Characterization of complex hereditary cancer associated germline variants with long read sequencing

Ambry Genetics®

Sami Belhadj, PhD[1*]; Felicia Hernandez, PhD[1]; Terra Brannan, PhD[1]; Colin Young, PhD[1]; Jesus Ramirez Castaño, BS[1]; Cassidy Carraway, BS[1]; Jessica Grzybowski, MS-CGC[1]; Carrie Horton, MS-CGC[1]; Marcy Richardson, PhD[1]; Rachid Karam, MD, PhD[1]

[1] Ambry Genetics. 1 Enterprise, Aliso Viejo, CA 92656 | * Email: sbelhadj@ambrygen.com

BACKGROUND

Long read sequencing (LRS) is emerging as a key technology to elucidate complex variants that are missed and/or not fully characterized by conventional short read (SR) sequencing approaches. Among these, structural variants and mobile elements are particularly challenging to resolve. We sought to apply LRS to resolve several complex germline variants (**Results I**) and to explore missing heritability in cases suspected to have Lynch Syndrome (LS; **Results II**).

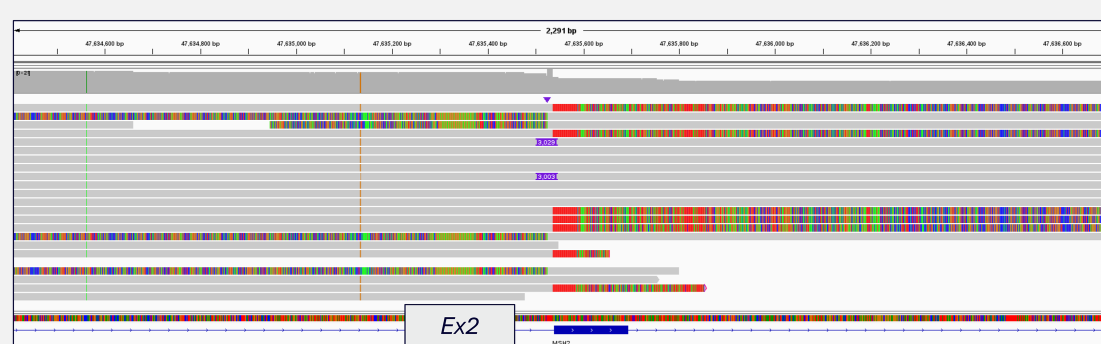
METHODS

DNA samples from patients with inconclusive results from multigene panel testing underwent multiple LRS-based strategies including long read genome sequencing (LR-GS) and targeted DNA sequencing (T-LRS) using a custom Twist panel on the PacBio Sequel IIe. Short read DNA and RNA sequencing data was available for all studied cases as part of clinical testing.

RESULTS I: Using LRS to resolve complex germline variants

A. Characterizing an SVA insertion in *MSH2* that causes skipping of Exon 2

Using LR-GS, we fully characterized a 3Kb SINE-VNTR-Alu (SVA) insertion near the acceptor site of *MSH2* exon 2 (c.212-15_212-14insSVA) in a family fulfilling Amsterdam II criteria.

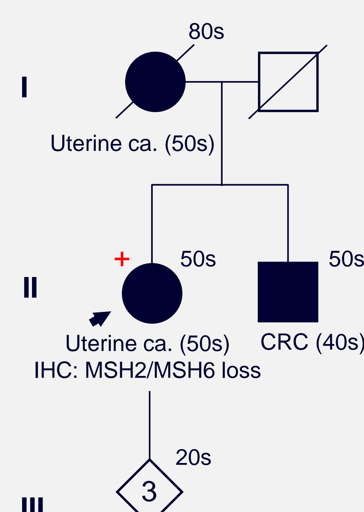


I — 80s; Uterine ca. (50s)
II — + 50s; Uterine ca. (50s) IHC: MSH2/MSH6 loss; 50s CRC (40s)
III — 3; 20s

Above: Reads supporting the SVA insertion, displayed by IGV as soft-slipped reads or by a purple box. **Right:** Pedigree of the family fulfilling Amsterdam II criteria. Cancer diagnosis ages are noted after cancer type. Current age or age at passing denoted in top-right corner. Carrier marked by a red "+" sign.

B. Refuting a *PMS2* EX11dup identified by MLPA

PMS2CL is a pseudogene with high sequence identity to *PMS2* leading to variant calling challenges. We performed T-LRS and separated PMS2/PMS2CL haplotypes using Paraphase[1]. This revealed a gene conversion event that made the *PMS2*–specific probe #364 (SALSA MLPA Probemix P008) bind to *PMS2CL* resulting in a false positive MLPA finding.

