HRD classification by ASCN Diversity and Indel Repair Signatures with Anchored Multiplex PCR and Next Generation Sequencing

ST105

Taylor R. Patterson, Devin Tauber, Allison Hadjis, Ryan Rogge, Mark Rogers, Laura Johnson, Brent Lutz | Archer, Boulder, CO. USA.

Introduction

Homologous Recombination Deficiency (HRD) results in improperly repaired double stranded breaks in the genome due to repair by mechanisms that have lower fidelity than homologous recombination repair. In next-generation sequencing (NGS) data this may present as an increase in the number of amplifications and deletions, or by changes in the types of indels present. However, measuring these events by NGS typically requires very deep and broad coverage, if not whole genome sequencing (WGS).

The conventional method of HRD measurement sums the number of losses of heterozygosity (LOHs), telomeric allelic imbalances (TAIs), and large-scale state transitions (LSTs) into a genome instability score (GIS). Such methods ignore other well known genomic scarring signals, such as certain types of indels, while requiring tens of thousands of probes and many millions of additional sequencing reads.

We have developed an Anchored Multiplex PCR (AMP™) primer module and tumor-only bioinformatic method for HRD classification that combines the signal present in genome-wide allele specific copy number (ASCN) profiles and indel mutational signatures from targeted coverage of approximately 5000 single nucleotide polymorphism (SNP) targets and 80 indel targets.

HRD Method Description

The new IDT VARIANT*Plex*™ HRD Module contains primers that target likely-heterozygous SNPs spaced approximately 500kb apart across the genome (“SNP array”) and primers that target inverted repeats, which are expected to be more susceptible to double stranded breaks and enriched for genomic scars of improper repair such as indels. These two groups of primers provide the two metrics used to generate HRD scores: allele-specific copy number (ASCN) entropy and percent of indels repaired by non-homologous end joining (NHEJ) (Figure 1).

SNP array primers contribute breadth of coverage across the genome and enable ASCN measurements, which can be useful outside of HRD to detect LOH and other CNV events. The ASCN process also makes an estimate of tumor purity and tumor ploidy allowing total and minor absolute copy number estimates from CNV fold changes and SNP allele frequencies. The ASCN entropy metric summarizes the number and abundance of different total and minor copy number states observed. ASCN entropy is higher in HRD samples than in HRR samples, because improper double stranded break repair results in more total and minor copy number states of larger spans.

Inverted repeat primers add distinct information on the mechanisms of repair which have been active. Inverted repeats, which are sequences followed by their own reverse complements, can form secondary structures that are resolved by double stranded breaks. The HRR pathway repairs breaks perfectly, in which case no indel will be observed. When indels are observed, they can be attributed to different repair pathways by the length of the indel and the length of flanking microhomology, which we summarize as the percent of indels repaired by NHEJ. Percent NHEJ tends to be lower in HRD samples than in HRR samples because a greater proportion of breaks that would otherwise be perfectly repaired by homologous recombination are instead repaired other pathways, alternative end joining or single strand annealing.

HRD scores are calculated by Archer Analysis from the percent of indels repaired by NHEJ and the ASCN entropy in a multiple linear regression model. Scores are on the same scale as LOH+TAI+LST methods. HRD classification can then be made on the scores by user-define thresholds. By default, Archer Analysis uses two thresholds, <33 for HRD Negative and ≥42 for HRD Positive, allowing for an optional intermediate range between the two thresholds.

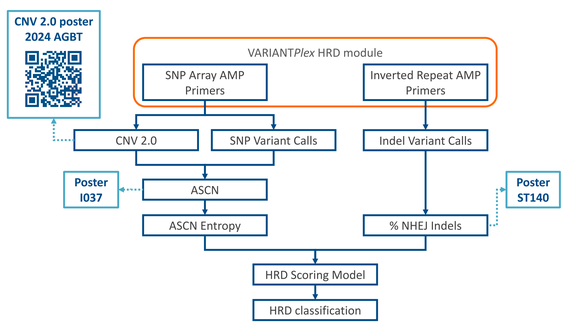


Figure 1. Archer VARIANT*Plex* HRD method diagram.

Information from the two components of the HRD module are used to generate distinct evidence of HRD from allele specific copy number states and double stranded break repair mechanisms then combined in a final scoring.

More information on our tumor-only CNV method can be found in a poster linked by the QR code.

More information on our tumor-only ASCN method can be found at Poster I037, and more information on the % NHEJ metric can be found at poster ST140.

