Epstein-Barr virus-associated primary nodal T/NK-cell lymphoma shows a distinct molecular signature and copy number changes

Siok-Bian Ng,[1,2,3]* Tae-Hoon Chung,[3]* Seiichi Kato,[4] Shigeo Nakamura,[4] Emiko Takahashi,[5] Young-Hyeh Ko,[6] Joseph D. Khoury,[7] C. Cameron Yin,[7] Richie Soong,[1,3] Anand D. Jeyasekharan,[3] Michal Marek Hoppe,[3] Viknesvaran Selvarajan,[1] Soo-Yong Tan,[1,2] Soon-Thye Lim,[8] Choon-Kiat Ong,[9] Maarja-Liisa Nairismägi,[9] Priyanka Maheshwari,[2] Shoa-Nian Choo,[1] Shuangyi Fan,[1] Chi-Kuen Lee,[1] Shih-Sung Chuang[10] and Wee-Joo Chng[3,11]

[1]Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore; [2]Department of Pathology, National University Hospital, National University Health System, Singapore; [3]Cancer Science Institute of Singapore, National University of Singapore; [4]Department of Pathology and Laboratory Medicine, Nagoya University Hospital, Nagoya, Japan; [5]Department of Pathology, Aichi Medical University Hospital, Nagakute, Japan, [6]Department of Pathology, Samsung Medical Center, Sungkyunkwan University, Seoul, Korea; [7]Department of Hematopathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [8]Lymphoma Genomic Translational Research Laboratory, National Cancer Centre Singapore, Division of Medical Oncology, National Cancer Center Singapore; [9]Lymphoma Genomic Translational Research Laboratory, Division of Medical Oncology, National Cancer Centre Singapore; [10]Department of Pathology, Chi-Mei Medical Center, Tainan, Taiwan and [11]Department of Haematology-Oncology, National University Cancer Institute of Singapore, National University Health System

**S-BN and T-HC contributed equally to this work.*

doi:10.3324/haematol.2017.180430

Received: September 9, 2017.

Accepted: October 27 2017.

Pre-published: November 2, 2017.

Correspondence: patnsb@nus.edu.sg or mdccwj@nus.edu.sg.

Supplemental Methods

Study group

The study group included 66 adult patients with no known immunodeficiency and diagnosed with EBV-positive T/NK-cell lymphoma (TNKL) involving extranodal and nodal sites between years 1992 to 2015. Some of the cases have been previously published.[1-3]. All the cases were positive for cCD3, EBV-encoded small RNAs (EBER) in the majority of tumor cells (>50%), and at least one cytotoxic marker (TIA1 and/or granzyme B). The majority of these cases, including the nodal cases, would be classified as ENKTL based on the 2008 WHO lymphoma classification which recognizes that rare cases of ENKTL may present with nodal disease in the absence of extranodal involvement.[4] Systemic and cutaneous EBV-positive T/NK lymphoproliferative diseases occurring in children such as systemic EBV+ T-cell lymphoma, aggressive NK-cell leukemia, and chronic active EBV infection of T/NK type and hydroa vacciniforme–like lymphoproliferative disorder were excluded.

Clinical data including age, sex, ethnicity, site of disease presentation, nasal involvement, stage and overall survival were obtained (Table S1) and the cases were then categorized based on the following features: i) disease presentation (extranodal [EN-group] vs nodal [N-group]); ii) nasal involvement (absence vs presence); and iii) cell of origin (COO, T-cell vs NK-cell). T-cell lineage was assigned based on a) positive expression of TCRB and/or TCRG protein by immunohistochemistry (IHC) and/or b) detection of monoclonal TCRG rearrangement by PCR using the BIOMED2 protocol. NK-cell lineage was assigned based on the lack of TCRB and TCRG expression by IHC and germline TCRG by PCR.[1,5,6] (Figure S1) For cases that lacked clonality testing and were negative for both TCRG and TCRB protein expression (denoted with [*] in Table S1 and Figure 1A): i) T-cell lineage was assigned based on lack of CD56 expression and/or loss of TCR loci and ii) NK- lineage was assigned based on expression of CD56 and/or intact TCR loci. Cases lacking complete information or remained indeterminate in lineage were assigned as indeterminate (IND) (Figure S1).

Control tissues

The normal B-cells, normal T-cells and normal NK-cells were isolated from peripheral blood. Benign lymph node tissues used for copy number analysis and GEP were benign or reactive FFPE lymph node tissues obtained from the archives of the department of pathology.

GEP

GEP was performed using Affymetrix HuGene 1.0 ST arrays in this study. Total RNAs from NKTL FFPE tissues and FFPE normal tissue controls were extracted using RecoverAll Total Nucleic Acid Isolation (ThermoFisher Scientific Inc., USA) according to the manufacturer’s instructions. All the sections were deparaffinized with Histo-Clear II (National Diagnostics, USA), subjected to proteinase K digestion, and extracted as per the manufacturer’s protocol. The concentration and purity of the total RNA extracted were measured using the NanoDrop 2000 spectrophotometer (ThermoFisher Scientific Inc., USA). RNA quality was assessed with the Agilent 2100 Bioanalyzer (Agilent Technologies, USA) and the RNA 6000 LabChip kit (Agilent Technologies, USA). Amplification and labeling of RNA samples were performed using the Ovation FFPE WTA System (NuGEN, USA). A total of 75 ng of input RNA was used to generate the amplified Ribo-SPIA product. Cleanup was performed with RNAClean magnetic beads (Agencourt Bioscience Corporation, USA) followed by Qiaquick PCR purification kit (Qiagen, Germany). Nanodrop 2000 spectrophotometer (ThermoFisher Scientific Inc., USA) was used to determine the amplified cDNA yield for the labeling and fragmentation steps. Amplified cDNA samples were fragmented and labeled using an Encore Biotine Module (NuGEN, USA). Hybridization cocktails were prepared according to NuGEN Encore Biotin Module user guide. Labeled products were hybridized onto Affymetrix Human Gene 1.0 ST arrays according to NuGEN specifications. Subsequently, the microarrays were washed and stained using user prepared wash and staining reagents. Scanning was performed on a GeneChip Scanner 3000.

GEP data processing and analysis

Data analysis was performed using the R/Bioconductor system.[7] Considering the geographical and temporal complexity of sample collection and handling, GEP data were preprocessed with frozen robust multiarray analysis (FRMA) algorithm, an extension of renowned RMA algorithm that utilizes probe-specific effects and variances. Since the parameters are frozen using estimations from large publicly available microarray databases in fRMA, individualized chip processing and simple data addition extending the flexibility of data processing of Affymetrix microarrays.[8,9] Signal intensities for probesets were finally transformed to represent expression values for genes by averaging intensities larger than half of maximum one for probesets of specific genes.

Unsupervised hierarchical clustering was performed using the top 500 most variable genes selected by median absolute difference (MAD) of log2-transformed median-normalized expression values where each gene's expression values were normalized against its own overall median and log2-transformation was taken afterwards. This was to account for gene-by-gene difference in natural expression levels and to avoid undesirable impacts due to outliers. Euclidean distance was used as the distance metric between two expression vectors and average linkage was used to measure the distance between two clusters.

Differentially expressed genes (DEG) between clusters and N-group vs EN-group were identified by using the significance analysis of microarrays (SAM) algorithm[10] implemented in Bioconductor library 'siggenes'.[11] We chose genes with fold change ≥1.5 in either groups, $p<0.05$ and false discovery rate ≤ 0.15. Gene Set Enrichment Analysis (GSEA) between N-group cases and normal lymph node control tissues was performed by using Broad Institute's java client program with MSigDB version 5.1.[12] based on $p <0.05$ and FDR < 0.2. We also performed GSEA analysis to investigate if there are signalling pathwayss which are significantly different between nodal group-T cell cases (N-T, n=16), extranodal group of T cell (EN-T, n=13) and extranodal group of NK cell origin (EN-NK, n=25) as follow: comparison (1): N-T vs EN-NK and comparison (2): N-T vs EN-T.

T-cell signature analysis was performed using genes CD27, CD3G, CD3D, ICOS, MAL, TCF7, PKIA[13] according to the procedure outlined in a previous publication.[14] In this case, log2-transformed median-normalized expression values of seven T-cell signature genes were simply added for each sample.

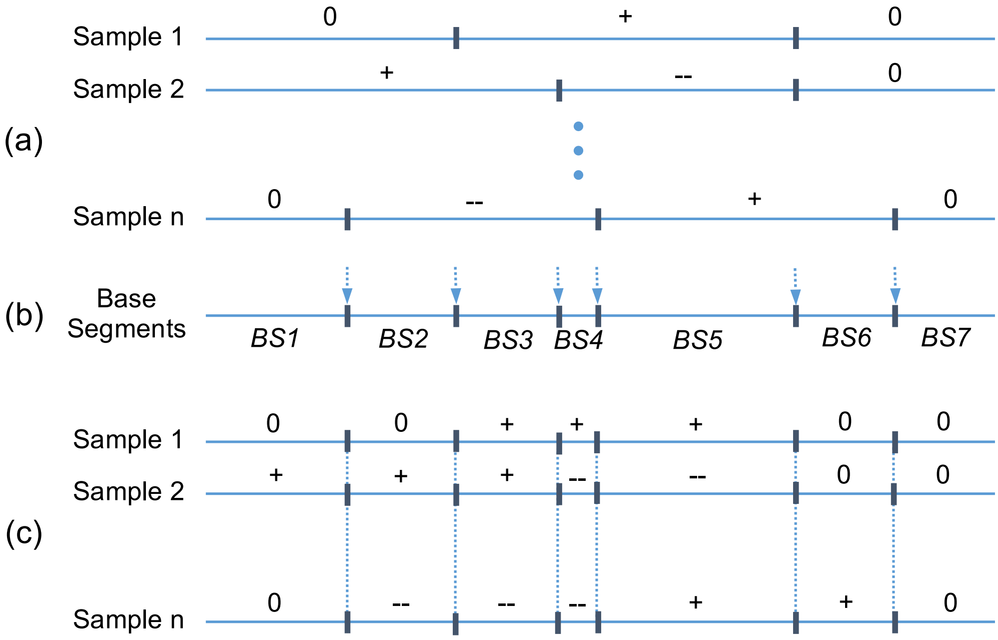
OncoScan® assay

The OncoScan® FFPE assay (Affymetrix) was performed according to the manufacturer's instructions as previously published.[15,16] The recommended genomic DNA input amount of 79.2ng was used whenever possible. Briefly, the genomic DNA targets were annealed to molecular inversion probes (MIP) followed by gap-filling with dNTPs (ATs or GCs) to circularize the MIP probes. Non-circularized probes were removed by exonuclease I treatment. Circularized probes were then cleaved and released from the DNA targets, and inverted to the correct orientation for first round of PCR amplification before assessing for successful amplification on 4% E-Gel® 48 (Invitrogen). Thereafter, 2uL of PCR product was used for second round of PCR amplification followed by restriction digestion with HaeIII. Complete digestion was confirmed on the 4% E-Gel® 48 before digested products were hybridized to the OncoScan® arrays for 16 to 18 hours at 49°C, 60 rpm. Finally, the arrays were washed and stained in GeneChip® Fluidics Station 450 prior to scanning using the GeneChip® 3000 7G scanner. The OncoScan® Console (v1.3) software was utilized to produce the QC metric and other analysis data.

CNA analysis

Segmentation results from OncoScan® Console and the probe-by-probe signal intensities were used for further analysis. Five levels of CNA (-2, -1, 0, +1, +2) were assumed in this study. Each segment was assigned to one of the possible CNA levels using the following rule: -2 (copy number (CN) ≤ 0.5), -1 ($0.5 < CN \leq 1.5$), 0 ($1.5 < CN < 2.5$), +1 ($2.5 \leq CN < 3.5$), +2 ($3.5 \leq CN$).

To derive CNAs specific to N-group and EN-group, ie. CNAs that are exclusively enriched or depleted in any of the groups, we split genomic regions into disjoint base segments (BSs) of respectively unique gain/loss profiles and performed Fisher's exact test between the groups and gain/loss profile of each BS (shown below). Since start and end positions actual CNA segments of different samples cannot be identical (part (a)), we marked the start positions of all CNA segments from all samples first and constructed BSs by delimiting genomes by those marked positions (part (b)). In this construction, since BSs were delimited by start positions of actual segments, no BS can cross the boundary of any two actual CNA segments and hence any actual CNA segment can be represented by one BS or multiple consecutive BSs. Finally, the copy number of each BS was assigned to the copy number of encompassing actual CNA segment (part (c)). When this was applied to all chromosomes, we produced the CNA matrix. Finally, the association between each BS CNA profile (row of CNA matrix) BSs with p ≤ 0.05 were retained for further examination.



CNA of four TCR loci (TCRA, TCRB, TCRG, TCRD) were also examined using CNA levels of segments that harbored respective TCR locus. The following locations are used (hg19): TCRA: chr14: 22265459 – 23016719, TCRB: chr7: 142008780 –

142495186, TCRG: chr7: 38279181 – 38403119, TCRD: chr14: 22918107 – 22934912.