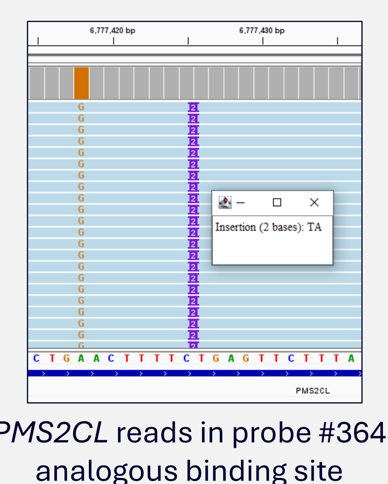


PMS2CL reads in probe #364 analogous binding site

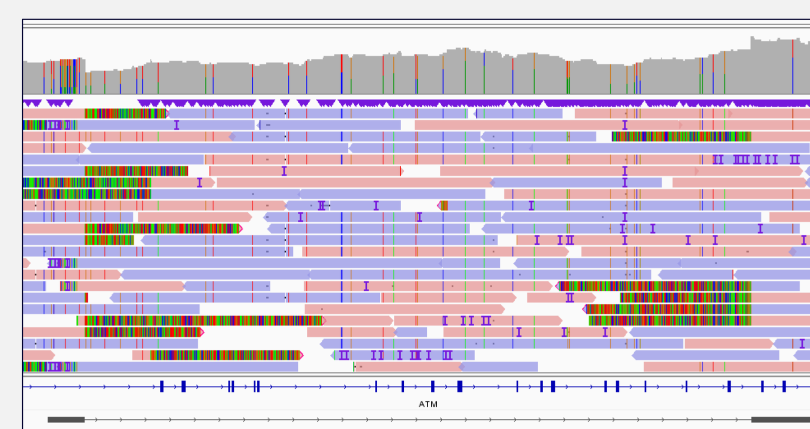
CTGGACTTTTCTATGAGTTCTTTA	*Probe #364 PMS2 binding site (24nt) – **PMS2 probe binds***
CTGAACTTTTC--TGAGTTCTTTA	*Analogous region for probe #364 in PMS2CL (22nt) – **PMS2 probe does not bind***
CTGGACTTTTCTATGAGTTCTTTA	*PMS2CL sequence after conversion event (24nt) – **PMS2 probe binds***

C. Resolving an *ATM* rearrangement that causes skipping of Exons 16-34

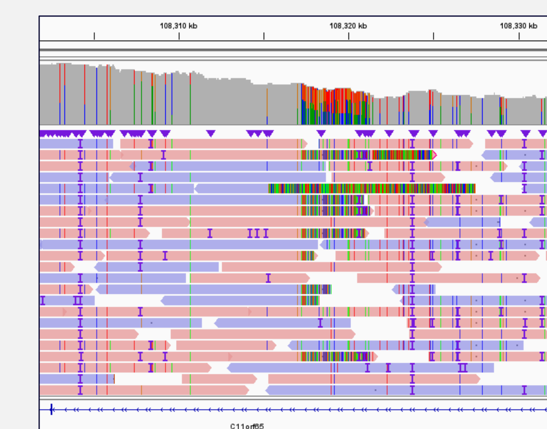
ATM EX34-3'UTRdup was identified by MLPA and classified as VUS. We used LR-GS and T-LRS to fully resolve a del/dup event consisting of a ~38kb deletion with breakpoints in *ATM* Introns 16-35, and a duplication with breakpoints in *ATM* intron 16 and C11orf65 Intron 2 (located 80kb downsteam of *ATM*).



This complex re-arrangement is expected to result in an NMD prone transcript and was re-classified to likely pathogenic (LP).



Deletion in *ATM* with breakpoints in introns 16 and 35 (chr11:108133562-108171771)
Detected by both LR-GS and T-LRS



