening segments of TTMV ORF1 and *RARA* intron 2. A previously reported case described an absence on the RNA level of 84-bp of viral sequence that was present in the inserted DNA.¹² Similarly, our cases with RNA data showed blastn alignments with high sequence homology to TTMV isolates and predicted splicing out of homologous 87-90 bp segments of viral sequence.

The methods presented here for detecting TTMV::*RARA* in NGS data depended on the collection of TTMV isolates deposited to the NCBI viral database and the ability to generate blastn alignments that scored sufficiently high under the default settings. The blast database used at the time of this study included 6686 TTMV accessions, where the contigs from Cases 1-4 had up to 95.7-100% identities to 4 specific TTMV isolates respectively, of which 3 were highly homologous to one another. Similarly, contigs generated from public RNA sequencing data had 97.6% (203/208) and 100% (285/285) identity to the highly homologous TTMV isolates

MN771920.1 and MN769771.1, while the other 2 cases described in the literature were reported to have 93-94% sequence identity to MN769771.1 once again,¹² and to unspecified isolates.¹³ The recurrence of highly homologous TTMV isolates is notable but could have many different explanations, ranging from spatiotemporal patterns of TTMV infection to potential biological differences between isolates. Overall, the very high identities of all known TTMV::*RARA* cases to date is reassuring for effectively detecting TTMV::*RARA* by blast-based molecular methods. However, these methods theoretically could miss a TTMV::*RARA* fusion arising from a TTMV isolate that has not been submitted to, and is different from, the isolates in the viral database. Thus, the development of other algorithmic approaches may be worth further investment and might increase the yield of screening efforts. Such enhancements could include experimentation with blastn settings and exploration of blastx. Finally, an orthogonal method might be based on screening for alignments with softclips starting at a common genomic location within *RARA* intron 2 to exon 3 but without supplementary alignments resolving the softclip; however, this approach may be challenging to automate from the perspective of specificity.

Several methods have been used to identify TTMV::*RARA* fusions, including whole transcriptome or genome sequencing and long-range DNA sequencing with nanopore and CRISPR guides targeting *RARA*, typically combined with manual review of data. We demonstrate that the fusion is similarly detectable by manual analysis of fusion data from a clinical Archer-based RNA fusion assay, which is a commercially available platform used by many diagnostic molecular laboratories. Moreover, we develop an efficient automated method to eliminate the need for labor-intensive manual review and for more complete, routine evaluation of the data. We also show how the Archer assay enables effective monitoring of disease burden during treatment. The Archer assay should be capable of detecting any TTMV::*RARA* fusion with an RNA breakpoint that is not too distant from *RARA* exon 3, which has been the case for all previously described cases in the literature since the retained segments of intron 2 on the RNA level have always been relatively small. Of note, fusion panel assays without the capability to detect unknown *RARA* partners, without specialized informatics, or without appropriate manual analysis may not identify the fusion. The ability to detect the TTMV::*RARA* fusion relatively rapidly using the Archer-based clinical assay may be advantageous in comparison to less widely-available and more time-intensive testing modalities. Since the assay will also detect a wide range of fusions, isoforms, and expressed variants with diagnostic, prognostic, and therapeutic relevance, it may be useful and reasonable to order this test early in a clinical workup.

Conclusion

We characterize four cases of APL with TTMV::*RARA* and summarize the clinicopathologic and molecular features of the cases described to date. We highlight characteristic features that should prompt clinician and pathologist recognition of this newly-described entity. We also provide a framework for how to identify and screen for the fusion by manual review of fusion panel data and automated analysis, and we demonstrate that the TTMV::*RARA* fusion may be otherwise cryptic on routinely-used diagnostic modalities including karyotype, *PML*::*RARA* fusion and *RARA* break-apart FISH. Determination of the true frequency of APL with TTMV::*RARA* will require larger studies, however our limited data suggests that it may be particularly rare in adults and enriched in pediatric leukemias with an APL phenotype. Future studies will be informative to characterize this entity in general and the biology of the TTMV insertions in particular, to identify more rapid means of fusion detection, and to investigate optimal therapeutic strategies.

References

1. WHO Classification of Tumours Editorial Board. *Haematolymphoid tumours [Internet; beta version ahead of print]*. Lyon (France): International Agency for Research on Cancer; 2022 [cited YYYY Mmm D]. (WHO classification of tumours series, 5th ed.; vol. 11). Available from: <https://tumourclassification.iarc.who.int/chapters/63>.
2. Weinberg, O. K., Porwit, A., Orazi, A., Hasserjian, R. P., Foucar, K., Duncavage, E. J. et al. The International Consensus Classification of acute myeloid leukemia. *Virchows Arch* **482**, 27–37 (2023).
3. Lo-Coco, F., Avvisati, G., Vignetti, M., Thiede, C., Orlando, S. M., Iacobelli, S. et al. Retinoic Acid and Arsenic Trioxide for Acute Promyelocytic Leukemia. *New Engl J Med* **369**, 111–121 (2013).
4. Kutny, M. A., Alonzo, T. A., Abla, O., Rajpurkar, M., Gerbing, R. B., Wang, Y.-C. et al. Assessment of Arsenic Trioxide and All-trans Retinoic Acid for the Treatment of Pediatric Acute Promyelocytic Leukemia. *JAMA Oncol.* **8**, 79–87 (2022).
5. Takahashi, K., Iwasa, Y., Hijikata, M. & Mishiro, S. Identification of a new human DNA virus (TTV-like mini virus, TLMV) intermediately related to TT virus and chicken anemia virus. *Arch Virol* **145**, 979–993 (2000).
6. Kaczorowska, J. & Hoek, L. van der. Human anelloviruses: diverse, omnipresent and commensal members of the virome. *Fems Microbiol Rev* **44**, 305–313 (2019).
7. Virgin, H. W., Wherry, E. J. & Ahmed, R. Redefining Chronic Viral Infection. *Cell* **138**, 30–50 (2009).