Multiplex immunofluorescence (MIF) and multispectral Imaging

MIF for PD-L1/CD3 and CD2/CD3 was performed on 50 cases using 3 microns formalin fixed paraffin embedded (FFPE) sections using Opal™ 7-color Flourophore TSA plus Fluorescence Kit (NEL797001KT, PerkinElmer, Inc.). Image acquisition and analysis were done with the Vectra 2 multispectral automated imaging system (PerkinElmer) and inForm 2.0 image analysis software. Nuance software (PerkinElmer) was employed to build the spectral libraries for the chromogens (Opal 570, Opal 690 and DAPI). These chromogen signature profiles were later used to spectrally unmix and quantitate CD3, CD2 and PD-L1 staining intensity, with appropriate regions for analysis chosen by 2 pathologists (NSB, PK).

For each case, 4 images containing at least 10000 cells were analyzed. The absolute optical density (OD) of each chromogen indicating the intensity of each antibody was obtained for every image was normalized against respective positive cutoff and the resulting median OD ratio was compared between N-group and EN-group. For PD-L1/CD3, slides were dewaxed and heat induced epitope retrieval (HIER) was performed at pH 9 for 20 min. This was followed by blocking of endogenous peroxidase activity and blocking with 10% albumin. Slides were subsequently incubated with PD-L1 (E1L3N, Cell Signalling 13684S, 1:100) for 60 min, incubated with Anti-Rabbit HRP labeled secondary antibody, and amplified with TSA plus Opal 570. HIER was performed again for 5 min followed by blocking with 10% albumin, incubation with CD3 (DAKO polyclonal, 1:100) for 30 min, incubation with

Anti-Rabbit HRP labeled secondary antibody, and amplification with TSA plus Opal 690. Slides were then counterstained with DAPI (FP1490). Appropriate controls were used.

For CD2/CD3; Slides were dewaxed and HIER was performed at pH 9 for 20 min. This was followed by blocking of endogenous peroxidase activity and blocking with 10% albumin. Slides were subsequently incubated with CD2 (EPR 6451, Abcam ab131276, 1:50) for 60 min, incubation with Anti-Rabbit HRP labeled secondary antibody, and amplification with TSA plus Opal 570. Slides were subjected to HIER for another 5 min, blocking with 10% albumin, incubation with CD3 (DAKO polyclonal, 1:100) for 30 min, incubation with Anti-Rabbit HRP labeled secondary antibody, and amplification with TSA plus Opal 690. Slides were then counterstained with DAPI (FP1490). Appropriate controls were also used.

FISH for TCRA

All samples used were fixed in neutral buffered formalin and paraffin-embedded. 4-µm sections were placed on electrostatically charged slides (Platinum Pro, Matsunami Glass Ind. Ltd, Japan). The sections were then subjected to direct FISH using the SureFISH 14q11.2 TCRAD probes (Agilent Technologies, TX, USA). The 468 kb probe specific for the 5' (centromeric) side spanning Human Chromosome 14:21645060-22113080 was labeled in P5 Green and the 305 kb probe specific for the 3' (telomeric) side spanning Human Chromosome 14:23031630-23336700 was labeled in Red. FISH was performed according to standard protocol as described previously.[17] A total of 100 non-overlapping nuclei were scored per case using fluorescence microscopy. Cells that demonstrate a breakapart will have a distinct red signal distanced from a distinct green signal, indicating a translocation involving one TCRAD allele, in addition to one fused red/green (or yellow) signal of the normal chromosome 14 (signal pattern: 1R, 1G, 1Y). In the event of a biallelic translocation, no fused signals would be present, and two distinct red and two green signals would be seen (signal pattern: 2R, 2G). Cells that do not demonstrate a breakapart will have two fused signals (signal pattern: 2Y).

The interpretation of fused and breakapart signals was based on generally accepted guidelines used for all other commercially available break-apart FISH assays in clinical laboratories performing testing using this method. This approach requires the space between two signals to be greater than one signal diameter in order to be considered a split signal. The FISH images were obtained using an Olympus BX61 microscope and captured on the Applied Image Analysis System v.3.93 (Applied Imaging, Pittsburgh, PA)

References

1. Kato S, Asano N, Miyata-Takata T, et al. T-cell receptor (TCR) phenotype of nodal Epstein-Barr virus (EBV)-positive cytotoxic T-cell lymphoma (CTL): a clinicopathologic study of 39 cases. Am J Surg Pathol. 2015;39(4):462-471.
2. Jeon YK, Kim JH, Sung JY, Han JH, Ko YH, Hematopathology Study Group of the Korean Society of P. Epstein-Barr virus-positive nodal T/NK-cell lymphoma: an analysis of 15 cases with distinct clinicopathological features. Hum Pathol. 2015;46(7):981-990.

3. Takahashi E, Asano N, Li C, et al. Nodal T/NK-cell lymphoma of nasal type: a clinicopathological study of six cases. Histopathology. 2008;52(5):585-596.

4. Chan JKC, Quintanilla-Martinez L, Ferry JA, Peh SC. Extranodal NK/T-cell lymphoma, nasal type. In: Swerdlow SH, Campo E, Harris NL, et al., eds. WHO classification of tumours of haematopoietic and lymphoid tissues. Lyon: IRAC Press; 2008:285-288.

5. Ko YH, Chan JKC, Quintanilla-Martinez L. Virally associated T-cell and NK-cell Neoplasms. In: Jaffe ES, Arber DA, Campo E, Harris NL, Quintanilla-Martinez L, eds. Hematopathology. Philadelphia: Elsevier; 2017:565-598.

6. Jhuang JY, Chang ST, Weng SF, et al. Extranodal natural killer/T-cell lymphoma, nasal type in Taiwan: a relatively higher frequency of T-cell lineage and poor survival for extranasal tumors. Hum Pathol. 2015;46(2):313-321.

7. Team RC. A language and environment for statistical computing. R Foundation for statistical computing, 2015; Vienna, Austria; 2016.

8. McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). Biostatistics. 2010;11(2):242-253.

9. McCall MN, Jaffee HA, Irizarry RA. fRMA ST: frozen robust multiarray analysis for Affymetrix Exon and Gene ST arrays. Bioinformatics. 2012;28(23):3153-3154.

10. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001;98(9):5116-5121.

11. Schwender H. siggenes: Multiple testing using SAM and Efron's empirical Bayes approaches. R package version 1.46.0.; 2012.

12. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-15550.

13. Iqbal J, Weisenburger DD, Chowdhury A, et al. Natural killer cell lymphoma shares strikingly similar molecular features with a group of non-hepatosplenic gammadelta T-cell lymphoma and is highly sensitive to a novel aurora kinase A inhibitor in vitro. Leukemia. 2011;25(2):348-358.

14. Chng WJ, Chung TH, Kumar S, et al. Gene signature combinations improve prognostic stratification of multiple myeloma patients. Leukemia. 2016;30(5):1071-1078.

15. Lee CS, Bhaduri A, Mah A, et al. Recurrent point mutations in the kinetochore gene KNSTRN in cutaneous squamous cell carcinoma. Nat Genet. 2014;46(10):1060-1062.

16. Harms PW, Fullen DR, Patel RM, et al. Cutaneous basal cell carcinosarcomas: evidence of clonality and recurrent chromosomal losses. Hum Pathol. 2015;46(5):690-697.

17. Ventura RA, Martin-Subero JI, Jones M, et al. FISH analysis for the detection of lymphoma-associated chromosomal abnormalities in routine paraffin-embedded tissue. J Mol Diagn. 2006;8(2):141-151.

Supplemental Table 1. Clinicopathologic features of 66 cases of EBV-positive T/NK cell lymphoma.

Table 1A. Clinicopathologic features of Nodal-group

Sample ID	Age	Sex	Ethnicity	DP	Stage	OS	Status	Nasal	CD8	CD56	TCR Beta	TCR Gamma	PCR	TCR loss	COO
NKTL50	35	M	Bangladesh	N	4	0.1	0	No	Neg	Neg	Neg	Neg	na	Yes	T*
NKTL59	85	F	Chinese	N	3	1.3	na	na	Neg	Neg	Neg	Neg	na	Yes	T*
TNK01	59	M	Japanese	N	3	1.4	1	No	Neg	Neg	Neg	Neg	na	Yes	T*
NKTL69	81	M	Chinese	N	na	1	1	na	Pos	na	Neg	na	R		T
A1	70	M	Japanese	N	4	0.2	1	No	Pos	Neg	Pos	Neg	R	Yes	T
K1	55	M	Korean	N	4	0.1	1	No	na	Neg	Pos	na	na		T
K2	89	M	Korean	N	4	1.4	1	No	Pos	Neg	Pos	na	na		T
NKTL40	58	M	Chinese	N	4	3.5	1	No	Pos	Neg	Pos	na	na	Yes	T
T114	64	M	Japanese	N	4	0.3	1	No	Pos	Neg	Neg	Neg	R	Yes	T
T242	68	M	Japanese	N	4	1	1	No	Pos	Neg	Neg	Neg	R	Yes	T
T85	63	M	Japanese	N	4	6	1	No	Pos	Neg	Pos	Neg	na		T
TNK03	59	F	Japanese	N	1	39.8	0	No	Pos	Neg	Pos	Neg	na	Yes	T
TNK09	59	F	Japanese	N	1	23.8	0	No	Pos	Neg	Pos	Neg	na	Yes	T
TNK10	79	F	Japanese	N	3	12.3	0	No	Pos	Neg	Pos	Neg	na	Yes	T
TNK04	55	M	Japanese	N	4	9.5	1	No	Pos	Pos	Pos	na	R	Yes	T
TNK06	32	M	Japanese	N	3	3.8	1	No	Pos	Pos	Neg	na	R	Yes	T
TNK05	49	M	Japanese	N	4	1.3	1	No	Neg	Pos	Neg	na	NR		NK
TW23	52	M	Chinese	N	4	5.2	1	No	Neg	Pos	Neg	na	NR		NK
NKTL65	46	M	Chinese	N	na	na	na	No	Neg	Neg	Neg	Neg	na		IND

Table 1B. Clinicopathologic features of Extranodal-group

Sample ID	Age	Sex	Ethnicity	DP	Stage	OS	Status	Nasal	CD8	CD56	TCR Beta	TCR Gamma	PCR	TCR loss	COO
MD05	64	F	Caucasian	EN	4	4.8	1	Yes	Neg	Neg	Neg	Neg	na	Yes	T*
NKTL43	58	F	Chinese	EN	4	25.7	1	No	Neg	Pos	Neg	Neg	na	Yes	T*
MD15	72	M	Caucasian	EN	1	27.8	1	Yes	Pos	Pos	Neg	Neg	na	Yes	T*
MD10	26	F	Hispanic	EN	2	147.5	0	Yes	na	Neg	Pos	na	na	Yes	T
NKTL23A	51	M	Chinese	EN	na	6.9	1	No	Neg	Neg	Pos	Neg	R	Yes	T
NKTL60	54	M	Chinese	EN	4	80.1	0	No	Neg	Neg	Neg	na	R	Yes	T
T202	63	F	Japanese	EN	4	1.5	1	No	Pos	Neg	Neg	Neg	R	Yes	T
T25	44	M	Japanese	EN	4	15	0	No	Neg	Neg	Neg	Pos	NR	No	T
TW40	43	M	Chinese	EN	1	14.7	0	Yes	na	Neg	Neg	Neg	R		T
MD02	47	F	Hispanic	EN	4	12.1	1	Yes	Neg	Pos	Neg	Neg	R		T
TW08	32	M	Thai	EN	1	7	0	na	Neg	Pos	Pos	neg	R		T
TW22	63	M	Chinese	EN	4	4.5	1	No	Neg	Pos	neg	pos	na	No	T
TW39	40	F	Chinese	EN	1	15	0	Yes	Neg	Pos	Neg	Neg	R	No	T
MD11	33	M	Hispanic	EN	na	0.4	1	No	na	Pos	Neg	Neg	na	No	NK*
NKTL17	69	M	Chinese	EN	na	na	na	Yes	Neg	Pos	Neg	Neg	na		NK*
NKTL36	56	M	Chinese	EN	na	na	na	na	Neg	Pos	Neg	Neg	na	No	NK*
NKTL41	37	M	Chinese	EN	na	na	na	Yes	Neg	Pos	Neg	Neg	na	No	NK*
NKTL52	65	M	Malay	EN	1	0.4	na	Yes	Neg	Pos	Neg	Neg	na	No	NK*
TW13	25	M	Chinese	EN	na	0.3	1	na	Neg	Pos	Neg	Neg	na	No	NK*
TW19	32	F	Chinese	EN	na	8	0	na	Neg	Pos	Neg	Neg	na		NK*
TW26	41	M	Chinese	EN	1	157	0	Yes	Pos	Pos	Neg	Neg	na	No	NK*
TW27	73	M	Chinese	EN	1	25	1	Yes	Neg	Pos	Neg	Neg	na	No	NK*
TW29	34	M	Chinese	EN	1	138.8	0	Yes	Neg	Pos	Neg	Neg	na	No	NK*

