



EUROPEAN  
HEMATOLOGY  
ASSOCIATION



Ferrata Storti  
Foundation

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

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

**Haematologica** 2018  
Volume 103(2):278-287

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

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

### ABSTRACT

The molecular biology of primary nodal T- and NK-cell lymphoma and its relationship with extranodal NK/T-cell lymphoma, nasal type is poorly understood. In this study, we assessed the relationship between nodal and extranodal Epstein-Barr virus-positive T/NK-cell lymphomas using gene expression profiling and copy number aberration analyses. We performed gene expression profiling and copy number aberration analysis on 66 cases of Epstein-Barr virus-associated T/NK-cell lymphoma from nodal and extranodal sites, and correlated the molecular signatures with clinicopathological features. Three distinct molecular clusters were identified with one enriched for nodal presentation and loss of 14q11.2 (TCRA loci). T/NK-cell lymphomas with a nodal presentation (nodal-group) were significantly associated with older age, lack of nasal involvement, and T-cell lineage compared to those with an extranodal presentation (extranodal-group). On multivariate analysis, nodal presentation was an independent factor associated with short survival. Comparing the molecular signatures of the nodal and extranodal groups it was seen that the former was characterized by upregulation of PD-L1 and T-cell-related genes, including CD2 and CD8, and downregulation of CD56, consistent with the CD8<sup>+</sup>/CD56<sup>-</sup> immunophenotype. PD-L1 and CD2 protein expression levels were validated using multiplexed immunofluorescence. Interestingly, nodal group lymphomas were associated with 14q11.2 loss which correlated with loss of TCR loci and T-cell origin. Overall, our results suggest that T/NK-cell lymphoma with nodal presentation is distinct and deserves to be classified separately from T/NK-cell lymphoma with extranodal presentation. Upregulation of PD-L1 indicates that it may be possible to use anti-PD1 immunotherapy in this distinctive entity. In addition, loss of 14q11.2 may be a potentially useful diagnostic marker of T-cell lineage.

### Correspondence:

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

Received: September 9, 2017.

Accepted: October 27 2017.

Pre-published: November 2, 2017.

doi:10.3324/haematol.2017.180430

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: [www.haematologica.org/content/103/2/278](http://www.haematologica.org/content/103/2/278)

©2018 Ferrata Storti Foundation

Material published in *Haematologica* is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>.

Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>,

sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.



## Introduction

Epstein-Barr virus (EBV)-associated T- and NK-cell lymphoproliferative diseases (TNKL) comprise a heterogeneous group of cytotoxic T/NK lymphoproliferative disorders, including extranodal NK/T-cell lymphoma, nasal type (ENKTL), and a group of cutaneous and systemic diseases in children such as systemic EBV-positive T-cell lymphoma, aggressive NK-cell leukemia, chronic active EBV infection of T/NK subtype, hydroa vacciniforme-like lymphoproliferative disorder, and mosquito bite hypersensitivity.<sup>1</sup> ENKTL is the prototypic example of TNKL in adults.<sup>2</sup> The majority of ENKTL are derived from NK cells but some show a cytotoxic T-cell phenotype. Although the nasal cavity is the most common site of involvement, ENKTL can involve a wide variety of extranodal sites including the skin, gastrointestinal tract and testes. Lymph node involvement is uncommon but may occur as a secondary event.<sup>3</sup>

In addition to ENKTL, there is a group of TNKL that occurs in adults and presents exclusively with lymph node disease but may involve a limited number of extranodal organs (nodal TNKL).<sup>4</sup> These cases show significant overlap with ENKTL but a few reports have described clinicopathological features distinct from ENKTL, including the lack of nasal involvement, frequent T-cell origin, and CD8<sup>+</sup>/CD56<sup>-</sup> phenotype.<sup>5-8</sup> The World Health Organization (WHO) recognizes that most of this group of primary nodal lymphomas are derived from T cells but that a minority have an NK-cell origin and might be different from the prototypic ENKTL, which rarely involve lymph nodes. Since it is currently unclear whether this group of nodal TNKL represents a distinct entity, the revised 2016 WHO lymphoma classification has recommended including primary nodal TNKL as an EBV-positive variant of peripheral T-cell lymphoma, not otherwise specified.<sup>2</sup>