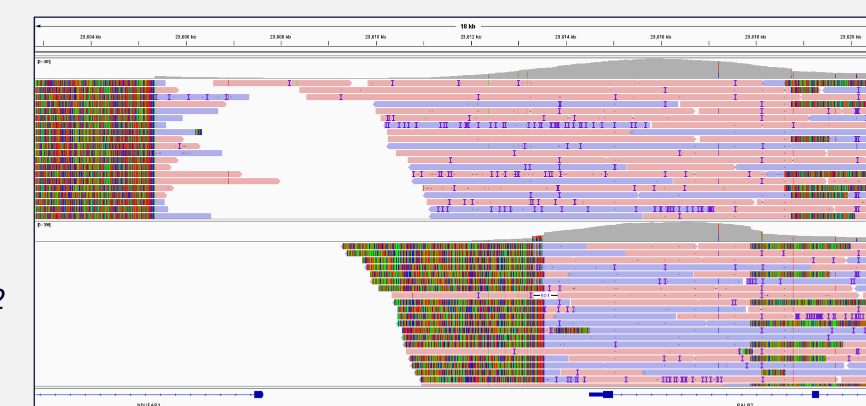
Duplication breakpoint (chr11:108321504), supporting C11orf65 intron 2- *ATM* intron 16 junction
Detected by LR-GS only.

D. *PALB2* Ex13 duplications have breakpoint-specific effects on RNA splicing

Previously described in the literature[2,3], conventional SR DNA-only testing is limited in distinguishing the two different alterations. Using T-LRS, we were able to delineate all breakpoints. The results are consistent with their effect on RNA.

Asian patients[2] (~13.4kb)
5'breakpoint in *NDUF1* intron 1
Chr16:23605287-23618695

Caucasian patients[3] (~4.6kb)
5' breakpoint downstream of *PALB2*
Chr16:23613291-23617887

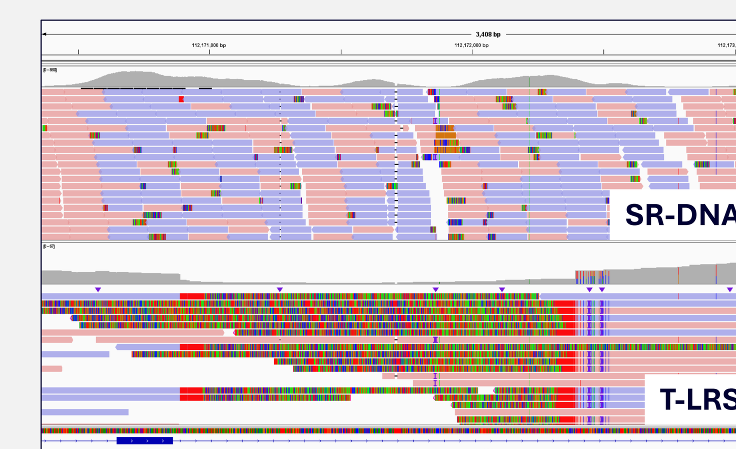


No effect on RNA

Abnormal RNA splicing (duplication of EX13 first 83bases, in tandem. LP)

E. *APC* Alu-driven recombination event resulting in Exon 14 skipping

Several cases presented with *APC* Ex14 skipping, where the underlying DNA variant couldn't be fully resolved by short-read approaches. Using T-LRS, we identified a previously described Alu-driven recombination event [4].



The event resulted in a ~1.5kb deletion (chr5:112170890-112172393), with the 5' breakpoint within ~23bp from Ex14 splice donor site.

RESULTS II: Using T-LRS to explore missing heritability

We successfully identified a *PMS2* EX1 inversion in a case suspected to have LS (IHC supporting PMS2 loss) that received previously negative germline testing. Allele specific expression analysis showed complete monoallelic expression in RNA, consistent with the inversion disrupting the transcription start site.

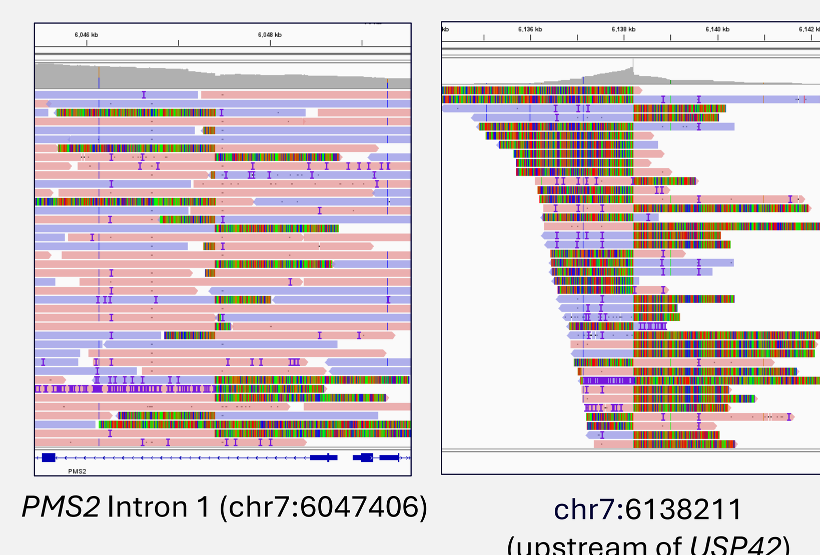
A. Study design

LS missing heritability cohort (n=17)
- 15/17 diagnosed with CRC/Endometrial cancer; 3/17 diagnosed with other LS-related cancer types (Dx age 38-77, median 50)
- 13/17 with ≥1 FDR and 4/17 with ≥2 SDRs diagnosed with LS-related cancers
- All cases with negative clinical testing and abnormal IHC results