TW34	77	M	Chinese	EN	1	7.3	1	Yes	Neg	Pos	Neg	Neg	na	No	NK*
T241	74	M	Japanese	EN	4	3.6	0	No	Neg	Neg	Neg	Neg	NR	No	NK
TW10	33	F	Chinese	EN	1	8	1	na	Neg	Neg	Neg	neg	NR		NK
MD09	38	F	Hispanic	EN	1	87.3	0	No	Neg	Pos	na	na	NR		NK
MD13	35	F	Asian	EN	1	20.2	1	No	na	Pos	Neg	Neg	NR		NK
T35	34	F	Japanese	EN	4	0.9	1	No	Neg	Pos	Neg	Neg	NR		NK
T72	21	M	Japanese	EN	1	29.1	1	No	Neg	Pos	Neg	Neg	NR		NK
TW07	57	F	Taiwanese	EN	na	na	na	na	Neg	Pos	Neg	Neg	NR	No	NK
TW09	37	M	Chinese	EN	na	2	1	na	Pos	Pos	Neg	Neg	NR		NK
TW11	78	M	Chinese	EN	1	24	1	na	Neg	Pos	Neg	Neg	NR		NK
TW24	49	M	Chinese	EN	2	33	0	Yes	Neg	Pos	Neg	Neg	NR		NK
TW33	54	F	Chinese	EN	1	107	0	Yes	Neg	Pos	Neg	Neg	NR	No	NK
TW35	53	M	Chinese	EN	1	80.7	0	Yes	Neg	Pos	Neg	Neg	NR		NK
TW37	71	F	Chinese	EN	4	12	1	Yes	Pos	Pos	Neg	Neg	NR	No	NK
TW38	57	M	Chinese	EN	4	40.5	0	No	Neg	Pos	Neg	Neg	NR	No	NK
MD04	36	M	Asian	EN	1	132.4	1	Yes	na	na	Neg	Neg	na	No	IND
NKTL22	46	M	Chinese	EN	na	na	na	na	Neg	Neg	Neg	Neg	na	No	IND
NKTL24	88	M	Chinese	EN	na	na	na	na	Neg	Neg	Neg	Neg	na	No	IND
NKTL27B	32	M	Malay	EN	na	na	na	na	Neg	Neg	Neg	Neg	na	No	IND
NKTL53	40	M	Chinese	EN	1	0.4	na	Yes	Neg	Neg	Neg	Neg	na	No	IND
MD01	39	M	Hispanic	EN	1	141.1	0	Yes	na	Pos	Neg	Neg	na		IND
MD07	35	M	Caucasian	EN	4	3	1	No	na	Pos	na	na	na		IND
SGH2010	24	M	chinese	EN	2	0.2	0	No	na	Pos	Neg	Neg	na		IND
TW14	58	M	Chinese	EN	na	119	0	na	Pos	Pos	Neg	Neg	na		IND

M;male, F;female, N;nodal, EN;extranodal, DP; disease presentation, Stage; disease stage, OS; overall survival (months), Status; Dead=1 or Alive=0,Nasal; nasal involvement, na; not available, Pos; positive, Neg;negative, R; rearranged, NR; not rearranged, IND; indeterminate, PCR; TCRG clonality, COO: cell of origin, CNA; copy number analysis,

* denotes cases with absence of TCRG/TCRB expressionand lacking TCR clonality data and. COO assigned based on combination of CD56 expression and/or TCR loci information

Table S2 List of 500 top variable genes used to perform unsupervised clustering of 66 cases

No.	GENE	No.	GENE
1	JCHAIN	251	RAD50
2	CCDC144B	252	EIF3M
3	KLRC4	253	TMBIM1
4	NOTCH2NL	254	TMEM126B
5	HLA-DRB4	255	YWHAB
6	CENPS	256	RABEP1
7	CDR1	257	SNRPB2
8	BMS1P5	258	GZMK
9	RNU5B-1	259	H3F3A
10	VTRNA1-1	260	AOAH
11	CD2	261	TIMM17A
12	RPS4Y1	262	KRT6C
13	GPR89B	263	EIF3A
14	S100A8	264	ZNF638
15	IGLV@	265	EOMES
16	KLRC4-KLRK1	266	ZNF700
17	HIST1H2BI	267	SLAMF6
18	LOC440434	268	PRH1
19	PMS2P4	269	LDHB

20	TCEAL6	270	HSP90AB1
21	LOC105375030	271	PAIP2
22	DHRSX	272	HIST1H2BM
23	MIR21	273	HADHA
24	COX7B	274	FMC1
25	TRAPPC2B	275	P2RY10
26	GPR174	276	ENY2
27	LRRC37A	277	SEC61G
28	ULBP2	278	IGHG4
29	SELENOP	279	HDGFRP3
30	FCER1G	280	IGKV3D-15
31	APOBEC3A	281	ANXA5
32	FAM230B	282	CTSB
33	IGHA2	283	USP7
34	GOLGA8A	284	TOMM5
35	ANKRD20A3	285	PTBP3
36	GOLGA8B	286	SMC3
37	SPECC1L	287	DNAJA2
38	FAM95B1	288	ARAP2
39	STAG3L3	289	EIF1AY
40	SNORD75	290	HIST1H2AE
41	LOC100233156	291	CLEC2D
42	NPIPB5	292	SNORD36B
43	KLRC1	293	PPP6C
44	OR2W3	294	TCEA1
45	P2RX5	295	GNAQ
46	ARPC3	296	ZNF474
47	FAM24B	297	U2AF1
48	THRAP3	298	NDUFB6
49	RPSA	299	DICER1

50	LDHA	300	SNORA33
51	RNU5D-1	301	SNORD12C
52	OVOS	302	BCAP29
53	OR10A2	303	SLMAP
54	SNORD116-14	304	HMGB2
55	TMEM189	305	FRG2
56	DDT	306	SLC7A5P2
57	LOC105371030	307	TOP1
58	HIST2H2AB	308	ROCK1
59	FAM21A	309	OCR1
60	LRRC37A4P	310	PFDN4
61	CXCL9	311	ATP1B3
62	PCDH11Y	312	HIST1H3D
63	PMS2P5	313	SNORA46
64	IGKV2-24	314	HIST1H4D
65	OR2T8	315	RAB18
66	MOP-1	316	RARRES1
67	AGAP6	317	MIR9-1
68	KLRC2	318	DAD1
69	USMG5	319	GPR65
70	THOC7	320	SCIMP
71	ANXA1	321	POLR2K
72	GYG1	322	MSH5
73	HSFY1	323	MRPL35
74	ABCC6	324	RPL4
75	SOD2	325	HLA-DQA1
76	NUCB2	326	MMADHC
77	ZNF737	327	SNRPF
78	SNORD44	328	RAB21
79	CD164	329	GOLGA8F

80	GCOM1	330	CHKB
81	CPT1B	331	LSM12
82	UBE2F	332	SHFM1
83	CLEC2B	333	CFLAR
84	MUC3A	334	TRAV20
85	NDUFA4	335	RBBP8
86	HSPA6	336	CDK12
87	GBP1	337	HSPB11
88	IFNG	338	A2M
89	RYBP	339	PCM1
90	PTPN22	340	IGLV1-44
91	ACAD11	341	DTHD1
92	CCL15	342	PSMB4
93	CD163	343	GOLGA8K
94	DPP4	344	SNORD38B
95	FCGR2A	345	ENPP2
96	PDE7A	346	SLX1A-SULT1A3
97	RRN3P1	347	ARRDC3
98	PWAR6	348	RPF1
99	LOC389765	349	ZNF721
100	SNORD78	350	ELK2AP
101	LGALS9C	351	MGAT4A
102	PILRB	352	HNRNPK
103	HIF1A	353	TMX1
104	GBP5	354	SNORD28
105	PLA2G7	355	ZNF733P
106	MT1H	356	FAM27E2
107	WARS	357	PRPF4B
108	KIR2DS4	358	SH3GLB1
109	EEF1E1	359	SAA1

110	SERPINB1	360	ACSL1
111	CCL2	361	ATP2A2
112	IFNA17	362	SNORD115-24
113	NME1	363	PSMA2
114	KIAA0408	364	AGTPBP1
115	KLRC3	365	CAP1
116	CYTIP	366	FN1
117	ABCE1	367	CCDC59
118	SNORA20	368	TCP1
119	ABCB8	369	GLRX
120	PLEK	370	UGP2
121	NCKAP1L	371	AIM2
122	ELOVL5	372	MAGEA3
123	LYST	373	SQRDL
124	FGF7P6	374	MLEC
125	PSTPIP2	375	PPP1R12A
126	ANKRD22	376	PUM2
127	PFDN5	377	LYZ
128	ELL2	378	HEXB
129	FGF7P3	379	ATP6V1C1
130	HSP90B1	380	TRMT13
131	NUCKS1	381	IQGAP1
132	COMMD3	382	SEPT7
133	SNORA5A	383	EFR3A
134	SLC2A3	384	RPS3
135	LY75	385	RBM12
136	C4A	386	PPA1
137	ARHGAP15	387	MAP4K4
138	DDX12P	388	KIR3DL1
139	SYF2	389	ADAM10

140	LAMP2	390	CMC4
141	RAD21	391	EIF3L
142	IGLV3-25	392	HSPA9
143	CCNH	393	SAT1
144	DDX18	394	CORO7
145	FCGR2C	395	CYBB
146	INO80B	396	GOLGA6C
147	SNORD4B	397	PITPNB
148	TMEM110	398	NDUFC2
149	SAMSN1	399	UTS2
150	HIST1H3G	400	ASPM
151	CKS2	401	RPS17
152	HSD17B7	402	WTAP
153	MIRLET7F1	403	FAM133B
154	CCL18	404	C1QA
155	PSMA4	405	EIF5B
156	POTEE	406	GNS
157	TRAC	407	TAX1BP1
158	ATXN1	408	CBWD2
159	EVI2A	409	IFIH1
160	CSE1L	410	TES
161	LSM6	411	MSL2
162	CRYGD	412	ATP5C1
163	SNORD116-1	413	CD3G
164	EIF3E	414	GAPDH
165	TIMM23	415	GBP4
166	FAM49B	416	PARP4
167	PSMC2	417	KTN1
168	FGL2	418	IGKC
169	DDX60L	419	MTHFD1

170	RAB8B	420	IQGAP2
171	SF3B6	421	IVNS1ABP
172	UBE2L3	422	ERAP2
173	TMEM56	423	CSDE1
174	TNFSF13B	424	CD3E
175	RTEL1	425	PSIP1
176	SNORD61	426	EXOC5
177	SLAMF7	427	ZNF587B
178	CDC123	428	CCL5
179	PA2G4	429	POM121C
180	INPP4B	430	SNX2
181	RGS1	431	SRI
182	GZMA	432	CCNI
183	IDO1	433	RBMY1A3P
184	JPX	434	BTBD1
185	CASP4	435	TPTE2P2
186	LOC100130278	436	ABI3BP
187	HSPH1	437	SMIM11A
188	S100A10	438	SNORD117
189	PIP4K2A	439	RAP1B
190	KIF2A	440	PIAS1
191	GTF2H2	441	EPRS
192	SUPT16H	442	RSL24D1
193	KLRG1	443	KRTAP10-9
194	DDX1	444	ATF2
195	CALR	445	CNOT1
196	CSH1	446	XRN1
197	GALNT1	447	COPB2
198	GTF3A	448	CENPW
199	ERBIN	449	CYB5B

200	SNORD41	450	FAM200B
201	SNORD56B	451	MCUR1
202	SNORD13P1	452	SNORA68
203	PYHIN1	453	DOCK10
204	CLU	454	ZC3H11A
205	CXCL11	455	ENTPD1
206	SRP72	456	MNDA
207	SNORD60	457	POLE4
208	TNFSF10	458	ZCRB1
209	LSM3	459	BHLHE40
210	SATB1	460	FCF1
211	RSL1D1	461	SH3KBP1
212	POMP	462	STON1-GTF2A1L
213	ITGA4	463	NPAT
214	C9orf78	464	MBD2
215	NUDT21	465	ERO1A
216	PAK2	466	FAM156B
217	DCUN1D1	467	SCARNA4
218	GTF2IRD2B	468	TRAM1
219	EIF2A	469	TLK1
220	GK	470	ME2
221	RPS6KA3	471	PRKACB
222	CD24	472	DENND4A
223	MDM2	473	PARP14
224	OSBPL8	474	CRYAB
225	GDNF	475	TMF1
226	EEF2K	476	RAB39B
227	ELF1	477	ATP5O
228	ANKHD1	478	HCRP1
229	RPS12	479	DNAJB11

230	SEPT5	480	HCLS1
231	SNORD116-24	481	GUSBP3
232	GNLY	482	SERPINB9
233	TYMS	483	HELZ
234	JAK1	484	SH2D1A
235	TSNAX	485	STIP1
236	CD274	486	RRP15
237	HBB	487	KPNA2
238	TIPRL	488	RACK1
239	CTSS	489	C11orf58
240	CD96	490	FABP5
241	ACTR3	491	GOLGA6B
242	KLRD1	492	GTF2I
243	HAUS6	493	NT5C3A
244	TXN	494	TLR4
245	CXCL10	495	KCNA3
246	NEDD8	496	EIF2S1
247	NAMPT	497	WDR43
248	H2AFZ	498	PICALM
249	SYNJ2BP	499	SSB
250	DMXL2	500	ARID4B

Table S3 Expression of CD8 and CD56 immunohistochemistry and cell of origin in the N-group and EN-group cases

			T-origin	NK-origin	Indeterminate origin	TOTAL
EN-GROUP (n=39)	CD8-/CD56+	23	5	18	0	23
	CD8-/CD56-	10	4	2	4	10
	CD8+/CD56+	5	1	3	1	5
	CD8+/CD56-	1	1	0	0	1
N-GROUP (n=17)	CD8-/CD56+	2	0	2	0	2
	CD8-/CD56-	4	3	0	1	4
	CD8+/CD56+	2	2	0	0	2
	CD8+/CD56-	9	9	0	0	9
TOTAL			25	25	6	56

10 cases with incomplete CD8 and CD56 IHC data are not included in the table.
Cases with indeterminate origin are negative for both TCRB and TCRG IHC but lacks clonality data.