8. Pan, S., Yu, T., Wang, Y., Lu, R., Wang, H., Xie, Y. et al. Identification of a Torque Teno Mini Virus (TTMV) in Hodgkin's Lymphoma Patients. *Front Microbiol* **9**, 1680 (2018).
9. Bal, A., Oriol, G., Josset, L., Generenaz, L., Sarkozy, C., Sesques, P. et al. Metagenomic Investigation of Torque Teno Mini Virus-SH in Hematological Patients. *Front Microbiol* **10**, 1898 (2019).
10. Astolfi, A., Masetti, R., Indio, V., Bertuccio, S. N., Messelodi, D., Rampelli, S. et al. Torque teno mini virus as a cause of childhood acute promyelocytic leukemia lacking PML/RARA fusion. *Blood* **138**, 1773–1777 (2021).
11. Rau, R. E. A viral cause of APL. *Blood* **138**, 1653–1655 (2021).
12. Sala-Torra, O., Beppu, L. W., Abukar, F. A., Radich, J. P. & Yeung, C. C. TTMV-RARA fusion as a recurrent cause of AML with APL characteristics. *Blood Adv* **6**, 3590–3592 (2022).
13. Chen, X., Wang, F., Zhou, X., Zhang, Y., Cao, P., Ma, X. et al. Torque teno mini virus driven childhood acute promyelocytic leukemia: The third case report and sequence analysis. *Frontiers Oncol* **12**, 1074913 (2022).
14. Chen, J., Zhou, X., Chen, X., Chen, Q., Yang, J., Lu, Y. et al. Pediatric TTMV::RARA-positive relapsed acute promyelocytic leukemia responsive to venetoclax and achieving long remission after allogenic transplantation. *Pediatr. Blood Cancer* **70**, e30665 (2023).
15. Chen, J., Zhou, X., Wang, Y., Zhang, Y., Chen, X., Wang, T. et al. TTMV::RARA-driven myeloid sarcoma in pediatrics with germline SAMD9 mutation and relapsed with refractory acute promyelocytic leukemia. *Int. J. Lab. Hematol.* (2023) doi:10.1111/ijlh.14189.
16. Zheng, Z., Liebers, M., Zhelyazkova, B., Cao, Y., Panditi, D., Lynch, K. D. et al. Anchored multiplex PCR for targeted next-generation sequencing. *Nat. Med.* **20**, 1479–1484 (2014).
17. He, J., Abdel-Wahab, O., Nahas, M. K., Wang, K., Rampal, R. K., Intlekofer, A. M. et al. Integrated genomic DNA/RNA profiling of hematologic malignancies in the clinical setting. *Blood* **127**, 3004–3014 (2016).
18. Frampton, G. M., Fichtenholtz, A., Otto, G. A., Wang, K., Downing, S. R., He, J. et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat. Biotechnol.* **31**, 1023–1031 (2013).
19. Borrow, J., Goddard, A. D., Sheer, D. & Solomon, E. Molecular Analysis of Acute Promyelocytic Leukemia Breakpoint Cluster Region on Chromosome 17. *Science* **249**, 1577–1580 (1990).
20. Liquori, A., Ibañez, M., Sargas, C., Sanz, M. Á., Barragán, E. & Cervera, J. Acute Promyelocytic Leukemia: A Constellation of Molecular Events around a Single PML-RARA Fusion Gene. *Cancers* **12**, 624 (2020).

21. Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
22. Zerbino, D. R. & Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**, 821–829 (2008).
23. Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P. C., Timony, M. A., Ottaviano, M. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).
24. Tsai, H. K., Gogakos, T., Lip, V., Tsai, J. M., Li, Y.-D., Fisch, A. S. et al. Outlier expression of isoforms by targeted or total RNA sequencing identifies clinically significant genomic variants in hematolymphoid tumors. *J. Mol. Diagn.* (2023) doi:10.1016/j.jmoldx.2023.06.007.
25. Zhou, X., Chen, X., Chen, J., Cao, P., Wang, F., Zhang, Y. et al. The Ingenious Multiple Efficacies of the Integrated Gene Fragment of Torque Teno Mini Virus in Orchestrating Promyelocytic Leukemogenesis Via Hijacking Rara. *Blood* **142**, 1576 (2023).
26. Kanagal-Shamanna, R., Luthra, R., Yin, C. C., Patel, K. P., Takahashi, K., Lu, X. et al. Myeloid neoplasms with isolated isochromosome 17q demonstrate a high frequency of mutations in SETBP1, SRSF2, ASXL1 and NRAS. *Oncotarget* **7**, 14251–14258 (2016).
27. Inamura, J., Ikuta, K., Tsukada, N., Hosoki, T., Shindo, M. & Sato, K. Acute Promyelocytic Leukemia with i(17)(q10). *Intern. Med.* **55**, 3341–3345 (2016).
28. Koczkodaj, D., Muzyka-Kasietczuk, J., Chocholska, S. & Podhorecka, M. Prognostic significance of isochromosome 17q in hematologic malignancies. *Oncotarget* **12**, 708–718 (2021).
29. Manola, K. N., Karakosta, M., Sambani, C., Terzoudi, G., Pagoni, M., Gatsa, E. et al. Isochromosome der(17)(q10)t(15;17) in Acute Promyelocytic Leukemia Resulting in an Additional Copy of the RARA-PML Fusion Gene: Report of 4 Cases and Review of the Literature. *Acta Haematol.* **123**, 162–170 (2010).
30. Duan, Y., Nie, J., Zhang, Z., Zhou, L., Zhu, F., Zhang, H. et al. A rare case with typical acute promyelocytic leukemia morphology associated with isolated isochromosome 17q without RAR α rearrangement. *Hematol. Oncol. Stem Cell Ther.* **6**, 42–45 (2013).
31. Jain, H., Shetty, D., Narula, G., Moulik, N. R., Tembhare, P., Patkar, N. et al. Cytogenetic profile and outcome of a pediatric acute promyelocytic leukemia patient presenting with isolated isochromosome 17q in absence of RARA rearrangement. *Blood Cells, Mol., Dis.* **88**, 102443 (2021).
32. Wen, L., Xu, Y., Yao, L., Wang, N., Wang, Q., Liu, T. et al. Clinical and molecular features of acute promyelocytic leukemia with variant retinoid acid receptor fusions. *Haematologica* **104**, e195–e199 (2019).

33. Fasan, A., Haferlach, C., Perglerová, K., Kern, W. & Haferlach, T. Molecular landscape of acute promyelocytic leukemia at diagnosis and relapse. *Haematologica* **102**, e222–e224 (2017).
34. Gurnari, C., Voso, M. T., Girardi, K., Mastronuzzi, A. & Strocchio, L. Acute Promyelocytic Leukemia in Children: A Model of Precision Medicine and Chemotherapy-Free Therapy. *Int. J. Mol. Sci.* **22**, 642 (2021).
35. Chen, Z., Brand, N. J., Chen, A., Chen, S. J., Tong, J. H., Wang, Z. Y. et al. Fusion between a novel Krüppel-like zinc finger gene and the retinoic acid receptor- α locus due to a variant t(11;17) translocation associated with acute promyelocytic leukaemia. *EMBO J.* **12**, 1161–1167 (1993).
36. Masi, A. di, Leboffe, L., Marinis, E. D., Pagano, F., Cicconi, L., Rochette-Egly, C. et al. Retinoic acid receptors: From molecular mechanisms to cancer therapy. *Mol. Asp. Med.* **41**, 1–115 (2015).
37. Miltiadous, O., Petrova-Drus, K., Kaicker, S., Mathew, S., Kluk, M., Geyer, J. T. et al. Successful treatment and integrated genomic analysis of an infant with FIP1L1-RARA fusion-associated myeloid neoplasm. *Blood Adv.* **6**, 1137–1142.