Results

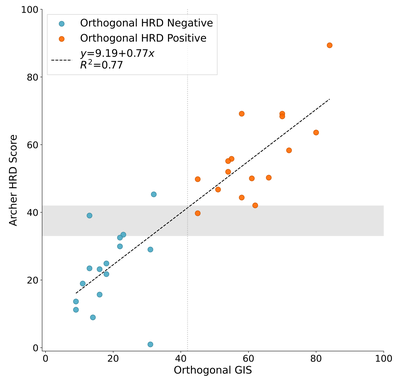


Figure 2. Archer HRD concordance with GIS.

VARIANT*Plex* libraries were prepared from 50 ng each of thirty-two samples, comprising FFPE tumor tissue samples, one blood sample, and three FFPE cell line reference inputs with LOH+LST+TAI scores from comparator assays using panels containing the HRD module, then sequenced and analyzed with Archer Analysis.

Each point represents one library, colored by the orthogonal HRD classification based on GIS scores with a threshold of 42 (indicated by vertical gray dotted line). The horizontal gray bar represents the Archer Analysis default HRD intermediate classification range, from 33-42. A score of 42 and above is classified as HRD Positive by Archer Analysis, while scores below 33 are classified as HRD Negative. The black dashed line is a line of best fit.

		Archer Classification	
		HRD Positive	HRD Negative
Orthogonal classification	HRD Positive	15	1
	HRD Negative	1	15

Table 1. Archer HRD concordance with GIS confusion matrix.

A threshold of 42 was applied to both the Archer HRD score and the orthogonal GIS scores for the libraries represented in fig 2.

When scores by either method ≥42 are classified as HRD positive, the positive percent agreement (PPA) is 93.75% (15/16 libraries), the negative percent agreement (NPA) is also 93.75% (15/16 libraries), and the overall accuracy is also 93.75% (30/32 libraries).

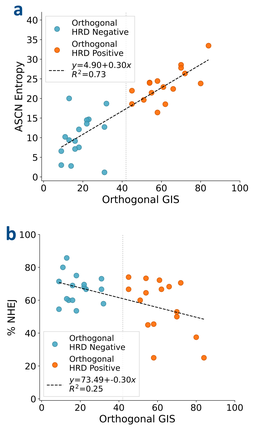


Figure 3. Concordance of metrics used to calculate HRD score with GIS.

ASCN entropy (a) and % NHEJ (b) for the same 32 libraries shown in fig. 2.

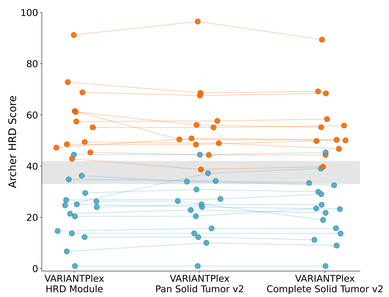
The dotted vertical gray line is the conventional GIS score threshold.

These two metrics are combined in the final HRD score by multiple linear regression.

Figure 4. Archer HRD scores across different HRD-enabled VARIANT*Plex* panels

Twenty-eight of the samples depicted in fig. 2 were prepared using VARIANT*Plex* Complete Solid Tumor v2. These libraries were re-analyzed using the panel content of VARIANT*Plex* Pan Solid Tumor v2 and the HRD Module alone, both of which are subsets of the Complete Solid Tumor v2 panel.

Each point represents one library, colored by the orthogonal HRD classification based on GIS scores with a threshold of 42. The horizontal gray bar represents the Archer Analysis default HRD intermediate classification range, from 33-42. A score of 42 and above is classified as HRD Positive by Archer Analysis, while scores below 33 are classified as HRD Negative.



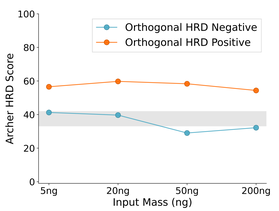


Figure 5. Archer HRD scores across input mass.

Libraries were prepared from 5, 20, 50 and 200 ng each of Seraseq FFPE HRD High-Positive and Seraseq FFPE HRD Negative reference materials using VARIANT*Plex* Complete Solid Tumor v2, sequenced, and analyzed with Archer Analysis. The orthogonal GIS score for the HRD Negative material is 31, the orthogonal GIS score for the HRD High-Positive material is 72.

The horizontal gray bar represents the Archer Analysis default HRD intermediate classification range, from 33-42.