The molecular biology of nodal TNKL is unknown and its relationship with ENKTL remains controversial. To date, mainly due to the rarity of such malignancies, efforts to resolve these controversies have been largely based on clinicopathological assessment of relatively few cases. In this study, we assessed the relationship between nodal and extranodal EBV-positive TNKL using gene expression profiling (GEP) and copy number aberration (CNA) analyses in a large cohort of adult patients. We correlated the molecular signature and copy number profile with lineage and with clinicopathological features in an attempt to understand whether cases with nodal disease at presentation are distinct from their extranodal counterparts. In addition, we also analyzed the GEP and CNA of nodal TNKL compared to control tissues in order to expand our knowledge on the tumor biology of this relatively unknown group of diseases.

## Methods

### Study group

The study group included 66 adult patients with no known immunodeficiency who had been diagnosed with EBV-positive TNKL involving extranodal and nodal sites. All the cases were positive for cCD3, EBV-encoded small RNA (EBER) in the majority of tumor cells (>50%), and at least one cytotoxic marker (TIA1 and/or granzyme B). Systemic and cutaneous EBV-positive T/NK

lymphoproliferative diseases occurring in children were excluded.

Clinical data were obtained (*Online Supplementary Table S1*) and cases were then categorized based on: (i) disease presentation (extranodal *versus* nodal); (ii) nasal involvement (absence *versus* presence); and (iii) cell of origin (T cell *versus* NK cell). Details of lineage assignment are provided in *Online Supplementary Figure S1* and the *Online Supplementary Methods*. This study was approved by NHG Domain Specific Review Board B (2009/212).

### Gene expression profiling

GEP was performed on 66 cases using the Ovation formalin-fixed paraffin-embedded (FFPE) WTA System (NuGEN, San Carlos, CA, USA) and labeled products were hybridized onto Affymetrix Human Gene 1.0 ST arrays (Affymetrix, ThermoFisher Scientific, Waltham, MA, USA).

### Copy number analysis

The OncoScan<sup>®</sup> FFPE assay (Affymetrix) was performed on 41 TNKL cases and 15 control tissues according to the manufacturer's instructions as previously described.<sup>9,10</sup> Copy number analysis was performed using OncoScan<sup>®</sup> Console (v1.3) software (ThermoFisher Scientific, Waltham, MA, USA).

### Multiplex immunofluorescence and multispectral imaging

Multiplex immunofluorescent staining was performed to assess CD3/PD-L1 and CD3/CD2 expression on 3 μm FFPE tissue sections using the Opal 7-color Fluorophore TSA plus Fluorescence Kit (NEL 797001KT, PerkinElmer Inc., Waltham, MA, USA). Images were acquired and analyzed with the Vectra 2 multispectral automated imaging system (PerkinElmer Inc., Waltham, MA, USA) and inForm 2.0 image analysis software (PerkinElmer Inc., Waltham, MA, USA).

### Gene expression profiling and copy number aberration analysis

GEP data were processed with a frozen robust multiarray analysis algorithm, an extension of the renowned robust multiarray analysis algorithm that utilizes probe-specific effects and variances.<sup>11</sup> Unsupervised hierarchical clustering was performed using the top 500 most variable genes (*Online Supplementary Table S2*). To account for the difference in inherent gene-by-gene signal intensity range, we divided signal intensities of a gene by its gross median and then applied the log<sub>2</sub> function before variability assessment.

CNA analysis was performed using segmentation results and probe-by-probe signal intensities from OncoScan<sup>®</sup> Console. Five CNA levels were determined as follows: -2 (CN ≤ 0.5), -1 (0.5 < CN ≤ 1.5), 0 (1.5 < CN < 2.5), +1 (2.5 ≤ CN < 3.5), +2 (3.5 ≤ CN). Cluster and nodal-group specific CNA, as well as CNA of four TCR loci (TCRA, TCRB, TCRG, TCRD) were derived.