Table S4. Multivariate survival analysis using clinicopathologic factors significantly different between N- and EN-groups.

Covariate	**Hazard Ratio**	**95% Confidence Interval**	**P-value**
Age*	1.01	0.98-1.05	0.54
Stage: I (ref) vs II to IV	2.06	1.32-3.22	0.0015
Disease Presentation: Nodal vs extranodal (ref)	4.04	1.3-12.51	0.016
Lineage: NK vs indeterminate (ref)	0.45	0.08-2.6	0.37
Lineage: T vs indeterminate (ref)	0.38	0.06-2.32	0.3
Nasal involvement : Present vs Absent (ref)	1.3	0.36-4.72	0.69

Ref: reference, *; age is assessed as a continuous variable

Table S5 Differentially expressed genes (DEG) between the N-group and EN-group (p<0.05, False Discovery Rate (FDR) <0.15, and Fold Change of ≥1.5)

A) Genes downregulated in nodal compared to extranodal group cases

No.	Symbol	Fold change	p-value	FDR
1	SPAG11A	2.752022755	0.000536472	0.06046579
2	POM121	2.73787768	0.006358416	0.108848315
3	HSPA6	2.313897848	0.006791769	0.111021531
4	KLRC1	2.236197649	0.00375309	0.092378803
5	IGHA2	2.140101459	0.012337541	0.130880083
6	TRIM49B	2.098392076	9.79E-05	0.049884668
7	ZNF733P	2.073871991	0.006579211	0.110222882
8	KLRC3	2.030809531	0.007921153	0.115341606
9	POSTN	2.010623514	0.000168494	0.055643815
10	SNORD113-3	2.00943472	3.41E-05	0.031600548
11	HSPB2	1.991362451	0.000288137	0.055643815
12	SNORD115-24	1.98406037	0.000518955	0.060340969
13	OR2T8	1.983829361	0.013812237	0.134530201
14	SERPINB3	1.980606516	0.000431284	0.058833612
15	AQP7P1	1.93274229	0.003784002	0.092378803
16	SH2D1B	1.931353467	0.000114292	0.050800369
17	MYZAP	1.908823486	0.005534174	0.105346258
18	SNORD113-4	1.904767316	0.000332296	0.055643815
19	CRYAB	1.903300013	0.001202919	0.071703864
20	FJX1	1.891186474	0.007706776	0.115239787
21	KLRC2	1.864937208	0.004718793	0.099965614

22	CD56 (NCAM1)	1.863394822	0.002867916	0.088566319
23	DEFB134	1.843630661	0.012251465	0.130700778
24	TSPAN8	1.841603079	0.0003026	0.055643815
25	RBMY1A3P	1.83560949	0.009987733	0.122826729
26	PIGR	1.82425852	0.00039374	0.058808812
27	MIR29A	1.801127343	0.000467875	0.059605989
28	HIST1H2AI	1.77312148	0.000259841	0.055643815
29	SNORD114-2	1.756407568	0.000617265	0.062455931
30	REG3A	1.753043521	0.001077218	0.071677338
31	ACOT1	1.742606035	0.000979269	0.071481072
32	MMP1	1.735956559	0.000466171	0.059605989
33	DEFB1	1.725891234	9.53E-05	0.049884668
34	SEPT7P9	1.723790525	0.003514055	0.090650839
35	AGR2	1.714314889	0.000208114	0.055643815
36	STATH	1.71074161	0.009648338	0.121289854
37	SAA2	1.694020058	0.002828698	0.088361283
38	SLFN13	1.689739797	0.002414576	0.087075086
39	CEACAM5	1.683523029	0.004456014	0.096959066
40	SLC4A10	1.679229271	2.53E-06	0.009640218
41	ZNF676	1.677101829	0.001491654	0.07497069
42	PLA2G10	1.669811312	0.00365259	0.09233847
43	SNORD114-3	1.661990354	0.008210489	0.115341606
44	UGT2B17	1.660172117	0.000874754	0.068813834
45	PLA2G2A	1.654146769	0.000183065	0.055643815
46	MEDAG	1.653607195	8.29E-05	0.049884668
47	S100A2	1.652834562	0.005484759	0.105309656
48	SULF1	1.650235645	0.002920924	0.088718267
49	TFPI2	1.640772478	5.04E-05	0.039175535
50	OR13C2	1.637028059	0.000186553	0.055643815

51	REG1A	1.632304812	0.003961696	0.092378803
52	EPCAM	1.631282165	0.000482379	0.059605989
53	SNORD1A	1.630094788	0.012897848	0.132733421
54	SLPI	1.62981236	0.018464045	0.145714351
55	DCN	1.623027824	0.010562448	0.124718574
56	Sep-10	1.616908908	0.002643373	0.087727213
57	SFRP2	1.616022404	0.005123126	0.101892325
58	TOMM7	1.604805558	0.000795938	0.068813834
59	SIGLEC17P	1.600014861	0.002499392	0.087131118
60	UGT2B28	1.597120494	0.001880577	0.081447106
61	LRRC66	1.586513274	7.01E-05	0.048449737
62	MSMB	1.585296096	0.002565499	0.087490105
63	POTEC	1.581898268	0.017425886	0.142960638
64	OR8H3	1.580503414	0.001249755	0.07234252
65	RNASE4	1.57912689	0.006944688	0.111798456
66	SLFN14	1.577384531	0.000540535	0.06046579
67	UGT2B15	1.5650055	0.002251575	0.083646712
68	MMP3	1.561452136	0.005917026	0.106984411
69	SCHIP1	1.557227862	0.006829324	0.111021531
70	SNORD7	1.556930649	0.00122542	0.071703864
71	CFHR1	1.555752693	0.014789454	0.136854896
72	EHF	1.5519857	0.000211363	0.055643815
73	FBXW10	1.551239624	0.014221967	0.136127417
74	KIR2DL4	1.54952909	0.017600046	0.143273554
75	TVP23C	1.544886038	0.002474583	0.087075086
76	TSPAN1	1.54330625	0.000130211	0.053344665
77	POU5F1	1.537683878	0.007008485	0.111798456
78	OR1J1	1.537674247	0.019559671	0.149353904
79	PIP	1.535316381	0.00590071	0.106984411

80	CFHR2	1.53522801	0.006676311	0.110486678
81	TMEM244	1.533828554	0.001878106	0.081447106
82	CDH17	1.532902578	0.000378365	0.058808812
83	ITGA1	1.529751333	0.017750077	0.143631963
84	MIR29B1	1.528479375	0.008538386	0.116711852
85	KRTAP12-2	1.523274212	0.007933628	0.115341606
86	FOXD2	1.519537047	0.00825884	0.115358214
87	CAMK2N1	1.512897411	0.000504529	0.060340969
88	PCAT4	1.512524772	0.005874164	0.106880173
89	FOXP2	1.511109247	8.21E-06	0.018168003
90	INHBA	1.509171843	0.00277549	0.088150892
91	SNORD115-33	1.503910687	0.001406446	0.073122324
92	GNGT1	1.503201013	0.003556865	0.091459582
93	MUC13	1.501311972	0.001904414	0.081447106

B) Genes upregulated in nodal compared to extranodal group cases

No.	**Symbol**	**Fold change**	**p-value**	**FDR**
1	MT1H	3.198607775	0.002746399	0.088150892
2	HIST1H2BI	3.191816504	0.000434917	0.058833612
3	CD3G	3.160975967	3.66E-07	0.003197336
4	GPR89B	3.133868248	0.012386039	0.131078757
5	ADAMDEC1	2.963129885	0.000552696	0.06046579
6	CD3D	2.920820951	0.00023652	0.055643815
7	THEMIS	2.823981741	0.001133232	0.071677338

8	PLA2G7	2.742710687	0.00088902	0.068813834
9	GPR174	2.731299024	0.002463297	0.087075086
10	ULBP2	2.710086742	0.011553016	0.128353156
11	ENPP2	2.672445720	0.000193565	0.055643815
12	CHI3L1	2.643527017	0.000885413	0.068813834
13	CD2	2.638380463	0.00527322	0.103187237
14	HIST1H4D	2.540562512	1.81E-05	0.028678054
15	MSH5	2.455639029	0.010472634	0.12436604
16	SNORA24	2.445075798	0.000912331	0.068813834
17	CCL19	2.376857525	5.14E-07	0.003197336
18	AGAP6	2.345487543	0.017243617	0.142791776
19	IL18	2.325313260	0.001133931	0.071677338
20	HIST1H2AE	2.310887436	0.001245432	0.07234252
21	CXCL9	2.283823709	0.008102457	0.115341606
22	CLU	2.280100585	0.000584957	0.061694854
23	GPR137B	2.278636641	0.000194517	0.055643815
24	ABI3BP	2.237063204	0.000805914	0.068813834
25	FCER1G	2.236071040	0.014606667	0.136497006
26	TMEM110	2.193587388	0.00400883	0.092378803
27	CTSB	2.158524499	0.002939469	0.088718267
28	CYBB	2.150772589	0.001161682	0.071703864
29	DENND2D	2.127225579	0.000236771	0.055643815
30	LCP1	2.116200749	0.001939644	0.081659013
31	CXCL10	2.094548400	0.017749398	0.143631963
32	SDCBP	2.086453434	0.001055753	0.071645798
33	GBP1	2.076209422	0.014792446	0.136854896
34	LEF1	2.067098565	0.000215889	0.055643815
35	VCAM1	2.053309359	0.002349951	0.086017532
36	CCL18	2.044612734	0.003152218	0.089648643

37	CD28	2.025651437	0.00051139	0.060340969
38	CD3E	2.021862094	0.003905216	0.092378803
39	SLMO2	2.002434787	0.002556688	0.087490105
40	LYST	1.998430548	0.007369609	0.113934632
41	HIST1H2BM	1.997351957	0.000593032	0.061826677
42	PSMB4	1.996276507	0.006034783	0.10760028
43	CD53	1.995806736	0.000658216	0.065013691
44	TXN	1.995647510	0.003091724	0.089648643
45	CSF2RB	1.988439413	2.50E-05	0.031108092
46	HLA-DQA1	1.977780757	0.005234716	0.103128054
47	HIST1H4L	1.974418702	0.000909549	0.068813834
48	CXCL11	1.974084710	0.01450475	0.136497006
49	HSPA8	1.972279736	0.010738119	0.125365539
50	GNS	1.970637759	0.003644996	0.09233847
51	HIST1H4H	1.960255701	0.004717804	0.099965614
52	CD8A	1.951339360	0.000344014	0.055643815
53	CNN2	1.944967330	0.00129085	0.072826782
54	STAT1	1.923116677	0.000897469	0.068813834
55	HIST1H2BE	1.922874216	6.03E-05	0.044142123
56	WARS	1.919290883	0.013640619	0.134214201
57	CHI3L2	1.916094763	0.001792608	0.080832023
58	GM2A	1.901418699	0.015530036	0.138234205
59	CTSL	1.885378385	0.012746824	0.132733421
60	ZNF841	1.884396199	0.012845296	0.132733421
61	COMMD3	1.883356505	0.002894073	0.088718267
62	GBP4	1.876347798	0.012836374	0.132733421
63	SMPDL3A	1.875800528	0.004115793	0.092559764
64	SMAP2	1.875679352	0.002466571	0.087075086
65	NPL	1.867044071	0.003452697	0.090463207

66	IFNGR2	1.865641406	0.001495023	0.07497069
67	CD274	1.862705629	0.017241287	0.142791776
68	LINC00266-1	1.862099974	0.001741677	0.080825881
69	BIRC3	1.862034259	0.001804578	0.081078026
70	IFNAR1	1.861794512	0.002770118	0.088150892
71	HSD11B1	1.860434535	0.000596142	0.061826677
72	SLC7A5P1	1.858683589	0.01553885	0.138234205
73	SGTB	1.857294604	0.000100207	0.049884668
74	RNU4-2	1.856464829	0.003163311	0.089648643
75	TUBB4A	1.855598707	0.00133959	0.073122324
76	ADA	1.853605758	0.000876878	0.068813834
77	STAMBPL1	1.852825183	0.004820047	0.099965614
78	EOMES	1.850281173	0.017339689	0.142960638
79	NCKAP1L	1.842757572	0.008512756	0.116550323
80	GABARAP	1.840570920	0.002018906	0.081659013
81	UCP2	1.839723538	0.000398618	0.058808812
82	RPS24	1.836608343	0.010634256	0.124973614
83	PSAP	1.830395907	0.001512022	0.07497069
84	TMEM56	1.826694966	0.012169082	0.130315464
85	PAG1	1.825983999	0.006615128	0.110486678
86	SCARNA1	1.819402121	0.009385952	0.120536614
87	SRP14	1.819104246	0.001928676	0.081659013
88	MT1JP	1.819094552	0.015716824	0.138234205
89	UBE2D1	1.814734350	0.012983511	0.132733421
90	HLA-DQB1	1.814389365	0.00324609	0.089883967
91	HIST1H3F	1.811179501	0.002670445	0.087922307
92	GPR65	1.805810349	0.003109656	0.089648643
93	ATP6AP1	1.802362609	0.002722146	0.088150892
94	COTL1	1.801237268	0.001182873	0.071703864