Figure legend

Figure 1: Manual review of Cases 1 and 2 in IGV. A) Case 1: Manual review in IGV of alignments over *RARA* exon 3 revealed split-reads with 3' alignments to exon 3 and a novel 5' soft clip pattern where the soft clips did not align to the human genome but were found to align by blastn to a TTMV isolate with high sequence identity. The primary split-read alignments in gray had unmapped paired reads that consisted predominantly or entirely of viral sequence. B) Case 2: Manual review in IGV of alignments within *RARA* intron 2 (left panel) and over exon 3 (right panel) revealed paired-end reads with primary and supplementary chimeric alignments consistent with aberrant cryptic splicing connecting a 26 bp segment within intron 2 (chr17: 38494366-38494391) to exon 3. Empirical analysis of historical fusion panel data showed this cryptic splice junction to be unique to this patient. The split-read alignments over the intron 2 segment moreover had 5' soft clips (left panel) that did not align to the human genome but were found to align by blastn to several TTMV isolates with high sequence identity.

Figure 2: Fusion contigs of Cases 1-4. A-B) RNA-based contigs for Case 1 (A) and Case 2 (B) were generated from BCH fusion panel data using Trinity, yielding lengths of 413 bp and 574 bp. The 5' ends of the contigs aligned best to the TTMV isolates MN768572.1 (Case 1; 96% identity) and MN770942.1 (Case 2; 98% identity), where both cases were characterized by 2 separate blastn alignments that overlapped in the same 7 bp sequence (AGATGAG). The remaining 3' ends of the contigs aligned with 100% identity to regions of *RARA*: a short segment within intron 2 (Case 2 only), the entire exon 3 (Cases 1-2), and the initial 60 bp of exon 4 terminating at the end of an assay primer (Cases 1-2). C-D) Merging of the blastn alignments in both Case 1 (C) and Case 2 (D) was most consistent with splicing out of 87-90 bp segments from the respective TTMV genomes, where the inferred splice acceptor sites directly preceded potential ATG start codons of open reading frames that were predicted by blastx to encode putative conserved domains from the TT viral ORF2 superfamily pfam02957. E-G) DNA-based contigs for Case 2 (E), Case 3 (F), and Case 4 (G) were generated from FMI fusion panel data using Velvet, yielding *RARA*::TTMV and TTMV::*RARA* contigs for Case 2 (lengths 454 bp and 441 bp) and Case 3 (402 bp and 341 bp) but only a TTMV::*RARA* contig for Case 4 (386 bp). All *RARA* portions of the contigs aligned with 100% identity to regions within *RARA* intron 2, consistent with viral integration of TTMV within *RARA* intron 2. The TTMV::*RARA* breakpoints on the TTMV side were within ORF2 for Cases 2-3, similar to most previously characterized cases, and within ORF1 for Case 4, similar to findings from our re-analysis of SRR14060871 (see Supplementary Figures 1B and 2B).

Figure 3: Pathologic features of Case 1. A-B) At initial diagnosis the debriement specimens demonstrated a sheet-like infiltrate of primitive cells with eosinophilic granular cytoplasm, ovoid to folded nuclei, dispersed chromatin, and variable small nucleoli, consistent with myeloid sarcoma (A). By immunohistochemistry the cells were positive for MPO (B) and negative for CD34 (B inset). C-E) Bone marrow involvement at the time of second relapse showed similar morphology with eosinophilic granular cytoplasm on the core biopsy (C). On the aspirate smear (D) the leukemic cells were large with ovoid to irregular nuclei, dispersed chromatin, variable pale nucleoli, and frequent azurophilic granules, including some heavily granulated forms. Flow cytometry (E) demonstrated that the leukemic cells had high forward and side scatter, dim CD45, and expressed CD33 with dim to negative HLA-DR, CD34, and CD117.

Figure 4: Pathologic features of Case 2. A) Peripheral blood smear demonstrated anemia, thrombocytopenia, and scattered large primitive myeloid cells with ovoid to irregular nuclei, dispersed chromatin, large nucleoli, and moderate cytoplasm with azurophilic granules. B-E) The bone marrow was involved by a sheet-like infiltrate of leukemic cells with eosinophilic granular cytoplasm and round to irregular nuclei on core biopsy (B-C). On aspirate smear (D) the leukemic cells were large with ovoid to folded nuclei, dispersed chromatin, variably prominent pale nucleoli, and frequent azurophilic granules, including some heavily granulated forms. Flow cytometry (E) demonstrated that the leukemic cells had increased side scatter and expressed CD45 (dim), CD34 (dim), CD33, CD117 (variable), with dim to negative HLA-DR.