T-LRS (*Hybrid capture) MLH1, MSH2, MSH6, PMS2 and EPCAM*) →

PMS2 EX1 INV (90.8kb)
Identified in 1/17 patients

B. Breakpoints of the 90.8kb inversion



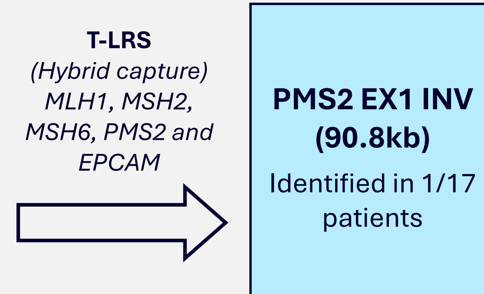
PMS2 Intron 1 (chr7:6047406)

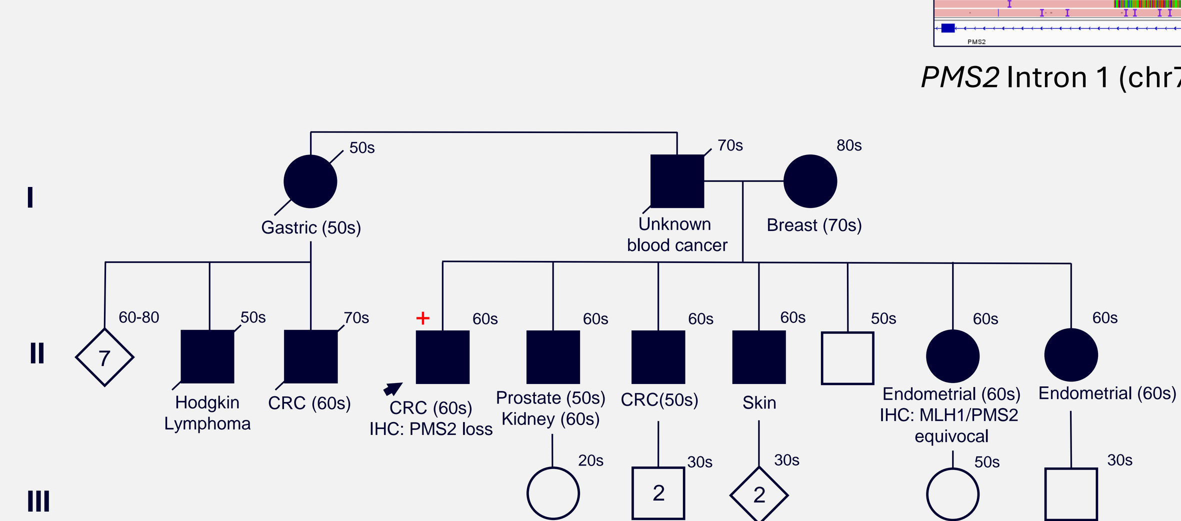
chr7:6138211 (upstream of *USP42*)

C. Allele specific expression

Ex 7: c.780C>G   Ex 11: c.1408C>T

D. Pedigree





D: Pedigree of the family where *PMS2* EX1INV was identified. Cancer diagnosis ages are noted after cancer type. Current age or age at passing denoted in top-right corner. Proband is also a carrier of *NBN* p.R43* (VUS)

TAKE HOME POINTS

1. LRS increases the accuracy of germline genetic testing by providing better characterization of complex variants and identifying alterations that were previously missed.
2. LRS is a useful complement to conventional assays such as short read DNA and RNA sequencing.
3. Compared to LR-GS, T-LRS applications are cost effective, allow testing of multiple cases simultaneously and increase read depth.

REFERENCES

1. https://github.com/PacificBiosciences/paraphase
2. Kwang A et al. JCO Precision Oncology 2021 (PMID: 34994627)
3. Yang C et al. Breast Cancer Res. Treat. 2016 (PMID: 27757719)
4. Tuohy TMF et al. Hum. Genetics 2010 (PMID: 20033212)