The genes differentially expressed between TNKL with a nodal presentation (N-group) and those with an extranodal presentation (EN-group) were identified by using the significance analysis of microarrays algorithm.<sup>12</sup> The T-cell signatures of the two groups were analyzed using published genes CD27, CD3G, CD3D, ICOS, MAL, TCF7, PKIA<sup>13</sup> according to the procedure outlined in a previous publication.<sup>14</sup> Gene set enrichment analysis (GSEA)<sup>15</sup> of the N-group compared to the EN-group and normal lymph nodes was performed using the gseapy package with MSigDB version 5.2 (Broad Institute, Cambridge, MA, USA).

GEP (GSE90597) and CNA (GSE90777, GSE90783) raw data have been deposited in the Gene Expression Omnibus database repository.

Full methods, including the fluorescence *in situ* hybridization for TCRA, are detailed in the *Online Supplementary Methods*. The overall study design is illustrated in *Online Supplementary Figure S2*.

Results

Clinicopathological features of the study group

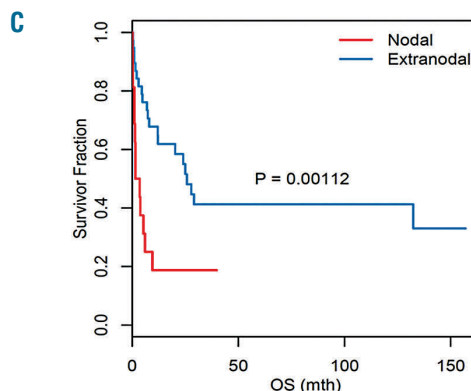
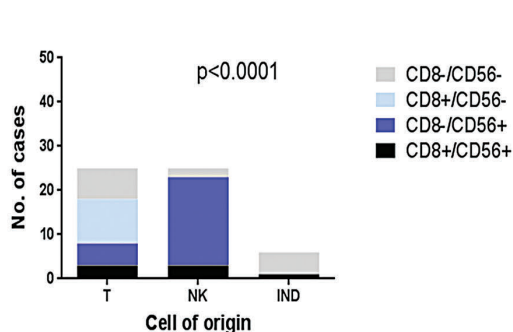
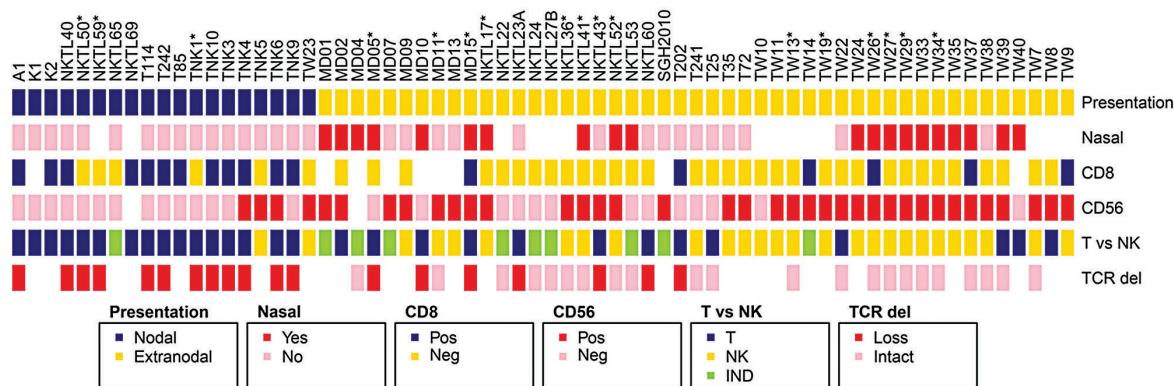
Overall, 47 patients presented with extranodal disease (EN-group) and 19 with nodal disease (N-group) (Table 1, Figure 1A). Twenty-nine cases were classified as having T-cell origin, 27 as NK-cell lineage and ten cases as being of indeterminate lineage (Online Supplementary Figure S4). For 56 cases with sufficient data, the most common CD8/CD56 phenotype was CD8<sup>+</sup>/CD56<sup>-</sup> (25/56, 44.5%), followed by CD8<sup>+</sup>/CD56<sup>+</sup> (14/56, 25%), CD8<sup>-</sup>/CD56<sup>-</sup> (10/56 cases, 18%), and CD8<sup>-</sup>/CD56<sup>+</sup> (7/56, 12.5%) (Online Supplementary Table S3). As expected, all ten cases with the CD8<sup>+</sup>/CD56<sup>-</sup> phenotype were of T-cell origin and the majority (20/25, 80%) of cases with the CD8<sup>+</sup>/CD56<sup>+</sup> phenotype were of NK-cell lineage ( $P=0.0001$ ) (Figure 1B), consistent with CD56 being a marker of NK-cell origin. Interestingly, cases with the CD8<sup>+</sup>/CD56<sup>+</sup> phenotype could be of T- or NK-cell derivation. This indicates that the presence of CD8 positivity, despite the presence of CD56, should prompt the need for clonality assessment to distinguish T-cell from NK-cell origin (Figure 1B). For the EN-group, three out of 45 cases (7%) were positive for TCRB and two out of 43 were positive for TCRG (5%). Thirty-nine of 44 cases (89%) were negative for both TCRB and TCRG. For the N-group, nine out of 17 cases