Figure 1

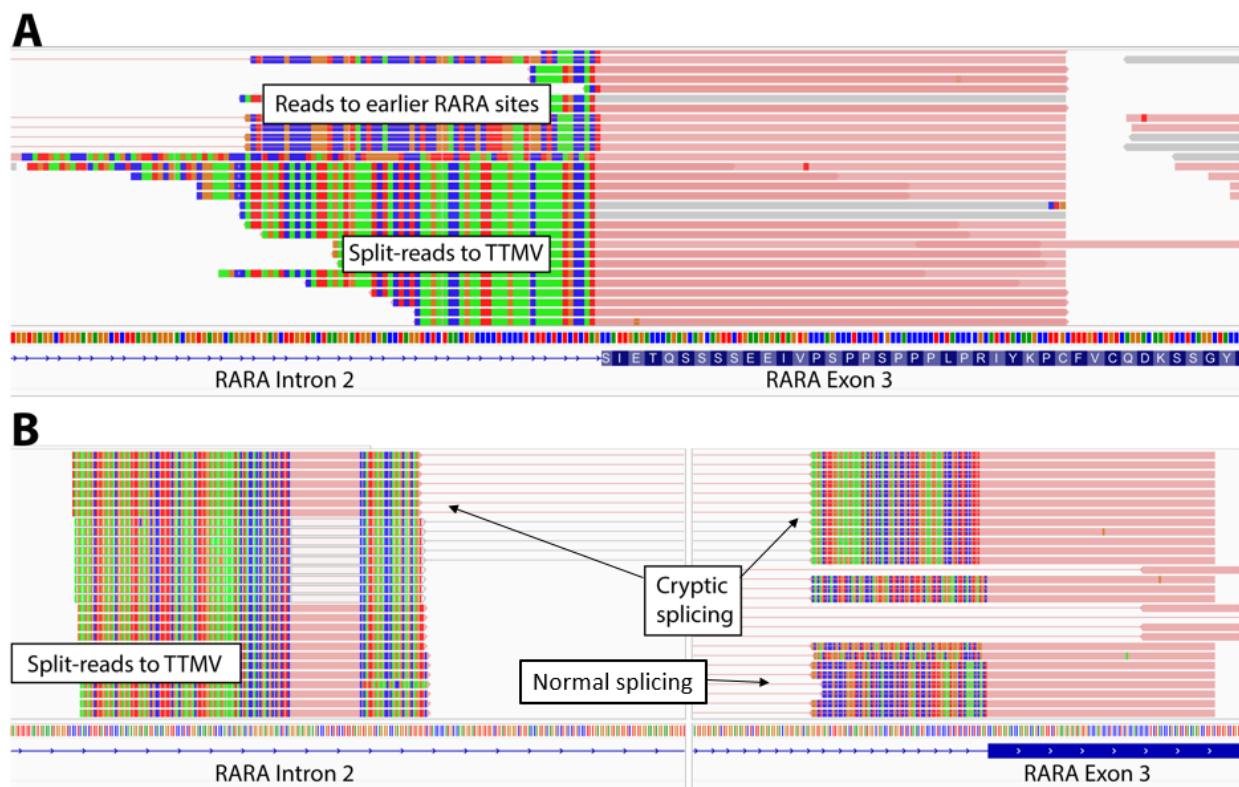


Figure 2

A. Case 1 (TTMV::*RARA* contig [413 bp] from RNA)

```
AGAGTCTGCAGGCTAATTCTGGGCGGGTGCCGAAGATGAGCAAATATGTACCAAAAAACCCCTAACACAAGAAATGGC
AGACGAGACTGGATTAACACAATTGTAATGATTCATGACCTCCAATGCGACTGCGGAAAGCCATTAGATCACACATT
AGATGAAATCTACTCTCAAGAACCAGAATTAAGAAACAAAAATGCCATGCCATTGAGACCCAGAGCAGCAGTTCTG
AAGAGATAGTGGCCAGCCCTCCCTCGCCACCCCTCTACCCCGCATCTACAAGCCTTGCTTTGTCTGTCTGAGACAAG
TCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGCTGCAAGGGCTTCTTCCGCCGCAGCATCCAGAAGAACAT
GGTGTACACGTGTACCCGGGACAAGAAC
```

TTMV MN768572.1 - 222-261* (39/40) [98%] + 345-515* (168/171) [98%]

RARA exon 3 - chr17:38504568-38504716 (149/149) [100%]

RARA exon 4 - chr17:38506036-38506095 (60/60) [100%]

*includes shared **AGATGAG** (7 bp) **Bold** = mismatch

B. Case 2 (TTMV::*RARA* contig [574 bp] from RNA)

```
GCCGGAGACGCGATAGCAACTATCAGCGGCTTAGCCTGGGCGGGTGCCGAAGATGAGCAAACCTTCTACCTATCAAAA
CTTCTGCTAGAAATCGCCGCCTAGACTGGATTAATGGAATTGTACACCTGCATAACCTTGTTTGTGGCTGCGACAAA
CCTCTAGAACACACCTTGGAAAGAAATCTACACCCAAGAACCTCAGCTTCATCCATATAAATAAAATATGCCCTACTAC
TACAAAAGACCACGGTACCACCGGCCTAGAAGAGCTTGGAGACGGAGATTTAGAGCGCCTTTTCGACGAAGACTTTG
GAGAAGAAAATACTACCACCGACGCCGCTTGGGAAGGAAGCCCGTCTTCTTTTAGCCATTGAGACCCAGAGCAGC
AGTTCTGAAGAGATAGTGGCCAGCCCTCCCTCGCCACCCCTCTACCCCGCATCTACAAGCCTTGCTTTGTCTGTCA
GGACAAGTCCTCAGGCTACCACTATGGGGTCAGCGCCTGTGAGGGCTGCAAGGGCTTCTTCCGCCGCAGCATCCAGA
AGAACATGGTGTACACGTGTACCCGGGACAAGAAC
```

TTMV MN770942.1 - 179-233* (53/55) [96%] + 314-602* (276/289) [96%]

RARA intron 2 - chr17:38494366-38494391 (26/26) [100%]

RARA exon 3 - chr17:38504568-38504716 (149/149) [100%]

RARA exon 4 - chr17:38506036-38506095 (60/60) [100%]

*includes shared **AGATGAG** (7 bp) **Bold** = mismatch

C. Case 1 (Merging of blastn alignments [MN768572.1])

```
AGAGTCTGCAGGCTAATTCTGGGCGGGTGCCGAAG-----
...CTTCAGAGTCTGCAGGCTAATTCTGGGCGGGTGCCGAAGGTGAGTGAACCACCGTAGTCAAGGGCCTAATGG
-----M S K Y V P K
-----ATGAGCAAATATGTACCAAAAA
GCAGAATTAGTCTTGGCGGAACGGGCTAGAAACTTAAAATATATTTCTTTTACAGATGAGCAAATATGTACCAAAAA
N P N T R N G R R D W I N T I V M I H D L Q C D C G
ACCCTAACACAAGAAATGGCAGACGAGACTGGATTAACACAATTGTAATGATTCATGACCTCCAATGCGACTGCGGA
ACCCTAACACCAGAAATGGCAGACGAGACTGGATTAACACAATTGTAATGATTCATGACCTCCAATGCGACTGCGGA
K P L D H T L D E I Y S Q E P E L K K Q K C H
AAGCCATTAGATCACACATTAGATGAAATCTACTCTCAAGAACCAGAATTAAGAAACAAAAATGCCATG
AAGCCATTAGACCACACTTTAGATGAAATCTACTCTCAAGAACCAGAATTAAGAAACAAAAATGCCATGGTAC...
```

BLASTX

Case1_RNA 17 PKNPNTRNGRRDWINTIVMIHDLQDCGKPLDHTLDEIYSQEP 61
Cdd:pfam02957 3 PPRYNPKGLELQWINSIVRSHDLFCGCGDPLKHLIHPILRQLPAL 47

D. Case 2 (Merging of blastn alignments to [MN770942.1])

```

GCCGGAGACGCGATAGCAACTATCAGCGGCTTAGCCTGGGCGGGTGCCGAAG-----
GTCGGAGACGCGATAGGAACTATCAGCGGCTTAGCCTGGGCGGGTGCCGAAGGTGAGTTTACCACCGGAGTCAAGGG

-----M S K L L
-----ATGAGCAAACCTTCTA
|||||
GCAATTCGGGCAGGCTAAGTCTGGCGGAACGGGCAAGAACTTAAAATATTGTCTTTTACAGATGAGCAAACCTTCTA

P I K T S A R N R R L D W I N G I V H L H N L V C G
CCTATCAAAACTTCTGCTAGAAAATCGCCGCCTAGACTGGATTAATGGAATTGTACACCTGCATAACCTTGTGTTGTGG
CCTATTAACCTTCTGCTAGAAAATCGCCGCCTAGACTGGATTAATGGAATTGTACACCTGCACAACCTGGTTTGTGG

C D K P L E H T L E E I Y T Q E P Q L H P Y N K I
CTGCGACAAACCTCTAGAACACACCTTGAAGAAATCTACACC CAAGAACCTCAGCTTCATCCATATAAATAAAATAT
CTGCGACAAACCTTTAGAACACACCTTGAAGAAATCTACACTCAAGAACCTCAACTACATCCATACAACAAAATAT

C P T T T K D H G T T G L E E L G D G D L E R L F D
GCCCTACTACTACAAAAGACCACGGTACCACCGGCCTAGAAGAGCTTGGAGACGGAGATTTAGAGCGCCTTTTCGAC
GCCCTACTACTACAAAAGACCACGCTACTACCGGCCTAGAAGAGCTTGGAGACGGAGATTTAGAGCGCCTATTTCGAC

E D F G E E N T T T D A A
GAAGACTTTGGAGAAGAAAATACTACCACCGACGCCGCTTC
GAAGACTTTGGAGAAGAAAATACTACCACCGACGCCGCTTCTACAAGCGGG...

