Identification of an Alu insertion in *MSH2* by Next-Generation Sequencing in a Family with Lynch Syndrome: An 8-year Diagnostic Odyssey

Amanda Jacquart[1], Stefano Rosati[2], Donald Basel[2], Jessica Grzybowski[3] Michael Muriello[2]

The Medical College of Wisconsin[1], Genomic Sciences & Precision Medicine Center (GSPMC)[2], Ambry Genetics[3]

Lynch syndrome (LS) is a well-known cause of hereditary colon cancer. Pathogenic and likely pathogenic variants in one of the mismatch repair (MMR) genes (*MLH1*, *MSH2*, *MSH6* and *PMS2* along with deletions of *EPCAM)* are known to cause LS. Alu insertions are retroelements (also referred to as "jumping genes") which have been shown to cause disease by either disrupting a coding region or a splice signal. Retroelements have been observed in cancer predisposition genes and were recently reported to be more common (1/325; 0.3%) than previously estimated (1/600; 0.16%) (Qian Y et al. Cancer Genet., 2017). There have been previous reports of Alu insertions in MMR genes in families with LS (Kloor et al. Hum. Genet., 2004, Solassol et al. Hum. Mutat., 2019).

We report a 16-year-old female who was diagnosed with stage 4 colon cancer and was posthumously found to have an Alu insertion in the *MSH2* gene. Her clinical history was unremarkable leading up to her diagnosis and she passed away 10 months later. Her colon tumor specimen showed abnormal microsatellite instability (MSI) and loss of protein expression of MSH-2 and MSH-6 by immunohistochemistry (IHC). Her mother had a history of multiple colon polyps starting in her mid-20s. Results from a LS screen performed on a colon tubular adenoma with focal high-grade dysplasia revealed abnormal MSI and the same absent protein expression by IHC. Her mother also had a history of a sebaceous adenoma and a squamous cell carcinoma of the scalp. The patient's maternal family history fulfilled Amsterdam Criteria II due to the patient's early-onset colon cancer, her mother's history of sebaceous gland skin tumors, and a great-grandfather with colon cancer. The patient's father had a family history of breast cancer in multiple paternal aunts concerning for hereditary breast and ovarian cancer (HBOC) syndrome.

Multiple genes, including *MSH2,* were analyzed prior to the patient's demise without identifying a causal variant. Whole genome sequencing (WGS) was performed in the mother a year after the patient's demise and did not reveal a causative variant to explain the family history of cancer. It did, however, incidentally identify a pathogenic variant in the *FBN1* gene, leading to a diagnosis of Marfan syndrome in the mother. Despite germline sanger sequencing and WGS failing to identify a pathogenic variant in one of the MMR genes, the family was counseled to follow LS clinical management guidelines. As part of this management, the patient's mother underwent a prophylactic total abdominal hysterectomy and bilateral salpingo-oophorectomy.

Eight years after her passing, testing on the patient's banked DNA utilizing a commercial laboratory's custom cancer panel on a Next-Generation sequencing (NGS) platform observed a c.1442_1443insAlu likely pathogenic variant in the *MSH2* gene. The mobile element (ME) detection software Mobster (Thung DT et al. Genome Biol., 2014) and the commercial laboratory's in-house developed software was used to detect unaligned and soft-clipped reads from the BAM file, and the variant was confirmed by sanger sequencing. The patient's parents underwent confirmatory genetic testing via the same commercial laboratory's custom cancer NGS panel considering this new genetic finding. The patient's father had negative/normal genetic findings. The patient's mother was found to carry the *MSH2*/Alu

insertion. The mother's WGS BAM file data were again reviewed and reads covering this insertion were not identifiable. Mobster, which was used to detect the Alu insertion in the daughter's case, was implemented to run on the maternal WGS. Split reads were detected on WGS in the same variant location. Upon further review, standard WGS BWA alignment (Li H et al. Bioinformatics, 2009) did not map the reads which contained more than 50% Alu reads and trimmed the reads with less than 50% Alu reads. This splice site was not detected by standard variant calling as reads are assessed for small variants and structural variation, the *MSH2*/Alu insertion was undetected by conventional NGS methods.

There is a subset of patients with a phenotype strongly suggestive of LS and no identifiable germline pathogenic variant. This case demonstrates the importance of critically assessing the testing methodologies previously performed in this patient cohort and specifically to consider testing capable of identifying retroelement insertions. In conclusion, this case demonstrates the value of reanalyzing short-read sequencing data for structural variants and retroelement events for cases that have not been previously diagnosed.