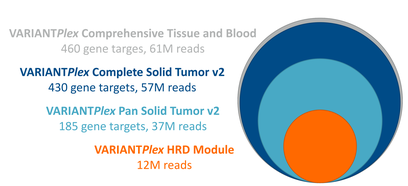


Figure 6. HRD-enabled catalogue VARIANT*Plex* Panels.

Three genomic profiling panels, VARIANT*Plex* Comprehensive Tissue and Blood, VARIANT*Plex* Complete Solid Tumor v2, and VARIANT*Plex* Pan Solid Tumor v2 contain the HRD module to enable ASCN and HRD in addition to SNV, indel, ITD, CNV, MSI, and TMB detection and measurement.

The HRD module may be combined with any other VARIANT*Plex* content to create custom, ASCN and HRD enabled panels.

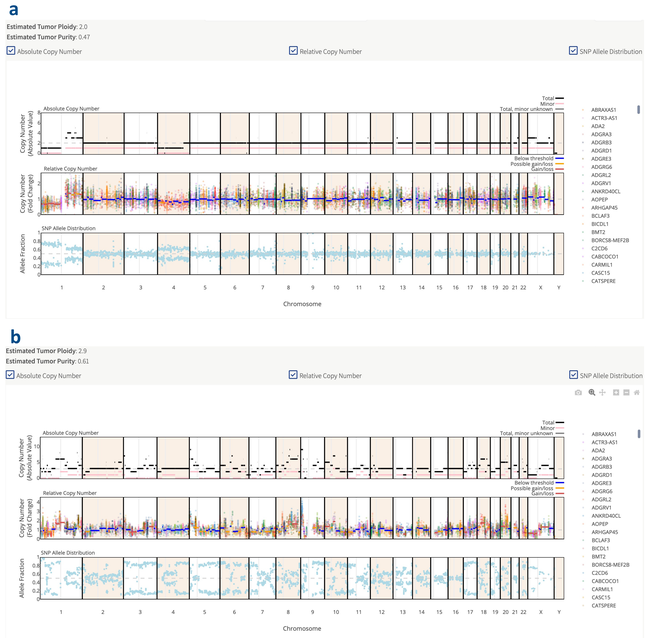


Figure 7. Archer Analysis v7.4 ASCN results visualization.

Archer Analysis produces interactive visualizations of ASCN results. The samples depicted in both images were prepared with VARIANT*Plex* Complete Solid Tumor v2 using 50 ng of input. The top image (a) is an HRD negative ovarian cancer FFPE, the bottom (b) image is the Seraseq FFPE HRD High-Positive reference material.

In both images, estimated tumor ploidy and tumor purity are presented in the upper left corner. In the middle of the screenshots are three panes, each with a different y-axis but with shared x-axes of genomic position on each chromosome.

The top pane shows the absolute total and minor copy numbers estimated by combining the information of the SNP allele frequencies, CNV fold changes, estimated tumor purity, and estimated tumor ploidy. The middle pane shows the fold change for each primer (point) and segment (lines) drawn among primers of similar fold change by the tumor-only CNV 2.0 segmentation process. The distinction between possible gains/losses and gains/losses in the figure legend and segment coloring is the significance level based on user defined alpha thresholds. In the bottom pane each point is a SNV, the y-axis is the allele frequency (AF). A dashed line is drawn at 0.5, where one would expect heterozygous SNVs to be.

The HRD negative sample (a) shows relatively few CNVs, as most of the genome is has two total copies, one minor copy. However, the HRD positive sample is very unstable, indicated by the many different copy number states, CNVs, and allele frequency splits away from 0.5. The HRD score for the HRD negative ovarian cancer FFPE (a) is about 1. The HRD score for the Seraseq FFPE HRD High-Positive reference material is about 58.

Conclusions

AMP chemistry-based targeted panels and analysis method enable highly concordant HRD classifications (fig. 2, table 1). They requires fewer reads than WGS or other NGS methods and does not require paired tumor-normal data or a panel of normals. The HRD module also enables ASCN detection (fig. 7), which can be used to detect losses of heterozygosity and other CNV events.

The HRD module can be combined with any other VARIANT*Plex* panel content for simultaneous measurement of HRD, other genomic signatures, and variant calling. Input recommendations are ≥10ng of DNA, depending on input quality, and ≥20% tumor cellularity.

Taylor R. Patterson
Assay Development R&D, Archer (Integrated DNA Technologies)
tpatterson@idtdna.com