```

Blastx

```

Case2_RNA      23 PIKTSARNRRLDWINGIVHLHNLVCGCDKPLEHTLEEIYTQEPQL 67
Cdd:pfam02957 3  PPRYNPKGLELQWINSIVRSHDLFCGCGDPLKHLIHPILRQLPAL 47

```

E. Case 2 (*RARA*::TTMV and TTMV::*RARA* contigs [454 bp and 441 bp] from DNA)

```

AGGGACAGCCCAGGTGGAGGGAGTGCTGCCTAGCAATGCCTTTCTTGGAGTCCGAAATGTAGAGTCAGACAAAAGAA
CATGACATGGAGTTTTTAATGAAAATGGGGTCCCTCACCTTTTATCTCCCAGCAAGGATAATGCACAGGTTCCCTCT
GGACAAGATCCCAGCATAGACGAACATCCTACAACCAAATTCGTTGCAGAGTCCAACAACACCCTCGAAAAATTCC
TCTACTCATTTGACTGGAGACGAGGTTACCTTACACAAAAGCTACTGCAAGAATTTAGCAGACGAAAGCACTGAA
AAAAATGTTTTTGGACCTACAGGATTCAACCTCTTCAATCCCAACCCACACAAGACTCATCGGAGGAGACAGAAAC
AGAGGAGAAAAGCGAAAAGAGACACTACAGATCCTCCTCAAGCAGCTCCGACAGCAACAACACAGATA

```

RARA intron 2 - chr17: 38494175-38494350 (176/176) [100%]

TTMV MN770942.1 - 2193-2470 (275/278) [99%]

```

CCTCAGCTTCATCCATATAAATAAAATATGCCCTACTACTACAAAAGACCACGGTACCACCGGCCTAGAAGAGCTTGG
AGACGGAGATTTAGAGCGCCTTTTCGACGAAGACTTTGGAGAAGAAAATACTACCACCGACGCCGCTTCGGGAAGGA
AGCCCGTCTTCTTTTAGGTGGGTGAGGGCCGAGGTATTGTGTCTACCCCATGCCAAGGGATGTGTTGCTTATGGGG
CTGTACCAGTGCAGGCCTGCCTGCCCTTTGAGGGAGTAGTGTCTGTGGGAAGAGGAAGTGGTTTCCATCTCTAAGTC
ATTTTGGGGCAGTGCCTTCTTCCCTCCTCCTCTGTGGAGTTGGGGCACAGCGAGGGATTGGGCCTCCAGGGTTG
ACTGGAGGCCACTCAACCTTTGGCTGTCAGAAATAGAGGTGCTCCCCTCCCCATTT

```

TTMV MN770942.1 - 457-602 (139/146) [95%]

RARA intron 2 - chr17: 38494366-38494660 (295/295) [100%]

F. Case 3 (*RARA*::*TTMV* and *TTMV*::*RARA* contigs [402 and 341 bp] from DNA)

```
CCCCCCCCCACCAGTCTGGATTGTCTATTGTTACTGCTTTTACGTCTTGGAAAAAGTTAGCACAACAAAGGGCTG
CTTTGTGGCTCACCCCTCTGCCTCCTGGCCTCACCCAGGCCCCCCAACCCCGCCCCCAGCAGCTGTTCTCAGGC
CTCTCAGCCTGTCTGATTTGCTTGTCTGGCCTGGGGAGAATGAGGTGGGAGAAAACCAGGCCAGGGCAGTTGGTGT
GGAGTGAAGAGCAGACGGCGGTGGTGGCTGTTACCTCCATCAAATCAATTAGCGACCCTGAAAACATGCCAAAA
CTACCGAATCCTATAATATCCTACAACCAAATTCGTTGCAGAGTCCAACAACACCACTCGAAAAATTCCTCTACTCA
TTTGACTGGAGACGAGG
```

 RARA intron 2 - chr17:38502765-38503019* (255/255) [100%]

 TTMV MN771618.1 - 2065-2214* (149/150) [99%]

*includes shared TGG (3 bp)

```
GGCTAAGTCTGGCGGAACGGGCAAGAACTTAAAATAATTTCTTTTACAGATGAGCAAACCTTCTACCAATTAACCT
TCTGCTAGAAATCGCCGCCTAGATTGGATTAATGGAATTGTACACCTGCACAACCTGGTGTGTGGCTGCGACAAACC
CCTAGAACACACCTTAGAAGAAATCTACACTCAAGAACCTCAACTTCATCCATACAACAGACGGCGGTGGGGAGGTC
AGGAGAGAATCTGCTGGGCTGGGGATGGTGTGGGCATCAACTGTCCCATTGCTGCAGGCTGGTCTTGGGGCAGGGAA
GGGGATGGGGGGCCATAGCAGTGTGGTTCAGCC
```

 TTMV MN771618.1 - 208-420* (209/213) [98%]

 RARA intron 2 - chr17:38503007-38503136* (130/130) [100%]

*includes shared CA (2 bp)

G. Case 4 (*TTMV*::*RARA* contig [386 bp] from DNA)

```
AAAAATTCCTCTACTCATTGACTGGAGACGAGGTTACCTTACACAAAAAGCTACTACAAGAATTTAGCAGACGAA
AGCACTGAAAAAATGTTTTGGACCTACAGGATTCAACCTCTTCAATCCACAACCCACACAAGATTCATCGGAGGA
GACAGAAACAGAGGAGAAAAGCGAAAAAGAGACATTACAGATCCTCCTCAAGCAGCTCCATTGCTGCAGGCTGGTC
TTGGGGCAGGGAAGGGGATGGGGGGCCATAGCAGTGTGGTTCAGCCAGGCTGGCCTGGGAAGTGGTGGCCAGGCACT
ACTAAGAGCCAGGAAAGCCCTGCCAAGGTTGTTGGCCTAGTTCCTGTTCATCAGCCGCCTAGCAGCCCCACTGTGT
C
```

