Investigation of microsatellite instability in RNA compared to DNA, using microsatellite targeted, anchored multiplex PCR and NGS

2943

Ryan Rogge, Taylor R. Patterson, Alicia Navetta, Dhruti Legare | Archer, Boulder, CO. USA.

Introduction

MSI measurements are typically made on DNA, but it would be useful to make similar measurements from RNA or total nucleic acid inputs (TNA), so that MSI determinations could be made in parallel to interrogation of the RNA. Doing so requires characterizations of microsatellite lengths from those which are transcribed and untranscribed, because both will be detected following cDNA creation. Microsatellite instability (MSI) occurs when small unit repeats of DNA undergo incorrect replication resulting in the growth or contraction of the number of repeats in the newly synthesized DNA. This accumulation of errors at microsatellite DNA is usually mitigated by the mismatch repair (MMR) pathway which can correct these errors. MSI is of interest because it serves as a phenotypic readout of the status of the MMR machinery, and MSI status can correlate to tumor progression and response to agents that regulate these cellular mechanisms. We investigated the status of microsatellites using Anchored Multiplex PCR (AMP™) targeted panels (RUO) and compared results from TNA inputs using an AMP MSI module and both VARIANT*Plex*™ chemistry to interrogate DNA, and FUSION*Plex*™ chemistry to interrogate both RNA + DNA. Our results indicate differences in the diversity of lengths of microsatellites present in RNA, and the resulting stability calls which would be made at those sites. Microsatellite length diversity was highly dependent on the genomic context of the microsatellite. Microsatellites located within transcripts showed increased length diversity when examining the RNA + DNA results compared to DNA alone. Microsatellites located in intergenic regions, which are unlikely to be transcribed, showed no increase in diversity of lengths when comparing the RNA + DNA results to DNA alone. While these results suggest that MSI determinations could be made from RNA or RNA + DNA NGS data, the unique behavior of microsatellites in RNA is likely to require additional investigations and unique data analysis techniques to correctly categorize microsatellites as stable or unstable.

Methods

The experiment described here evaluates whether MSI calls can be made from DNA when TNA (RNA + DNA) is used as input in FUSION*Plex* libraries. TNA extracted from 40 FFPE samples will be used to make FUSIONPlex and VARIANT*Plex* libraries, and MSI calls will be compared between libraries. The libraries prepared with FUSION*Plex* chemistry utilize reverse transcription to produce cDNA and so the sequencing results from these libraries will contain reads from both the RNA and DNA present in the original input nucleic acids. VARIANT*Plex* libraries utilize chemistry intended for DNA analysis only, and so the resulting libraries will report only on the DNA from the original input nucleic acids.

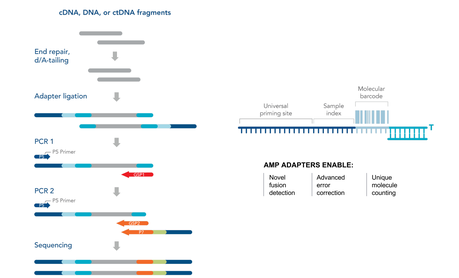

Figure 1. Overview of AMP chemistry for NGS library preparation.
Pictured here is a graphical abstract of how anchored multiplex PCR enables single sided, multiplexed, targeted NGS libraries. Briefly, adapters with universal priming sites provide one PCR primer for all fragments, while opposing gene specific primers provide targeted amplification for only sequences of interest. GSPs are used in a nested fashion providing for excellent on-target percentage. The adapters added also contain molecular barcodes, which enable deduplication of molecules for unique counts, and error correction of sequenced bases. Additionally, adapters include sample index sequences to enable demultiplexing of multiplexed NGS workflows. Not pictured here is an upstream cDNA creation step for FUSION*Plex* library preparations.

