The present study involved the analysis of raw SNP or aCGH data from the analytical approach across different clinical datasets, to identify common features where the inclusion of molecular biomarkers into risk assessment could impact the potential to identify those patients most likely to have clinical correlation of CNAs and genomic complexity.

Methods

DNA Isolation and Array

DNA was isolated from frozen tumor specimens using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer’s instructions.

Arrays

SNP arrays were performed on Affymetrix 6.0 (n=172), Illumina HumanOmni2.5 BeadChip (n=51HR), and Illumina 450K BeadChip (n=180). Array data were preprocessed and normalized using the BioConductor package RMA (Robust Multiarray Average) and quantile normalization. Genomic coordinates were adjusted to the UCSC Genome Browser assembly hg19.

Results

In-silico Identification of Overlapping CNAs in DLBCL

GISTIC identified regions of significant CNAs and peak-CMAs in the DLBCL datasets IS-180 (n=172) and IS-51HR. A DNA region of overlap was defined as the largest targeted genomic gandies including all overlapping regions across least 2 of the 3 datasets. The overlapped region of overlap within a DNA region was defined as its genomic region.

A peak-CMA was defined as the largest targeted genomic gandies across least 2 of the 3 datasets. The overlapped region of overlap within a peak-CMA was defined as its peak-CMA.

Genomic complexity was assessed by 3 different methods and tested for outcome association in three RCHOP treated datasets (IH, IS-124, IS-51HR).

Cross-Platform Assessment of Genomic Imbalance in Diffuse Large B-Cell Lymphoma Identifies Candidate Novel Loci and Genes with Prognostic Value and Roles in Lymphomagenesis.

Lizalyn Dias1, Venkata Thodima1, Geetu Mendiratta1, Asha Guttapalli1, Camille Gonzalez2, Jocelyn Marquaglia2, Andrew D. Zelenetz3, Julie Tenuya-Feldstein4, R.S.K. Chaganti5,6, Jane Houldsworth7

1 Cancer Genetics, Inc., Rutherford, NJ, 2 Department of Pathology, 3 Department of Medicine, and 4 Cell Biology Program, Memorial Sloan-Kettering Cancer Center

Cancer Genetics

For the 162 samples of the IS-124 dataset with available gene expression data genes that pairwise correlated with copy number values of the region CMAs were identified using the t-affy Bioconductor (R) package.

Gene ontology and pathway analysis was performed with Ingenuity.

The present study involved the analysis of raw SNP or aCGH data from the following publicly available studies (IS) and in-house (Rh) datasets:

- IS-172

(iii) 172 fresh frozen DLBCL tumor biopsy specimens diagnosed CHOP treated patients

(iii) Human Genome HapMap Trios array (R01K05)

(iii) Gene expression data was available for 162 of the 172 specimens profiled in the IS-172 dataset

- IS-51HR

(iii) 51 high risk DLBCL tumor tissues from newly diagnosed CHOP and RCHOP treated patients

(ii) SNP array Eu (Affymetrix)

(iii) Datasets which were analyzed in this study: IS-180, and IS-51HR were treated datasets (IH, IS-124, IS-51HR) and IS-51HR (n=46) diagnostic, respectively as one of the 2 cohorts.

(iii) RH-63

(iii) All 63 affy Affymetrix probes from RCHOP treated patients with the disease (DLBCL)

(iii) Custom designed high-resolution Affymetrix array

- Methods

DNA Isolation and array-CGH

Genomic DNA was extracted from 1.5-MLP core punches (1 mm diameter) from the IS-172 dataset following the protocols by Vihm et al. [7]. Labeled DNA was hybridized to a custom designed high-resolution Affymetrix array consisting of genomic regions designed to assist in the validation of various mature B-cell neoplasms including DLBCL (2).

GISTIC Analysis and Aberration Detection

Document/Chromosomes were identified by applying GISTIC2 algorithm to SNP or aCGH segment data from the datasets in the IS-124 and IS-51HR for the 2.5-2.5 log2 and 1.5-1.5 log2. TGF-β (value of 0.26). Raw aCGH SNP data were loaded into the Nexus Copy Number Analysis Software 6.1 where aberration detection was performed using the Rank segmentation algorithm.

Clinical Correlation Analyses

The clinical outcomes and genomic complexity were assessed in RCHOP treated datasets IS-180, IS-124 and IS-51HR using t-test (Unpaired 2sided) or Fisher’s exact test (Fisher’s exact test) as applicable. Where appropriate the Bonferroni multiple testing correction was performed.