 TTMV MN771921.1 - 2215-2427* (213/213) [100%]

 RARA intron 2 - chr17:38503070-38503245* (176/176) [100%]

*includes shared TCC (3 bp)

Figure 3

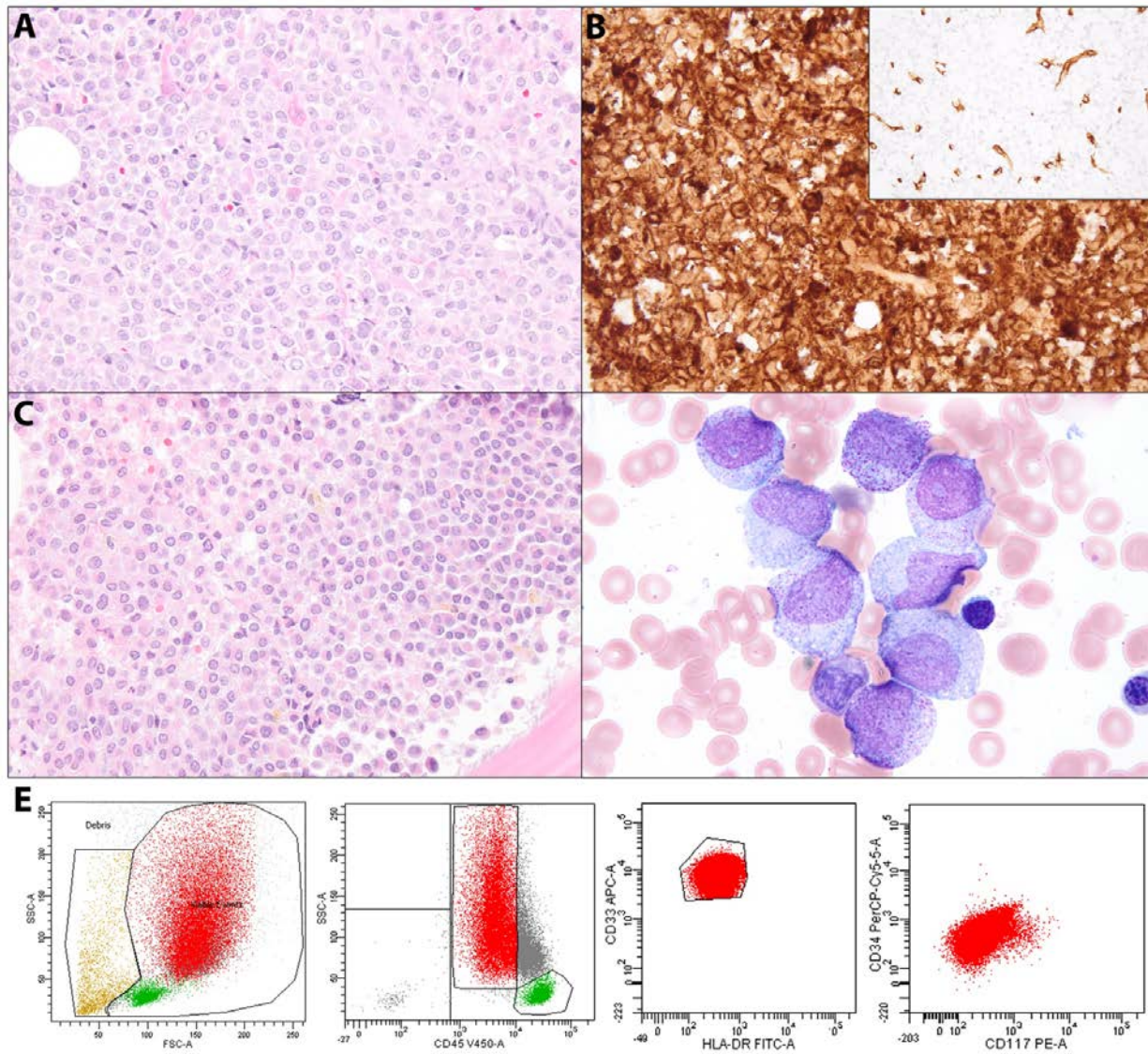


Figure 4

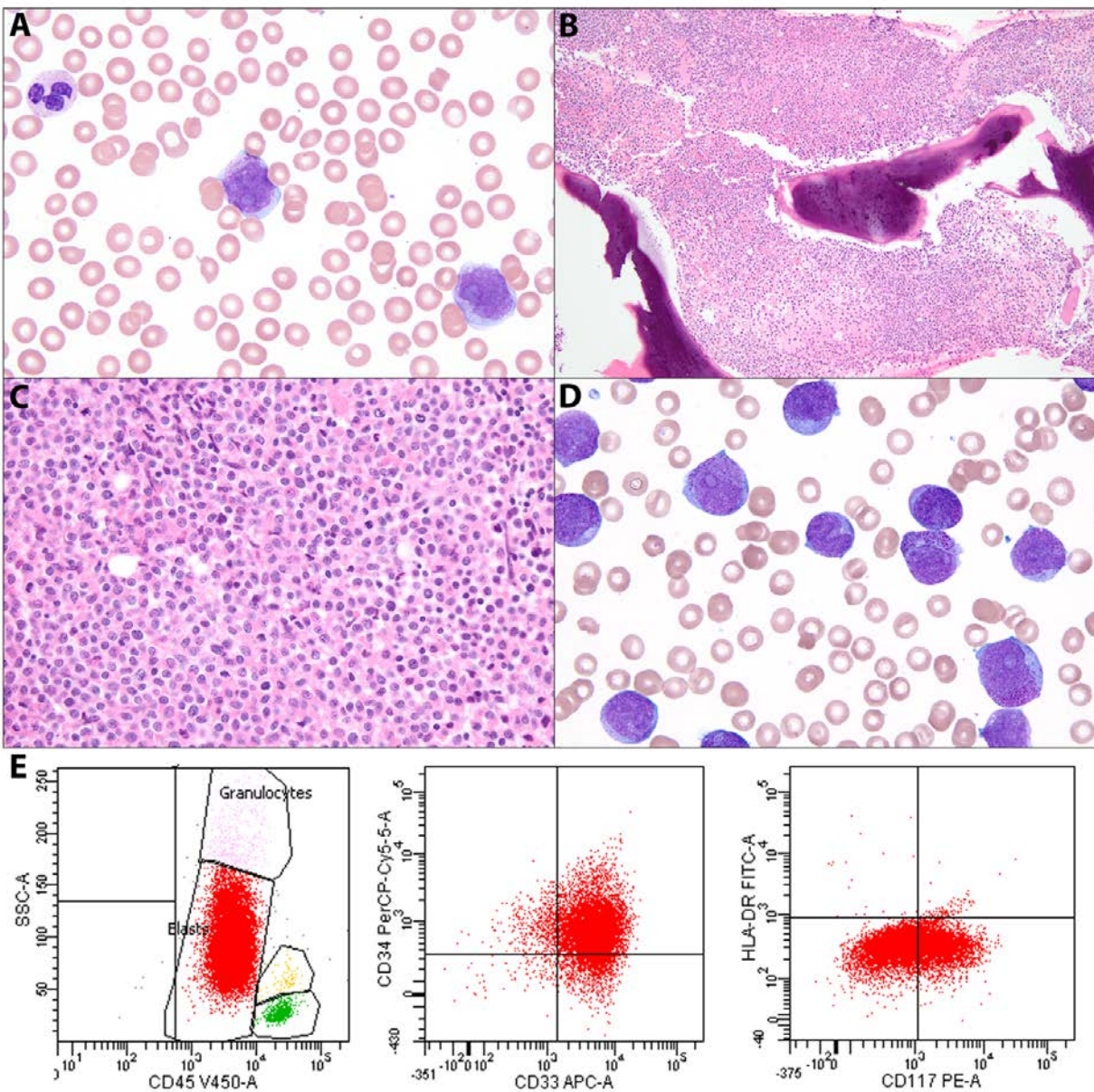


Table 1: Number of TTMV::RARA split-reads (and discordant reads)

GenBank_id	Case1_RNA	Case2_RNA	Case2_DNA	Case3_DNA	Case4_DNA	SRR14060870	SRR14060871
MN768572.1	67 (14)						
MN771921.1		4472 (1911)	15 (590)	8 (96)	9 (209)		
MN774502.1		4472 (1911)	15 (632)	8 (99)	9 (204)		
MN771618.1		4472 (1911)	15 (632)	8 (99)	9 (204)		
MN772671.1		4472 (1911)	15 (632)	8 (97)	9 (204)		
MN770942.1		4475 (2272)	15 (646)	4 (70)	9 (204)		
MN771042.1		4472 (1911)	15 (632)	4 (77)	9 (204)		
MN771954.1		895 (1163)					
MN770247.1		2 (0)	9 (101)				
MN770240.1		2 (0)	9 (101)				
MN773107.1			9 (88)				
MN773565.1			9 (88)				
BK032512.1			9 (79)				
MN772994.1			9 (88)				
MN771972.1			9 (101)				
MN770927.1			9 (101)				
MN770954.1			9 (101)				
MN770239.1			9 (101)				
MN770241.1			9 (101)				
MN770932.1			9 (101)				
MN771787.1						355 (1269)	110 (277)
MN774000.1						354 (1269)	110 (277)
MN769771.1						354 (1269)	110 (277)
MN771920.1						354 (1273)	110 (277)
MN768731.1						351 (1266)	110 (277)
MN772835.1						313 (1071)	110 (276)
MN771783.1						71 (1260)	97 (276)

Shaded cells are the isolates for which the assembled contig had the highest blastn score. Of note, it generally was also associated with the highest number of blastn alignments. The split-reads and discordant reads are from alignments to TTMV isolates by blastn and to the human genome by bwa-mem.

Table 2: Clinical features and outcome of APL with TTMV::RARA fusion (continued on next page)