Methods (continued)

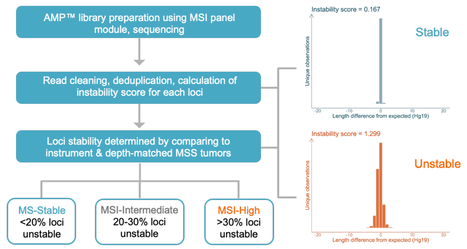

Figure 2. Outline of MSI pipeline for AMP libraries as part of Archer Analysis.
MSI analysis starts with deduplicated reads from Archer Analysis and builds a distribution of microsatellite lengths for each microsatellite locus. That distribution of lengths is used to calculate an instability score per locus, and then that score is compared against instrument matched normal baselines to determine if each microsatellite locus is stable or unstable. Finally, results are reported as percentage of microsatellite loci unstable, with sample classifications of MS-stable, MSI-Intermediate, or MSI-High. The thresholds for the sample classifications as well as many pipeline parameters can be set by the user but should work well with default settings.

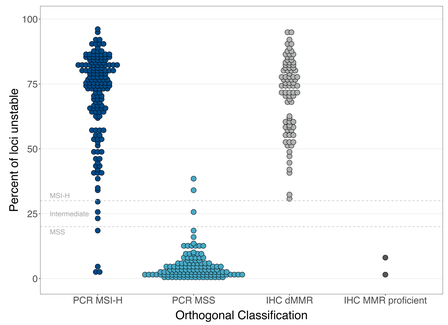

Figure 3. Performance of MSI with DNA inputs and VARIANT*Plex* chemistry.
362 de-identified FFPE samples previously characterized by PCR and fragment analysis or by IHC analysis compared to results from Archer Analysis on the same samples prepared with VARIANT*Plex* chemistry. These results have excellent concordance to the compared assays with 98.3% PPA (95% CI=95.6–99.5%) and 98.4% NPA (95% CI=94.4–99.8%). Libraries were prepped with 15-250 ng of DNA.

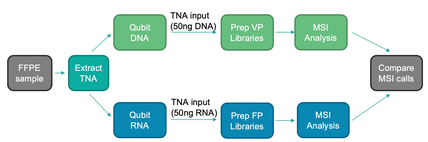

Figure 4. Experimental design for testing TNA input with FUSION*Plex* and VARIANT*Plex* chemistry.
40 de-identified FFPE samples were extracted for TNA, and then DNA and RNA present were quantified separately. NGS libraries were prepared with either VARIANT*Plex* chemistry and 50ng DNA to generate reads from the DNA, or with FUSION*Plex* chemistry and 50 ng RNA to generate reads from both RNA and DNA molecules. All libraries were sequenced using a NextSeq550, and the resulting data were analyzed using Archer Analysis.

Results

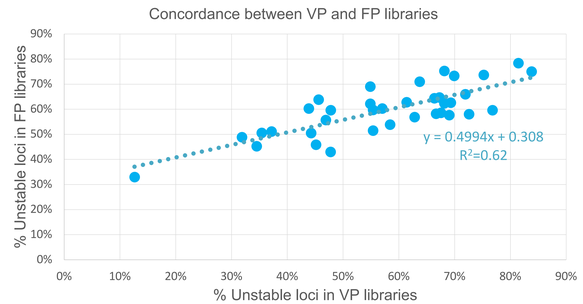

Figure 5. Correlation between MSI % unstable loci for each sample processed with FUSION*Plex* and VARIANT*Plex* chemistry.
The concordance between FUSION*Plex* (FP) and VARIANT*Plex* (VP) MSI results on the 40 TNA inputs. FP results include reads from both RNA and DNA, while the VP results include only DNA molecules. MSI % unstable are highly correlated categorically as all MSI-High samples would still be categorized as MSI-High. However, overall correlation indicates an increase in % unstable loci seen when using FP chemistry and interrogating both RNA and DNA molecules. This effect warranted further investigation on the effect of transcription across microsatellites.