95	ATM	1.801050838	0.013047426	0.132793468
96	FYB	1.798388025	0.013694455	0.134214201
97	HIST1H1D	1.797259387	0.000253045	0.055643815
98	HIST1H2BN	1.796704962	0.00561599	0.105654844
99	HSP90AB1	1.794418654	0.014426536	0.136403828
100	UHMK1	1.791259044	0.001365876	0.073122324
101	CMKLR1	1.790705667	0.001028104	0.071481072
102	SERPINB9	1.788035489	0.01164536	0.128598584
103	DICER1	1.785080904	0.007893811	0.115341606
104	ALAS1	1.782868716	0.003118178	0.089648643
105	IL10	1.780708172	0.002109413	0.081659013
106	RAC2	1.780369254	0.003020047	0.088963367
107	ELOVL5	1.779852926	0.01395499	0.135366145
108	CADM1	1.778077912	0.00189093	0.081447106
109	TNFAIP2	1.773708042	0.000299313	0.055643815
110	RNF213	1.772360823	0.005603431	0.105654844
111	PPA1	1.771568552	0.008431626	0.115965844
112	KIAA1551	1.766541408	0.016946177	0.142120942
113	SLC39A8	1.762658580	0.008126535	0.115341606
114	TMEM30A	1.761055690	0.002709129	0.088150892
115	CREG1	1.760045813	0.005774237	0.106305333
116	LGMN	1.752338788	0.004530044	0.097390266
117	RGL1	1.751073557	0.003990206	0.092378803
118	GRINA	1.750256888	0.00185867	0.081447106
119	ERH	1.741958778	0.00630398	0.108848315
120	LIMS1	1.738007164	0.006994866	0.111798456
121	FTL	1.735750745	0.00347359	0.090568757
122	CD68	1.733731266	0.016308055	0.140055385
123	KCNJ10	1.731434483	3.10E-06	0.009640218

124	CD81	1.722429268	0.009814264	0.122019942
125	PRKCB	1.720101276	0.001733413	0.080797522
126	TAP1	1.716159761	0.003045127	0.089381319
127	GYPC	1.714852425	0.000344271	0.055643815
128	AGTPBP1	1.714823999	0.006519843	0.110222882
129	POMP	1.710099297	0.014518274	0.136497006
130	ARL2	1.707930435	0.004407451	0.096393361
131	CRTAM	1.707894512	0.015933358	0.138862916
132	FAS	1.706527675	0.002153217	0.082201067
133	C1QB	1.698885565	0.000479818	0.059605989
134	HIST1H2BL	1.697279660	0.003761334	0.092378803
135	HIST2H4B	1.694139056	0.010257514	0.123469201
136	RNU4-1	1.693899145	0.000517506	0.060340969
137	HIST1H4I	1.692124920	0.005122721	0.101892325
138	SELL	1.689122058	0.008154554	0.115341606
139	TLR8	1.689099516	0.017314885	0.142960638
140	ABCD2	1.688857959	0.016944273	0.142120942
141	C3orf38	1.688345403	0.000443792	0.05938864
142	RNF13	1.688125943	0.002928807	0.088718267
143	HLA-DMA	1.685960855	0.000448601	0.059393587
144	PTPN1	1.682396811	0.001065472	0.071645798
145	MYL12A	1.680948735	0.005475432	0.105309656
146	ITGB2	1.680236266	2.34E-05	0.031108092
147	ADAM19	1.679403549	0.001970854	0.081659013
148	DDX5	1.675547957	0.008039477	0.115341606
149	IRAK4	1.673954098	0.001163717	0.071703864
150	PIAS1	1.665275665	0.015008782	0.137558831
151	RASGRP1	1.662312866	0.003851125	0.092378803
152	ATOX1	1.662308426	0.009457415	0.120536614

153	LILRB4	1.661862580	0.003990915	0.092378803
154	TOP2A	1.661787630	0.010381429	0.124077787
155	SERPINE1	1.661315292	0.013600073	0.134214201
156	DNM1P46	1.654518930	0.003404952	0.090463207
157	MYL12B	1.653631304	0.008924732	0.118539411
158	NR1H3	1.653113032	7.88E-05	0.049884668
159	VCP	1.651442035	0.002029216	0.081659013
160	STK17B	1.650746533	0.011583972	0.128376321
161	UBD	1.647670114	0.003964954	0.092378803
162	NAIP	1.642693978	0.015443122	0.138234205
163	RPL4	1.640822498	0.01628134	0.140055385
164	SNORD94	1.638986441	0.017020575	0.142331631
165	RASSF4	1.638739442	0.001506103	0.07497069
166	ADTRP	1.636723127	0.003720037	0.092378803
167	HMOX1	1.636408070	8.76E-06	0.018168003
168	MADD	1.635351176	0.00749992	0.114526556
169	MSL2	1.634733016	0.019802915	0.149651222
170	CBLL1	1.633764444	0.013563947	0.134187669
171	OLA1	1.632046570	0.01182902	0.129094977
172	LAP3	1.631293411	0.012346233	0.130880083
173	TIFA	1.630751960	0.00589636	0.106984411
174	SGPP2	1.625196655	0.000602789	0.061999351
175	COPB1	1.624701000	0.011761195	0.128885247
176	ATG7	1.623524085	0.000678552	0.065463661
177	PNRC1	1.621727494	0.000555429	0.06046579
178	FCRL3	1.621233559	0.016284156	0.140055385
179	FAM91A1	1.619859288	0.02059429	0.14997261
180	RNF144B	1.619689001	0.006236287	0.108666257
181	IGF2R	1.616818805	0.003357737	0.090450683

182	IREB2	1.615427348	0.003605928	0.091773104
183	EIF3L	1.613704843	0.019389227	0.149353904
184	CPM	1.613461671	0.013436464	0.13406093
185	SCD	1.612815978	0.004027653	0.092378803
186	TNS3	1.612315164	0.003239398	0.089883967
187	NIN	1.612212673	0.008503752	0.116550323
188	CAPG	1.611691405	0.005696979	0.105713659
189	CYCS	1.611492794	0.009981202	0.122826729
190	MPP1	1.611475306	0.004253864	0.094537196
191	CTSH	1.607650694	0.006221122	0.108666257
192	HIST1H4B	1.607647852	0.018419044	0.145543797
193	YWHAZ	1.606749166	0.01186112	0.129094977
194	DECR1	1.605914207	0.007830219	0.115295005
195	HNRNPLL	1.605162074	0.019521014	0.149353904
196	CST3	1.604933014	0.000701652	0.066331864
197	LRBA	1.603361794	0.019504844	0.149353904
198	TRAT1	1.596587529	0.010262834	0.123469201
199	HNRNPF	1.595108326	0.005169587	0.102447969
200	HMGCS1	1.593274281	0.001297375	0.072826782
201	ITM2B	1.591995253	0.009092308	0.119489941
202	BMP2K	1.591044539	0.009755497	0.121653901
203	ATP6V0E1	1.590953883	0.011978018	0.129721211
204	SEC23B	1.590230344	0.010081278	0.123367794
205	ASAH1	1.588297243	0.0146107	0.136497006
206	HIST1H3J	1.587711933	0.011501751	0.128264644
207	ETS1	1.587349284	0.009518983	0.120884793
208	IGSF6	1.584296910	0.006917276	0.111657501
209	HIST1H3B	1.581922559	0.017839833	0.143809556
210	NCOA4	1.581615874	0.018752863	0.147153836

211	RSRP1	1.578722542	0.008240599	0.115341606
212	COPA	1.577450426	0.00683152	0.111021531
213	PRELID1	1.576148584	0.001049884	0.071645798
214	TRAJ17	1.575449524	0.014004049	0.135736224
215	ARL6IP5	1.575307092	0.000316083	0.055643815
216	FTH1	1.574858380	0.019189598	0.148983964
217	MARCKS	1.574286671	0.003006082	0.088963367
218	TFEC	1.573175078	0.019784733	0.149651222
219	ABI1	1.572926331	0.018069563	0.144247636
220	RPS15A	1.572381635	0.008852428	0.118313181
221	ACTR2	1.572059513	0.015514804	0.138234205
222	CD4	1.569697956	0.009429867	0.120536614
223	CORO1A	1.569502409	0.009714672	0.12137789
224	DAZAP2	1.568960721	0.017043834	0.142331631
225	UBE2J1	1.567815834	0.003997185	0.092378803
226	POLK	1.567808505	0.001989814	0.081659013
227	CHIT1	1.566364327	0.011538643	0.128353156
228	RNASET2	1.565783744	0.001394793	0.073122324
229	MBOAT1	1.565754125	0.009686047	0.12137789
230	SNX1	1.565075491	0.006994312	0.111798456
231	SLC25A46	1.563809526	0.005966126	0.107453742
232	LINC00161	1.562951281	0.006684167	0.110486678
233	CDS2	1.562130529	0.00880295	0.118313181
234	SIRPG	1.561047205	0.001105617	0.071677338
235	ATP6V1A	1.560317734	0.008848671	0.118313181
236	SPTBN1	1.559566735	0.004678162	0.099694123
237	IL2RG	1.556242732	0.01793565	0.144247636
238	HNRNPC	1.555365368	0.015314841	0.138045884
239	GLS	1.555102053	0.015995629	0.139016223

240	LPXN	1.554808224	0.003507191	0.090650839
241	SFXN1	1.554108139	0.007683308	0.115167931
242	BCL2A1	1.554090721	0.001395228	0.073122324
243	ARPC1B	1.554069887	0.00048093	0.059605989
244	HSP90AA6P	1.553491070	0.020173145	0.149734356
245	ELMO1	1.552779443	0.011487104	0.128216184
246	HIST1H2AJ	1.552204562	0.002741857	0.088150892
247	TMEM243	1.552102223	0.010828141	0.125826355
248	HIST1H2BC	1.552074419	0.002094161	0.081659013
249	PARP12	1.551571246	0.012467875	0.131491523
250	SS18	1.550331175	0.001188645	0.071703864
251	PSMD11	1.549461028	0.020297827	0.149734356
252	OCIAD1	1.548767917	0.011335118	0.128012282
253	HLA-DMB	1.548554175	0.019914568	0.149734356
254	FAM26F	1.547851220	0.002939433	0.088718267
255	SMIM15	1.547657910	0.01741305	0.142960638
256	LOC400590	1.546678241	0.004067322	0.092378803
257	ANAPC16	1.546462104	0.004145706	0.092897186
258	KMO	1.545830110	0.012431136	0.131444233
259	PIK3AP1	1.542686656	0.009972556	0.122826729
260	LSM2	1.542454507	0.00310243	0.089648643
261	UBE2G1	1.541639689	0.007750747	0.115295005
262	FLNA	1.540982013	0.007493725	0.114526556
263	PPP1R3C	1.540564438	0.003793805	0.092378803
264	TCN2	1.540537670	0.00084452	0.068813834
265	CD27	1.540384781	0.000130593	0.053344665
266	LARP4B	1.539952048	0.017405201	0.142960638
267	TIPARP	1.539780965	0.005527653	0.105346258
268	MMP9	1.539625200	0.011441034	0.128109323

269	MTA2	1.538890570	0.001411259	0.073122324
270	FOXB2	1.538722262	0.005500148	0.105309656
271	UBB	1.535056844	0.002773338	0.088150892
272	CSNK2B	1.534564147	0.015244278	0.138045884
273	MORF4L1	1.533566310	0.004101893	0.092559764
274	TRAFD1	1.533369055	0.010556955	0.124718574
275	AP2B1	1.533223933	0.014026311	0.135824769
276	GPX7	1.532662262	0.007046882	0.111798456
277	EXOSC9	1.532228674	0.008019216	0.115341606
278	NBN	1.531217575	0.020449895	0.149734356
279	HLA-C	1.530921502	0.000483731	0.059605989
280	SIT1	1.530522516	0.000405099	0.058808812
281	TSPAN14	1.529186596	0.002476795	0.087075086
282	PFN1	1.527712050	0.012035186	0.129721211
283	NEIL3	1.526016304	0.00854331	0.116711852
284	CSTB	1.524864739	0.009075245	0.119455847
285	FAM96A	1.524345967	0.010304587	0.123469201
286	CTSD	1.520507871	0.006415387	0.109672725
287	ATF7IP2	1.520044905	0.004989943	0.100820117
288	IL2RA	1.516209255	0.010191606	0.123469201
289	HIST1H2AB	1.515801447	0.012921411	0.132733421
290	USP4	1.514163876	0.012847456	0.132733421
291	EEF2	1.513126407	0.006256997	0.108666257
292	LITAF	1.512016288	0.002027995	0.081659013
293	BASP1	1.511973241	0.00026951	0.055643815
294	ATP5G3	1.511687303	0.00523705	0.103128054
295	C16orf87	1.511163505	0.010633695	0.124973614
296	VEZT	1.509201006	0.005753538	0.106238578
297	MDH1	1.507802990	0.014375376	0.136403828

298	LGALS2	1.506903365	0.004120262	0.092559764
299	ACP2	1.506858127	0.013812085	0.134530201
300	CNOT7	1.506624790	0.017685102	0.143623593
301	CNDP2	1.506378277	0.00451284	0.097337739
302	CARD11	1.503767626	0.001849302	0.081447106
303	AIF1	1.503359045	0.004798258	0.099965614
304	XRN2	1.500567729	0.018008621	0.144247636

Table S6. Recurrent copy number aberrations occurring in at least 20% of Nodal-group and Extranodal-group cases.