Reference	Clinical parameters						Bone marrow involvement	CSF involvement
	Age (yrs)	Sex	CBC at diagnosis	Coagulopathy	Presentation	Myeloid sarcoma		
Current study Case 1	11-20	M	Normal	Not present at initial diagnosis. DIC developed after relapse.	Hearing loss and imaging showing multifocal bone and soft tissue of the head.	Yes, yes bone and soft tissue of head	Not at initial diagnosis At relapse	+
Current study Case 2	0-5	F	WBC 17, Hgb 9.9, PLT 44, 82% blasts by flow cytometry	Slightly prolonged PT/INR Slight coagulopathy after initiation of ATRA	Hip pain	No	At initial diagnosis	+
Current study Case 3	0-5	M	WBC 4.7, Hgb 8.4, PLT 146, no blasts	No	Leg & hip pain, recurrent fevers, weight loss	No	At initial diagnosis	-
Current study Case 4	21-30	F	WBC 10.1, Hgb 10.7, PLT: 223, no blasts	No	Large destructive lesion of distal femur and lymphadenopathy. Subsequent kidney involvement.	Yes, femur, lymph nodes, kidney	No	-
Astolfi et al. 2021 Case 1	6	F	Hgb 9.7, WBC 2.8, + PLT 101	+	Shoulder pain, fever, pallor, ecchymosis	No	At initial diagnosis	Unknown
Astolfi et al. 2021 Case 2	3	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown
Sala-Torra et al. 2022	39	M	Unknown	+	Thrombocytopenia, easy bleeding	No	At initial diagnosis	Unknown
Chen et al. 2022	3	M	Hgb 8.6, WBC 36.8, PLT 117	Slightly prolonged PT	Persistent fever and headache	Yes, at relapse ("multiple metastases")	At initial diagnosis	Unknown
Chen et al. 2023 PBC	11	F	Unknown	Unknown	Pelvic pain, intermittent fever	No	At initial diagnosis	Unknown
Chen et al. 2023 IJLH	7	M	Unknown	At time of relapse	Back pain, leg numbness, lumbar spinal lesion on MRI	Yes, spinal lesion	Not at initial diagnosis At relapse	Unknown

HSCT: hematopoietic stem cell transplant. NA: Not applicable.