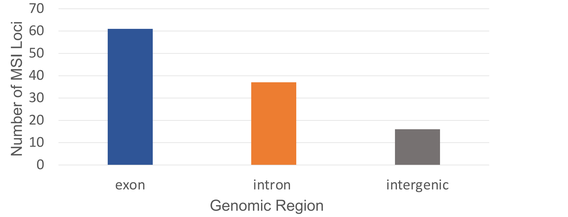

Figure 6. Distribution of microsatellite loci interrogated by location in or out of transcribed regions.
Microsatellite loci included in the MSI module are likely to be affected by whether the microsatellite is transcribed into RNA or found only in the DNA molecules present in the TNA inputs. This figure displays the location of the microsatellites as either exonic, intronic, or intergenic. The majority of the loci are transcribed and located within exons. The increased polymerase activity at these regions is likely to generate additional microsatellite lengths, and thus increased instability.

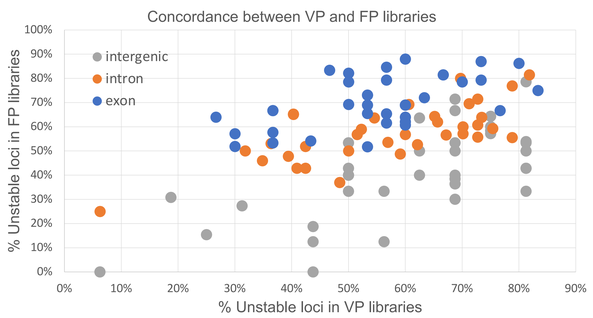

Figure 7. Percentage of microsatellites unstable when only calculated from a subset of loci based on genomic region.
The same 40 samples analyzed by the MSI pipeline in Figure 4, repeated using only the subset of loci from each genomic region. Each point represents a sample, and each color represents those samples data when only considering intergenic, intronic, or exonic loci. When separating loci used by the MSI pipeline based on genomic regions, those in transcribed regions produce a higher % unstable (blue and orange) data points.

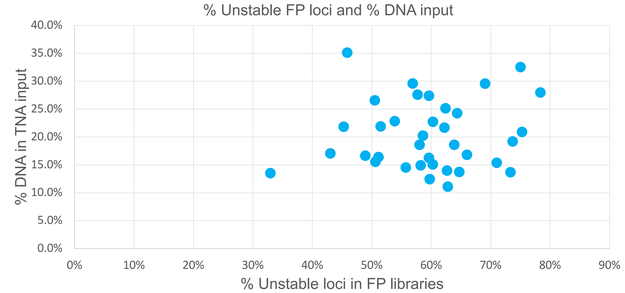

Figure 8. Lack of correlation between % DNA of TNA inputs, and MSI % unstable scores seen with FUSION*Plex* libraries.
Little to no concordance between the amount of DNA compared to RNA in the TNA inputs. Because the MSI pipeline was designed for DNA inputs we hypothesized that perhaps the amount of DNA responsible for our final readout was influencing the % of sites unstable. However, little to no correlation between the amount of DNA vs RNA and the MSI results suggests that this is not a contributing factor to differences noted between the FUSION*Plex* and VARIANT*Plex* chemistries.

Conclusions

MSI % of sites unstable calls appear feasible from total nucleic acid (TNA) inputs, based on strong correlation between results processed with FUSION*Plex* and VARIANT*Plex* chemistries and analyzed with Archer Analysis.

Transcribed microsatellites appear to have a higher diversity of lengths and a resulting higher instability compared to those from intergenic regions.

Development work would be necessary to address inherent differences in microsatellite stability between transcribed and un-transcribed microsatellites, in either TNA or RNA inputs, but would allow for simultaneous fusion, variant, and MSI calls from FUSION*Plex* products.

Possible solutions would be to adapt the microsatellite loci analyzed to those which are most correlated in the RNA results, and/or to re-calibrate the MSI pipeline for RNA/TNA results to better match the DNA results.

Contact

Ryan Rogge
Associate Director R&D, Archer (Integrated DNA Technologies)
rrogge@idtdna.com

Dhruti Legare
Staff Scientist, Archer (Integrated DNA Technologies)
dlegare@idtdna.com

Alicia Navetta
Applications Scientist, Archer (Integrated DNA Technologies)
anavetta@idtdna.com

Taylor Patterson
Research Scientist III, Archer (Integrated DNA Technologies)
tpatterson@idtdna.com