A) Recurrent loss in Nodal group cases (n=12)

chr	start	end	cytoband	Frequency (%)	gene count	gene name (selected list)
3	162347085	162672041	chr3q26.1	58	0	
14	22328365	23000062	chr14q11.2	100	50	TRAJ60, TRDV2, TRDV3, TRDJ4, TRDV1, TRD@, TRAV38-1, TRAV35, TRAV36DV7, TRAV40, TRAV41,TRAV39, TRAV33, TRAV34, TRAV12-2, TRAV26-2, TRAV26-1, TRAV28, TRAV29DV5, TRAV32,...
22	24346428	24390318	chr22q11.23	33	6	LOC768328, GSTTP2, LOC391322, LOC100420769, GSTTP1, GSTT1

B) Recurrent loss in Extranodal group cases (n=29)

chr	start	end	cytoband	Frequency (%)	gene count	gene name (Selected list)
3	162373665	162659705	chr3q26.1	34	0	
6	124681503	125097427	chr6q22.31	21	1	NKAIN2
6	132600719	132874814	chr6q23.2	21	5	TAAR8, RPL21P66, STX7, MOXD1, TAAR9
6	135057244	135739355	chr6q23.2-q23.3	21	8	HBS1L, MYB, MEMO1P2, ALDH8A1, MIR3662, AHI1, LOC100420631, MIR548A2

6	136625541	147583868	chr6q23.3-q24.3	21	90	ZC2HC1B, LOC100507477, MRPL42P3, PHACTR2, ABRACL, LOC100131041, LOC100507462, NMBR, FLJ46906, HYMAI, PEX7, RPS3AP23, LOC100130476, LOC100129577, RAB32, LOC100507390, FUCA2, RPL35AP3, TNFAIP3, IL22RA2,...
6	160186920	162627290	chr6q25.3-q26	21	19	SLC22A2, MAS1, LPA, LOC729603, AGPAT4-IT1, SNORA20, IGF2R, PLG, MRPL18, PNLDC1, MAP3K4, ACAT2, SNORA29, SLC22A3, SLC22A1, PARK2, LPAL2, AGPAT4, TCP1
9	21993468	22014524	chr9p21.3	21	4	UBA52P6, CDKN2B-AS1, CDKN2B, CDKN2A
22	24346428	24390318	chr22q11.23	48-59	6	LOC768328, GSTTP2, LOC391322, LOC100420769, GSTTP1, GSTT1

C) Recurrent gains in Extranodal group (n=29)

chr	**start**	**end**	**cytoband**	**Frequency (%)**	**gene count**	**gene name (selected list)**
1	156037357	184599514	chr1q22-q25.3	28	489	TIPRL, METTL13, FMO11P, SONP1, CD48, SHCBP1L, SDCCAG3P2, NIT1, XCL1, FMO9P, RPS17P6, RPL4P3, RPS15P3, LOC391136, SLC19A2, OR6P1, QRSL1P1, ARPC5, LOC100505686, OR10J9P,...

1	185727070	249212878	chr1q25.3-q44	28	754	PPP1R15B, OCLM, C1orf106, LOC100419489, NVL, KCNK2, KCNK1, SNORA77, MIR4742, ZNF124, MYOG, SUSD4, ZNF670, MIR135B, NID1, SPHAR, LEFTY1, LOC100130137, PLEKHA6, C1orf27,...
2	143731244	145490269	chr2q22.2-q22.3	21	5	KYNU, DKFZp686O1327, ARHGAP15, LOC100505471, ZEB2
2	158177489	158422586	chr2q24.1	21	4	ERMN, CYTIP, LOC728066, ACVR1C
2	161181830	161473973	chr2q24.2	21	3	LOC100419108, RBMS1, MIR4785
2	162139812	162414307	chr2q24.2	21	4	TBR1, PSMD14, AHCTF1P1, LOC100506055
2	195980760	198621506	chr2q32.3-q33.1	24	26	SLC39A10, ATP5G2P3, LOC729342, GTF3C3, LOC100419110, RFTN2, HSPD1, E2F3P2, CCDC150, HSPE1-MOB4, LOC391470, HECW2, LOC100130452, C2orf66, PGAP1, BOLL, COQ10B, HSPE1, LOC100507039, MARS2,…
2	199506556	199766349	chr2q33.1	21	0	
6	32340236	32582736	chr6p21.32	21	2	HLA-DRB6, HLA-DRB5
7	88259445	90063761	chr7q21.13	21	9	STEAP2, EEF1A1P28, GTPBP10, STEAP1, C7orf63, CLDN12, ZNF804B, C7orf62, DPY19L2P4
7	90372863	91184277	chr7q21.13-q21.2	21	3	NIPA2P1, CDK14, FZD1

7	92203716	94435106	chr7q21.2-q21.3	21	25	CCDC132, RPS27P17, TFPI2, RPS3AP25, CASD1, FAM133B, ATP5F1P2, MIR653, CALCR, CDK6, LOC100506027, COL1A2, SAMD9, HEPACAM2, SAMD9L, SGCE, MIR489, GRPEL2P3, RN7SL7P, BET1,...
7	99816878	99981895	chr7q22.1	21	6	GATS, PVRIG, SPDYE3, PMS2P1, PILRA, PILRB
7	103871529	104154069	chr7q22.2	21	1	LHFPL3
7	104245132	104619563	chr7q22.2-q22.3	21	4	EIF4BP6, LOC645591, LHFPL3, LOC723809
7	148158904	150483512	chr7q36.1	21	53	GIMAP5, GIMAP2, RNU7-20P, ZNF212, REPIN1, ALDH7A1P3, ZNF777, GIMAP7, LOC643438, LOC100128542, CUL1, LRRC61, COX6B1P1, PDIA4, KRBA1, ACTR3C, ZNF467, ZNF862, GIMAP1, LOC728743,...
7	151699597	159118443	chr7q36.1-q36.3	24	56	RNF32, LOC100132707, ACTR3B, RPS27AP12, XRCC2, MIR595, GALNT11, LMBR1, RPL21P76, LOC100506585, RPL26P23, LOC154822, LOC100506557, WDR60, YBX1P4, LOC100506380, LOC777650, SHH, UBE3C, LOC100506534,...
11	128383077	128519816	chr11q24.3	21	1	ETS1
13	48910204	49004132	chr13q14.2	24	2	RB1, LPAR6

17	45196190	47450057	chr17q21.32-7q21.33	21	68	MRPL45P2, HOXB-AS3, NPEPPS, MIR196A1, LOC100506252, LOC100303748, LOC100506325, KPNB1, UBE2Z, ITGB3, LOC100506228, GIP, TBKBP1, LOC100421523, MYL4, SKAP1, MIR10A, LOC100506373, NFE2L1, ATP5G1,...
17	50106958	50134132	chr17q21.33	21	1	CA10
17	57822608	58015149	chr17q23.1	21	6	RPS6KB1, RPS29P21, NDUFB8P2, TUBD1, MIR21, VMP1
17	59917389	60017888	chr17q23.2	21	2	INTS2, BRIP1
17	65064908	65309815	chr17q24.2	21	2	HELZ, RPL36AP48
17	66011057	66891543	chr17q24.2	21	15	KPNA2, WIPI1, LOC651250, ARSG, MIR635, FAM20A, LOC440461, AMZ2, ABCA8, PRKAR1A, LOC100418928, SLC16A6, SH3GL1P3, LOC732538, LOC100499466
17	73230856	73509541	chr17q25.1	21	11	MIF4GD, KIAA0195, NUP85, GRB2, LOC100287042, RPL36AP7, GGA3, MIR3678, SLC25A19, MRPS7, CASKIN2
17	78244300	78513632	chr17q25.3	21	6	LOC100294362, NPTX1, MIR4730, LOC100507440, ENDOV, RNF213

Chr, chromosome

Table S7. Results of TCR loci in all samples tested with OncoScan

TCR loci deletion status of TNKL samples					
Sample	**Cell Lineage**	**TCRα**	**TCRβ**	**TCRγ**	**TCRδ**
A1	T	-1			-1
MD04	Indeterminate				
MD05*	Indeterminate			-1	
MD10	T	-1			-1
MD11*	Indeterminate				
MD15*	Indeterminate	-1			
NKTL22	Indeterminate				
NKTL23A	T	-1/-2			-2
NKTL24	Indeterminate	1			
NKTL27A	Indeterminate				
NKTL36*	NK				
NKTL40	T	-1			-1
NKTL41*	Indeterminate				
NKTL43*	Indeterminate			-1	
NKTL50*	Indeterminate	-1			
NKTL52*	Indeterminate				
NKTL53	Indeterminate				

NKTL59*	Indeterminate	-1/-2			-2
NKTL60	T	-1/-2			-2
T114	T	-1	-1		-1
T202	T	-2			-2
T241	NK				
T242	T	-1			-1
T25	T				
TNK1*	Indeterminate	-1			-1
TNK10	T	-1/-2			-2
TNK3	T	-1			-1
TNK4	T	-2	-1		-2
TNK6	T	-1			-1
TNK9	T	-1/-2			-2
TW7	NK				
TW13*	Indeterminate				
TW22	T				
TW26*	Indeterminate				
TW27*	Indeterminate				
TW29*	Indeterminate				
TW33	NK				
TW34*	Indeterminate				

TW37	NK				
TW38	NK				
TW39	T				

Diagnosis	**Diagnostic category**	**TCRA**	**TCRB**	**TCRG**	**TCRD**
AITL	T cell lymphoma				
ALCL	T cell lymphoma	-1			-1
PTCL	T cell lymphoma	-1			-1
Tonsil	Benign				
Tonsil	Benign				
Lymph Node	Benign				
NK cells (1)	Benign				
NK cells (2)	Benign				
T cells (1)	Benign	-1/-2	-1		-2
T cells (2)	Benign	-1/-2	-1	-1	-2
B cells (1)	Benign				
B cells (2)	Benign				
PTCL	T cell lymphoma	-1/-2			-2
AITL	T cell lymphoma	-1			-1
ALCL	T cell lymphoma	-2	-2		-2

MCL; mantle cell lymphoma, MZL; marginal zone lymphoma, AITL; angioimmunoblastic T cell lymphoma, DLBCL; diffuse large B cell lymphoma, FL; follicular lymphoma, ALCL; anaplastic T cell lymphoma, PTCL; peripheral T cell lymphoma.

* denotes cases with absence of TCRG/TCRB expressionand lacking TCR clonality data and. COO assigned based on combination of CD56 expression and/or TCR loci information

Table S8. Association between TCR loci loss and cell lineage

P<0.0001	**TCR loci**		
Lineage	Loss	Intact	
T-origin (n=24)	20	4	24
Non-T-origin (n=9)	0	9	9
	20	13	33

T-origin samples includes TNKL of T-origin (n=16), PTCL (N=6) and normal T-cells (n=2). Non-T samples include TNKL of NK origin (n=5), normal B cells (n=2) and normal NK cells (n=2). Only those TNKL cases with complete TCRB/TCRG IHC and clonality data were included for this analysis. TNKL cases denoted with (*) in supplemental Table 1 were excluded for this analysis.

Table S9. Differentially expressed genes (DEG) between the Nodal-group and lymph node control tissues ($p<0.05$, False Discovery Rate (FDR) <0.15 and fold change ≥4.0)

A) Genes downregulated in nodal compared to control cases

No.	Symbol	Fold change	p-value	FDR
1	PSG10P	12.0893	0.005268	0.086546
2	SNX29P1	8.982119	0.000233	0.027088
3	SNORD38A	8.437966	0.001454	0.057091
4	SNORD115-24	7.990969	0.023649	0.127334
5	SNORD117	6.990204	0.006598	0.090213
6	MIR30C2	5.843239	0.00201	0.062918
7	SNORD55	5.723527	0.002954	0.07092
8	IGHV3OR16-7	5.151226	0.029267	0.134647
9	SNORD42A	4.750073	0.029835	0.13551
10	SCARNA4	4.495444	0.000126	0.022036
11	MIR622	4.422818	0.000944	0.050073
12	ZNF709	4.405266	0.007408	0.092071
13	PRADC1	4.338643	0.006369	0.089917
14	MIR30C1	4.230802	0.013879	0.109141
15	RNASE1	4.051542	0.000133	0.022392

B) Genes upregulated in nodal compared to control cases

No.	Symbol	Fold change	p-value	FDR
1	RSPH10B2	19.5085	0.021828074	0.123967389
2	CXCL10	18.0743	1.78E-08	0.000213117
3	SENP3	17.6259	0.001889945	0.061785449
4	HIST2H2AA3	14.3689	0.000439776	0.038061433
5	CXCL9	14.2346	4.74E-06	0.007119642
6	KLRC2	13.5148	0.000854121	0.049111583
7	GZMK	11.5879	9.68E-05	0.019022759
8	CENPS	9.3318	0.003408028	0.07443628
9	OR2W3	9.3005	0.028342249	0.133077304
10	LGALS9C	9.1842	0.001023157	0.051692812
11	FCGR3A	8.9654	7.90E-07	0.003155627
12	SNORD56B	8.9202	0.004784764	0.083513677
13	GZMA	8.8795	5.94E-07	0.003155627
14	SNORD102	8.8791	0.019856982	0.121471441
15	IFNG	8.5649	5.83E-05	0.016269408
16	FCER1G	8.4917	0.000104836	0.019953767
17	GBP1	8.4726	0.000173246	0.024348196
18	SNORA22	8.3726	0.006508978	0.089916595