Table 2 continued on next page

Table 2 (continued): Clinical features and outcome of APL with TTMV::RARA fusion

Reference	Treatment/Outcome			
	Initial Therapy	Induction failure/Relapse	Secondary therapy	Survival
Current study Case 1	Standard dose cytarabine, daunorubicin, & etoposide	Relapse #1: 3 months after diagnosis Relapse #2: 13 months after diagnosis, 7 months after 1st HSCT Relapse #3: 24 months after diagnosis	After relapse #1: High dose cytarabine and mitoxantrone, radiation therapy, HSCT following busulfan & cyclophosphamide conditioning. After relapse #2: Daunoxome & FLAG, 2nd HSCT following cyclophosphamide/TBI conditioning	Dead of disease
Current study Case 2	Initial ATRA therapy, discontinued after negative PML/RARA Low dose cytarabine, daunorubicin with dexrazoxane, gemtuzumab, and intrathecal methotrexate, cytarabine, and hydrocortisone	Suboptimal response to induction	Induction II with idarubicin, fludarabine, ara-c, G-CSF, and ATRA resulting in complete remission, HSCT following busulfan, fludarabine, and clofarabine conditioning	Alive, in remission after HSCT
Current study Case 3	Cytarabine + daunorubicin + etoposide	Induction failure	Cytarabine + mitoxantrone followed by HSCT.	Alive, in remission after HSCT
Current study Case 4	Bone resection & reconstruction Induction with ytarabine + daunorubicin Consolidation with HiDAC Core blood HSCT	No	NA	Alive, in remission 2 years after HSCT
Astolfi et al. 2021 Case 1	ATRA and standard AML induction chemotherapy, 3 high-dose cytarabine-based courses of consolidation chemotherapy	Relapse, 8 months after end of treatment	2 courses of ATRA and ATO with complete morphologic remission on day +21, followed by HSCT	Alive 4 years after HSCT
Astolfi et al. 2021 Case 2	Unknown	Unknown	Unknown	Unknown
Sala-Torra et al. 2022	Standard 7 + 3 cytarabine + daunorubicin regimen	Induction failure, day 15	Mitoxantrone, etoposide, cytarabine, decitabine, venetoclax (with response)	Unknown
Chen et al. 2022	ATRA and induction chemotherapy with daunorubicin, cytarabine, and etoposide. ATRA discontinued at day 7 due to absence of PML::RARA. Consolidation with idarubicin, cytarabine, and etoposide	Relapse, 1 month after achieving CR	Re-induction with haploidentical HSCT. 2nd relapse 50 days later with "multiple metastases"	Dead of disease
Chen et al. 2023 PBC	ATRA + idarubicin	Relapse, 15 months after diagnosis, 3 months after treatment	Multiple regimens without remission. Eventual remission after ATRA + azacitidine + venetoclax, followed by allo-HSCT	In remission at 20 months after HSCT
Chen et al. 2023 IJLH	6 cycles daunorubicin, cytarabine and etoposide, then 4 cycles of cytarabine with remission.	Relapse 6 months after treatment	Multiple regimens without remission, including in combination with ATRA.	Dead of disease

HSCT: hematopoietic stem cell transplant. NA: Not applicable.

Table 3: Pathologic features of leukemias with TTM::RARA

Reference	Morphology	Immunophenotype	Karyotype	FISH	Molecular/NGS
Current study Case 1	Atypical promyelocytes with frequent eosinophilic/azurophilic granules No Auer rods	Positive: MPO, lysozyme, CD43, CD33, CD13 (bright), CD64, CD56, CD4 (dim), CD71 (dim) Partial: CD34 (dim/-), CD117 (dim/-), CD15 (dim/-) Negative: HLA-DR, CD16, CD11b, CD7, CD5, CD41, CD61, CD10, CD20, CD19	46,XY,add(7)(q22)[5]/46,XY[15] (at time of marrow relapse)	AML & MDS FISH panels (including PML/RARA) negative	<i>NRAS</i> c.35G>T (p.G12V) Copy number analysis with loss of genes on 7q.
Current study Case 2	Atypical promyelocytes with frequent azurophilic granules No Auer rods	Positive: MPO, CD33, CD13 (bright), CD38, CD64, CD123, CD38 Partial: CD117, CD34 (dim/-), CD11c (dim/-), CD71 (dim/-) Negative: HLA-DR, CD3, CD4, CD5, CD7, CD8, CD10, CD11b, CD15, CD16, CD19, CD20, CD36, CD41, CD61, CD56, CD65	46,XX,dic(11;17)(p11.2;p11.2)[14]/46,XX,idelic(17)(p11.2)[4]/46,XX[2]	PML-RARA & RARA breakapart FISH negative AML & MDS FISH panels negative	<i>WT1</i> c.1142C>A (p.S381*) 2% VAF Copy number analysis with loss of genes on 11p and 17p, and gains on 17q
Current study Case 3	Atypical promyelocytes with frequent azurophilic granules Rare Auer rods	Positive: MPO, lysozyme, CD33, CD13 Partial: CD117, Tdt (focal by IHC) Negative: HLA-DR, CD34, CD1a, CD2, CD15, CD25, CD35, CD42b, CD61, CD163, langerin, mast cell tryptase, B-cell and T-cell markers	46,XY,der(7)ins(7;8)(q22;q13q22)[4]/26,XY[20]	PML-RARA & RARA breakapart FISH negative Extra copy of RUNX1T1 (8q)	No pathogenic or likely pathogenic variants.
Current study Case 4	Atypical promyelocytes No Auer rods	Positive: Flow: CD13 (bright), CD33 (bright), CD117, CD38, CD64. IHC: MPO, CD43, CD68, CD99 Negative: Flow: HLA-DR, CD34, CD14, CD15, CD36, CD56	Failed	Failed	Pathogenic: KMT2C p.Q2635* WT1 p.R369* Likely pathogenic: ATR splice site 6320-2A>C
Astolfi et al. 2021 Case 1	Atypical promyelocytes with azurophilic granules and Auer rods	MPO+ CD33+ CD13+ CD38+ CD99+ HLA-Dr low	Normal	RARA breakapart negative	Unknown
Astolfi et al. 2021 Case 2	Unknown	Unknown	No cytogenetic abnormality of pathogenic mutation		
Sala-Torra et al. 2022	APL-like with increased promyelocytes	MPO+ CD33+ CD34- HLA-DR-	46,XY,i(17)(q10)[18]/47,XY,+8,i(17)(q10)[2]	PML-RARA FISH negative	Unknown
Chen et al. 2022	Hypergranular promyelocytes No Auer rods	MPO+ CD33+ CD13+ CD117+ CD123+ CD64+ CD34partial HLA-DR- CD11b- CD16-	Normal	PML-RARA FISH negative	FLT3-ITD mutation RT-PCR negative for <i>PML::RARA</i> Whole transcriptomic screening for oncogenic fusion genes negative
Chen et al. 2023 PBC	Atypical promyelocytes	Reported as APL-like immunophenotype	47,XX,+21[10] at diagnosis 47,XX,+21[8]/48,XX,+8,+21[1]/46,XX[11] at relapse	Unknown	<u>Diagnosis:</u> RT-PCR negative for <i>PML::RARA</i> <u>Relapse:</u> Gene expression profile similar to APL with <i>PML::RARA</i> . <i>RARA</i> c.1160T>C/p.1387T 68% VAF <i>ARID1B</i> c.2837_2838insGT/p.A947Sfs*4 26% VAF
Chen et al. 2023 IJLH	Aberrant promyelocytic cells	Flow cytometry reported as APL-like immunophenotype. IHC: CD33+ MPO+ CD34- Cytochemical stains: MPO+ chloroacetate esterase+, specific and nonspecific esterase+	At relapse: 46,XY,dup(17)(q23q25)[15]/45,X,-Y,der(16)t(Y;16)(q12;q22),dup(17)(q23q25)[4]/45,X,-Y,del(4)(p14),der(16)t(Y;16)(q12;q22),dup(17)(q23q25)[1]	RARA breakapart negative	<u>Relapse (bone marrow):</u> RT-PCR negative for <i>PML::RARA</i> . Gene expression profile similar to APL with <i>PML::RARA</i> . WT1 c.1117_1118insCAGGATG/p.V373Afs*19 (not present in the prior myeloid sarcoma) <u>Germline:</u> SAMD9 c.153dupA/p.E52Rfs*

IHC: Immunohistochemistry.