19	COX7B	8.2459	2.73E-05	0.011215562
20	GBP5	7.9687	3.66E-05	0.012927228
21	TXN	7.6375	4.93E-06	0.007119642
22	SNORA24	7.5750	0.015877016	0.113170809
23	MT1H	7.4586	1.23E-06	0.003627475
24	HSPE1	7.1976	0.002323169	0.064361587
25	HIST1H3G	6.9808	0.001073488	0.052671265
26	CPT1B	6.7547	0.03103765	0.137502892
27	HIST1H2BI	6.6695	0.000459019	0.038061433
28	CXCL11	6.4341	3.48E-05	0.012927228
29	RNU4-1	6.3379	0.043980613	0.156620112
30	WARS	6.2585	8.12E-06	0.007119642
31	SAT1	6.2318	0.000146912	0.023399989
32	DHRSX	6.2035	0.012321809	0.105083179
33	S100A8	5.9660	0.000131764	0.022392449
34	IDO1	5.8322	9.01E-06	0.007119642
35	SNORD61	5.6844	0.02161399	0.123917586
36	APOBEC3G	5.5190	0.005456938	0.087327309
37	PSMB4	5.5175	0.001847851	0.061785449
38	SPECC1L	5.5057	0.031145118	0.137694111

39	S100A9	5.4266	4.22E-05	0.014065098
40	CORO7	5.3699	0.034331614	0.141954793
41	CTSL	5.3103	1.81E-05	0.00869966
42	GBP2	5.2105	1.37E-05	0.007651241
43	HCRP1	5.1754	0.001507048	0.05773467
44	SPDYE6	5.0042	0.036474064	0.145253709
45	FRG1JP	4.9858	0.045729097	0.157930768
46	GYG1	4.9725	9.10E-05	0.01865914
47	CD163	4.9649	1.04E-05	0.007119642
48	TLR8	4.9199	4.90E-05	0.015072634
49	AGAP6	4.8820	0.019330668	0.120589689
50	PRELID3B	4.8173	0.035477918	0.144200738
51	FTL	4.7947	0.003587532	0.07603778
52	SERPINB1	4.7894	0.000780782	0.047067147
53	GZMB	4.6843	4.35E-06	0.007119642
54	ANKRD36BP1	4.6798	0.007724414	0.092070695
55	CCL4	4.6704	1.13E-05	0.007119642
56	PTGES3	4.5630	0.001322991	0.056859941
57	HIST1H2AC	4.5564	0.005652328	0.087506193
58	ATP5F1	4.5153	0.003510479	0.07603778

59	MT2A	4.5027	0.00020827	0.026107298
60	LYZ	4.4466	0.003933849	0.079145331
61	ANKRD22	4.4058	0.001170169	0.053760317
62	ANXA5	4.3839	0.000672738	0.043742579
63	SLC25A52	4.3747	0.036974236	0.146177536
64	UBC	4.3711	0.012011585	0.104714687
65	ADAMDEC1	4.3227	0.000524356	0.039544241
66	SCD	4.2805	1.10E-05	0.007119642
67	ZBED6	4.2675	0.01229404	0.105083179
68	PSTPIP2	4.2620	0.000403487	0.036932771
69	HSP90AA1	4.2509	0.005753436	0.087610681
70	LDHA	4.2420	0.001204186	0.054571342
71	CKLF	4.2401	0.002395472	0.06453272
72	HSP90B1	4.2394	0.010676276	0.101040857
73	ANP32E	4.2187	0.000419867	0.037413247
74	ZCRB1	4.1548	0.004495087	0.082786778
75	RAB1A	4.1052	0.00364067	0.07603778
76	SLC39A8	4.0932	1.46E-05	0.007651241
77	HIST2H4A	4.0659	0.047183176	0.159291067
78	DDIT4	4.0633	9.46E-06	0.007119642

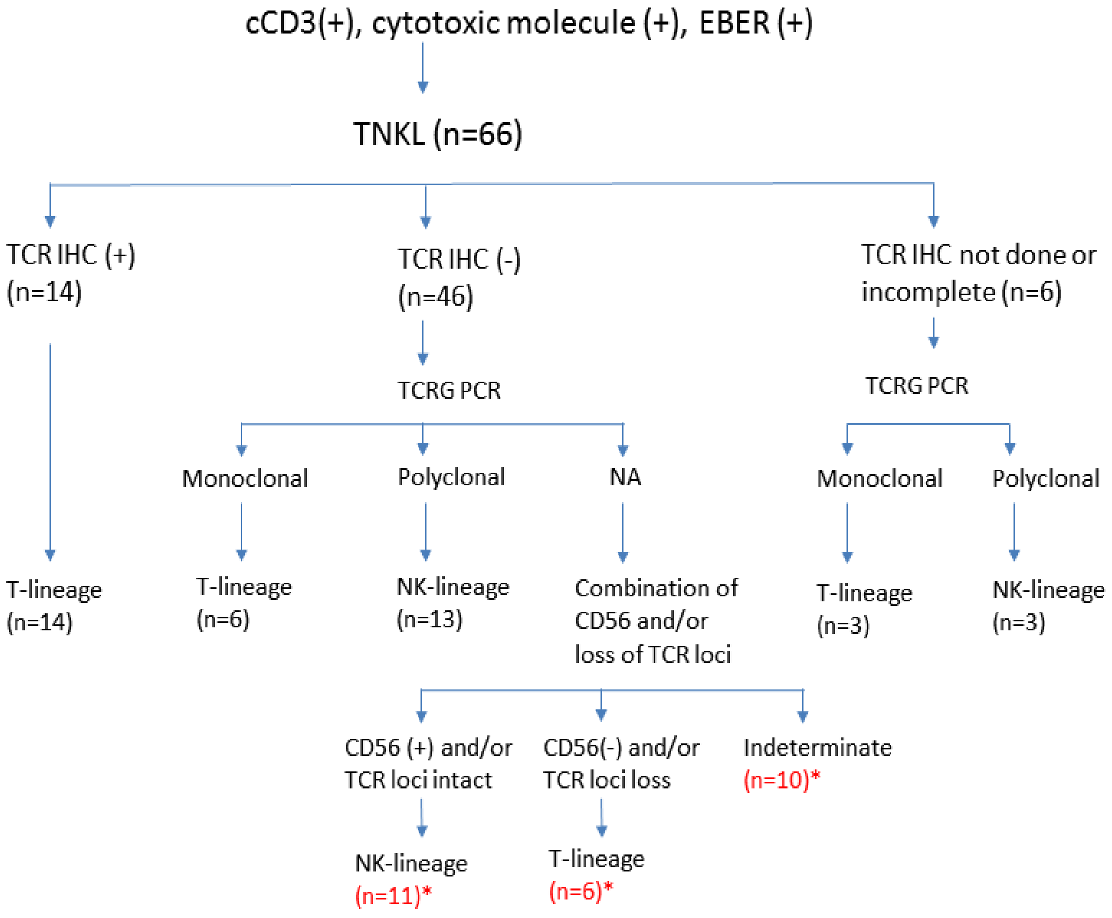
79	SLAMF7	4.0552	0.000344603	0.033850156
80	HSPA8	4.0377	0.005683418	0.087506193

Table S10. Gene set enrichment analysis showing hallmark gene sets enriched in N-group compared to normal lymph nodes (p <0.05 and FDR < 0.2)

NAME	SIZE	Normalized Enrichment Score	p-value	FDR
HALLMARK_G2M_CHECKPOINT	198	1.510794	0.00618392	0.18899
HALLMARK_ALLOGRAFT_REJECTION	200	1.470902	0.01670312	0.115202
HALLMARK_MTORC1_SIGNALING	198	1.450891	0.01779644	0.098907
HALLMARK_FATTY_ACID_METABOLISM	155	1.49483	0.01947286	0.171056
HALLMARK_IL6_JAK_STAT3_SIGNALING	87	1.491728	0.02123515	0.142304
HALLMARK_E2F_TARGETS	198	1.468775	0.02297243	0.10512
HALLMARK_APOPTOSIS	159	1.423059	0.02585867	0.129119
HALLMARK_MYC_TARGETS_V1	194	1.417661	0.03167238	0.119439
HALLMARK_MYC_TARGETS_V2	58	1.488854	0.03328738	0.105651
HALLMARK_COMPLEMENT	194	1.453511	0.03343109	0.114747
HALLMARK_INFLAMMATORY_RESPONSE	197	1.490906	0.04525237	0.120396
HALLMARK_GLYCOLYSIS	198	1.403631	0.04927255	0.132971

Supplemental Figures

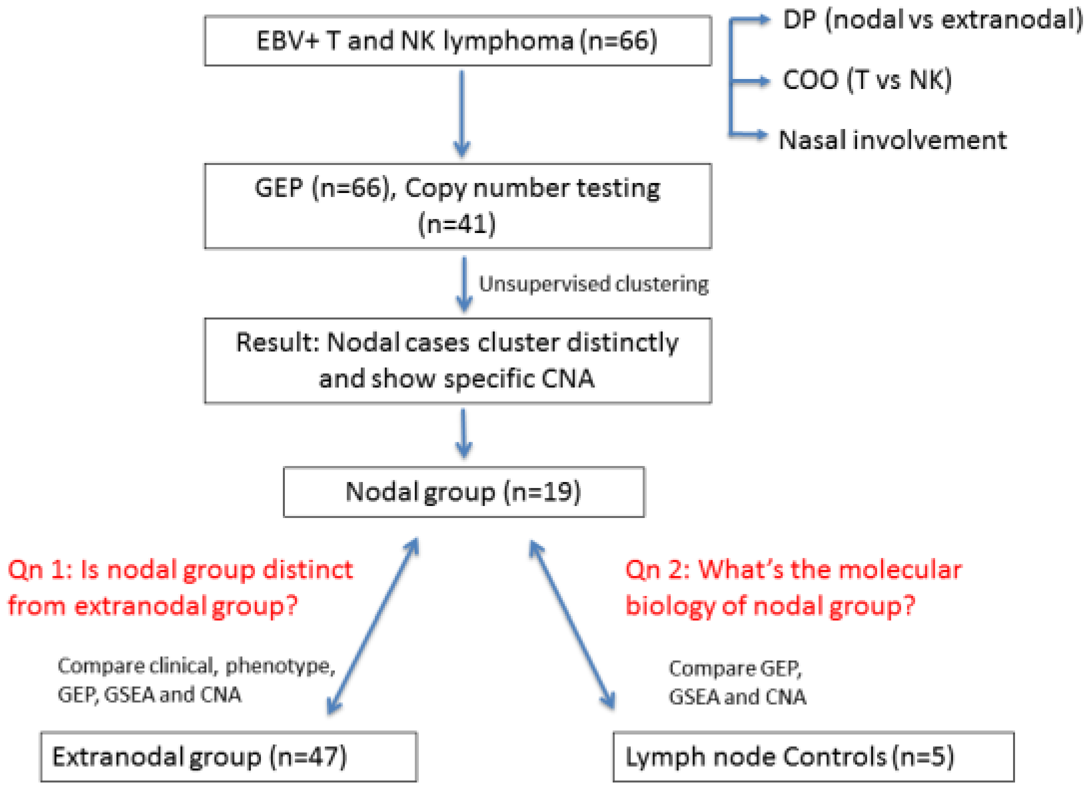
Figure S1



IHC; immunohistochemistry, PCR; polymerase chain reaction, NA; not available (+); positive expression, (-); negative expression, (*) cases denote those using TCR loci data to assign lineage, these cases excluded for TCR loci and lineage analysis

Figure S1. Flow chart illustrating the determination of T vs NK-cell lineage based on the expression of TCRB and TCRG immunohistochemistry and clonality for TCRG gene rearrangement by PCR. Cases with negative expression for TCRB/TCRG IHC and lacking clonality data had lineage assignment based on combination of CD56 expression and TCR loci data.

Figure S2



COO: cell of origin, DP: disease presentation, GEP: gene expression profiling, CNA: copy number aberrations, GSEA: gene set enrichment analysis, Qn: question

Figure S2 illustrating overall study design. It is currently unclear whether the group of EBV-positive T/NK cell lymphoma with nodal presentation (nodal TNKL) represents a distinct entity and the relationship with extranodal NK/T cell lymphoma, nasal type (ENKTL) remains controversial, we undertook this multicenter study in an attempt to better understand the relationship of nodal TNKL with ENKTL and analyzed the molecular biology of nodal TNKL. We collected a total of 66 samples of EBV-associated lymphomas of T and NK cell lineage (TNKL) and classified the cases according to their disease presentation (nodal vs extranodal), nasal involvement and cell of origin (COO). In order to maintain an unbiased approach from the onset, we performed unsupervised clustering of the GEP and copy number changes of all 66 cases and this revealed 3 distinct clusters which were associated with specific copy number changes. This supported the notion that nodal TNKL is distinct from those with extranodal disease. We therefore proceeded to study the clinical, pathologic and molecular signature and copy number profile of nodal cases and compared them to the extranodal group. In addition, we also compared the GEP and CNA of the nodal-group with normal lymph node controls to further understand the molecular biology of this group of tumors since there is currently little known about them.

Figure S3

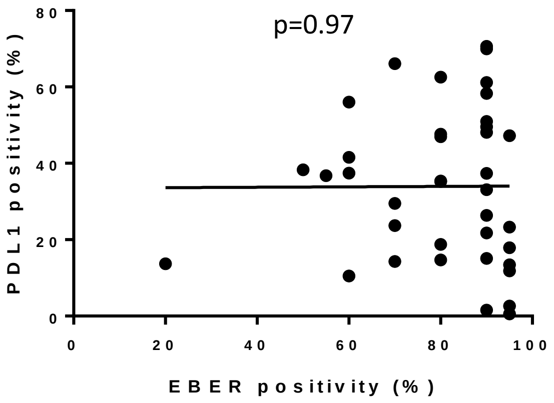


Figure S3. There was no significant correlation of PDL1 expression with EBER expression in tumor cell population.

Figure S4

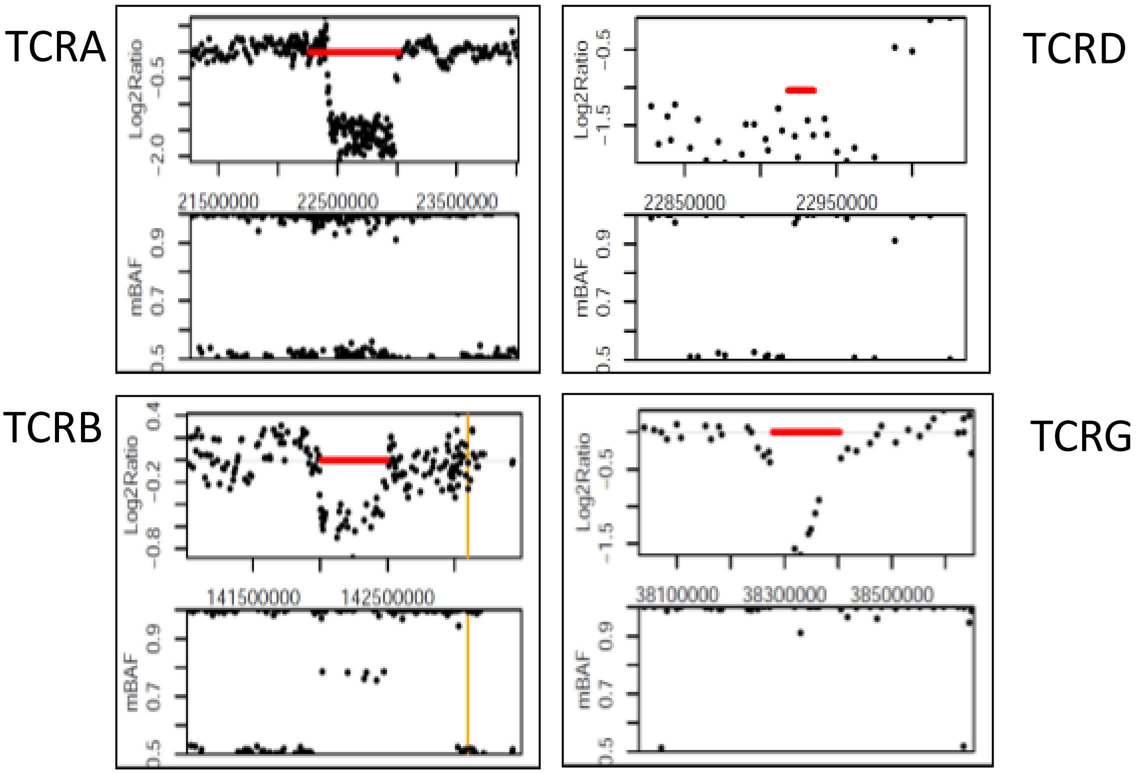


Figure S4. Deletion of TCR loci in TNKL of T-cell origin. Log2 Ratio and Mirrored B Allele Frequency (mBAF) plots of a case of TNKL of T-cell origin (TNK4) demonstrating loss of TCRA, TCRB and TCRD loci. The horizontal red lines indicate the positions of the TCR loci and the orange vertical line demarcates cytoband boundary. Though TCRG locus shown here seems to indicate deletion, the OncoScan Console did not mark it as deleted possibly due to short span.

Figure S5

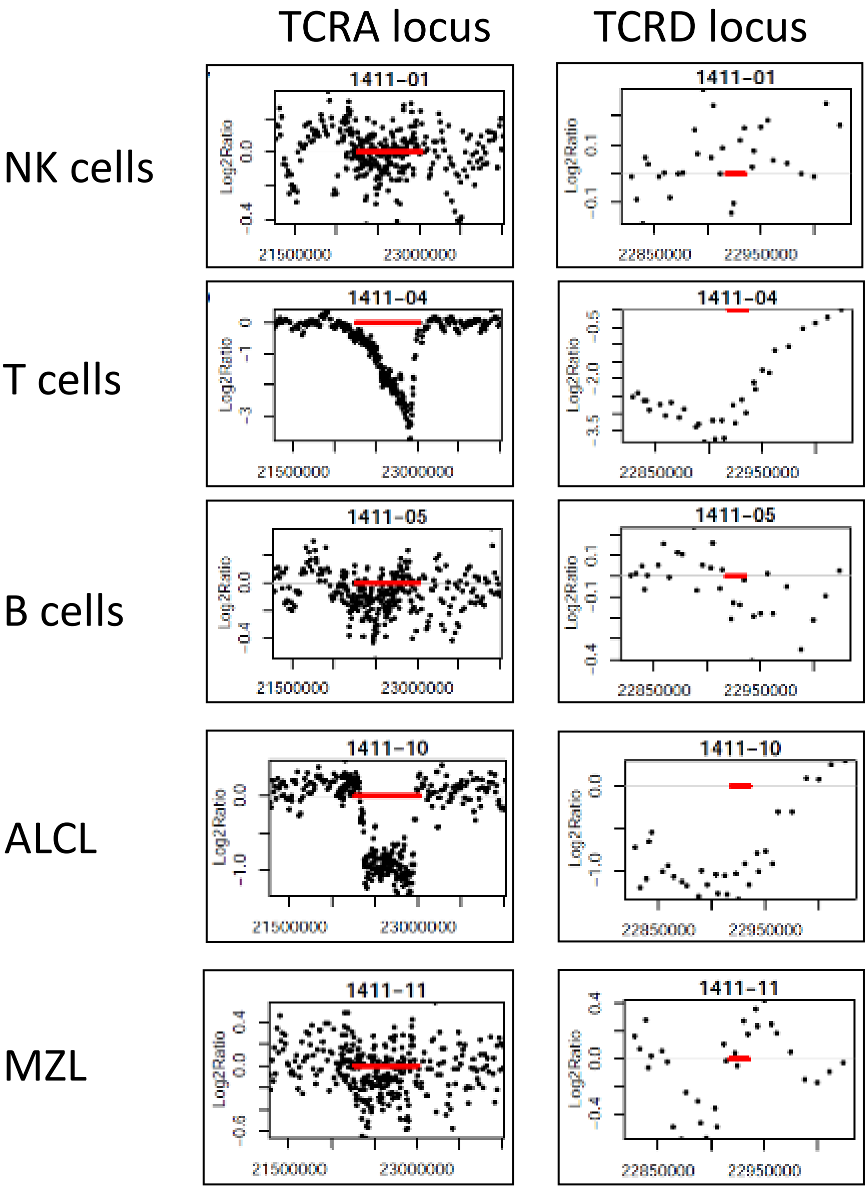


Figure S5. TCRA and TCRD loci in normal NK cells, B cells, T cells, Anaplastic large cell lymphoma (ALCL) and marginal zone lymphoma (MZL) samples. Log2 Ratio plots illustrate loss of TCRA and TCRD loci in normal T cells, and ALCL, but not in normal NK cells, B cells, and MZL.

Figure S6

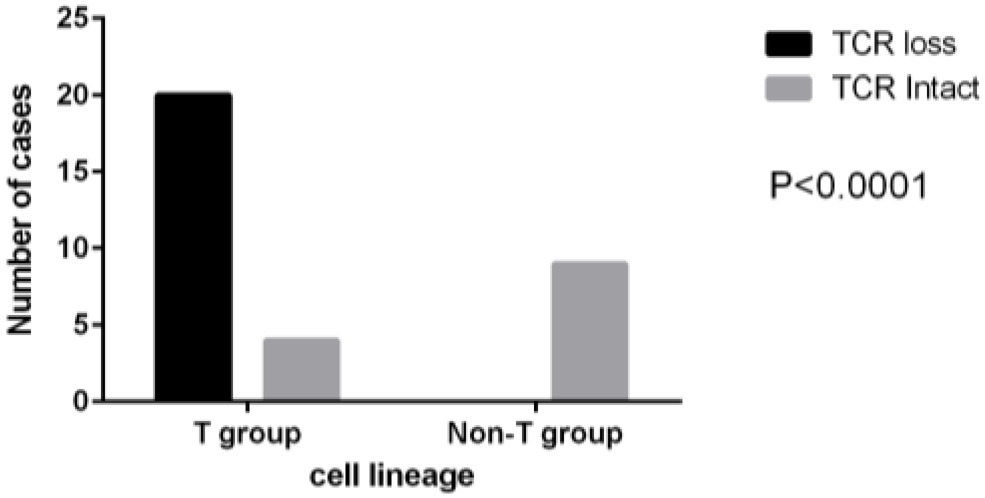


Figure S6. Association between status of TCR loci and cell of origin. T-group samples includes TNKL of T-origin (n=16), PTCL (N=6) and normal T-cells (n=2). Non-T group samples include TNKL of NK origin (n=5), normal B cells (n=2) and normal NK cells (n=2)

Figure S7

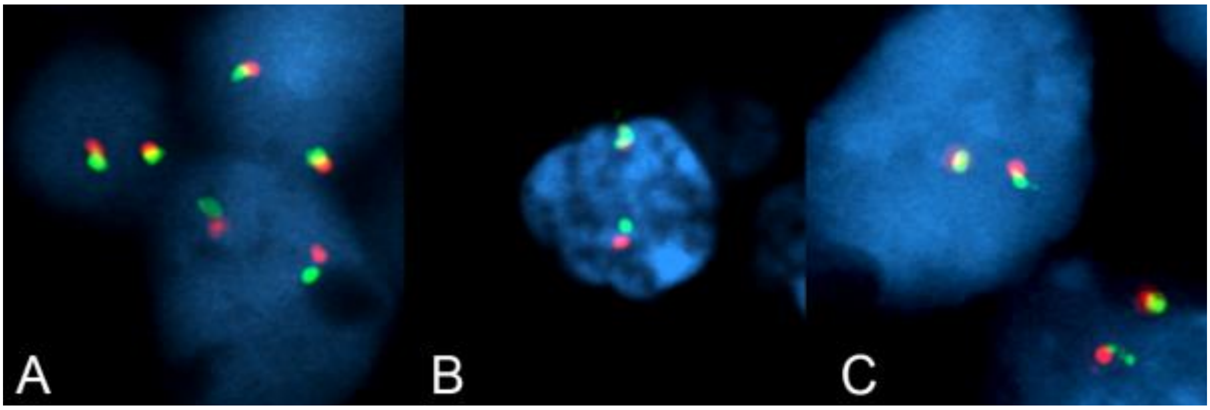


Figure S7. Fluorescence in situ hybridization for *TCRA* locus using breakapart probe. The probe specific for the 5' (centromeric) and 3' (telomeric) regions of the TCRA locus were labeled in P5 Green and Red, respectively. Our results showed an intact *TCRA* locus in normal T cells (A), normal NK cells (B), and T-cell lymphoma (C), indicating an absence of translocation of the *TCRA* locus.

Figure S8

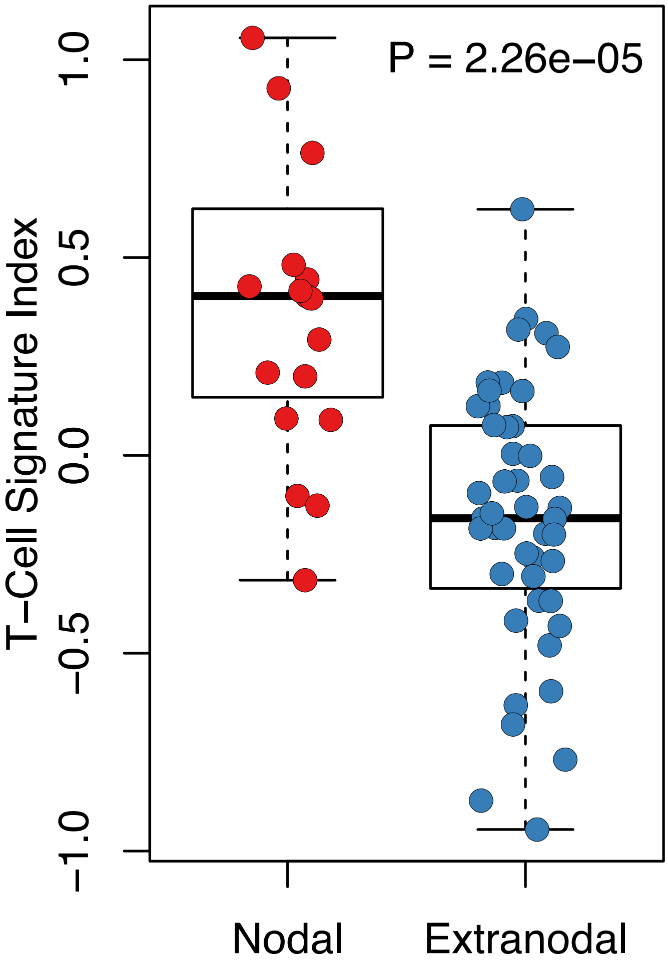


Figure S8. Comparison of T-cell signature between Nodal-group and Extranodal-group. Results showed a significantly higher expression of the T-cell related genes in Nodal-group compared to Extranodal-group.