(53%) were positive for TCRB and none of the 11 cases analyzed was positive for TCRG. Six of 15 cases (40%) were silent for both TCRB and TCRG.

**Table 1.** Comparison of clinicopathological features, gene expression profiling clusters, PD-L1 expression, and loss of TCR loci in the nodal and extranodal group.

	Nodal (n=19)	Extranodal (n=47)	P-value
Age (SD), years	60.95 (15.43)	48.68 (16.65)	0.0071 <sup>(3)</sup>
Sex (female/male)	4 / 15	14 / 33	0.554 <sup>(4)</sup>
Nasal involvement (yes/no)	0 / 17	20 / 15	$3.49 \times 10^{-5}$ <sup>(4)</sup>
COO: T vs. NK vs. Indeterminate	16 / 2 / 1	13 / 25 / 9	0.000116 <sup>(4)</sup>
Stage (1 & 2 / 3 & 4)	2 / 15	23 / 12	0.00031 <sup>(4)</sup>
CD8 (positive/negative)	12 / 6	6 / 33	0.00021 <sup>(4)</sup>
CD56 (positive/negative)	4 / 14	33 / 13	0.00053 <sup>(4)</sup>
Cluster (1 vs. 2 & 3)	9 / 10	10 / 37	0.069 <sup>(4)</sup>
PD-L1 (tumor) <sup>2</sup>	1.12 (0.29)	0.86 (0.34)	$9.1 \times 10^{-7}$ <sup>(3)</sup>
PD-L1 (normal) <sup>2</sup>	1.44 (0.47)	1.15 (0.48)	0.00031 <sup>(3)</sup>
TCR (loss/intact)	12 / 0	7 / 22	$6.38 \times 10^{-6}$ <sup>(4)</sup>
Median survival (months)	2.47	25.7	0.00112 <sup>(5)</sup>

SD: standard deviation; COO: cell of origin; <sup>2</sup>average optical density ratio (standard deviation); <sup>3</sup>t-test; <sup>4</sup>Fisher exact test; <sup>5</sup>log-rank test.



**Figure 1.** Clinicopathological features of the nodal and extranodal groups. (A) Composite map showing clinicopathological features and loss of TCR loci in the N- and EN-groups. The map highlights lack of nasal involvement, CD8<sup>+</sup>/CD56<sup>-</sup> phenotype, T-cell lineage and loss of TCR loci in the N-group compared to the EN-group. (B) Graph illustrating the association of CD8/CD56 phenotype with cell lineage. The CD8<sup>+</sup>/CD56<sup>-</sup> phenotype is associated with NK-cell lineage while the CD8<sup>-</sup>/CD56<sup>+</sup> phenotype is associated with T-cell origin. (C) Survival curve between the N- and EN-groups. Patients in the N-group had significantly shorter overall survival (OS) compared to those in the EN-group.

### Molecular signature and copy number profile of the study group revealed distinct molecular subsets

Unsupervised clustering of GEP performed on 66 TNKL cases revealed three distinct clusters (Figure 2). Cases with upregulation of top variable genes formed a distinct cluster in the middle (cluster 1), which showed enrichment for nodal presentation and nine of 13 cases (69%) tested showed loss of 14q11.2 (Figure 2). The second cluster (cluster 2, left) showed predominant downregulation of genes, loss of 13q14.3-q21.33 (3/7 cases, 43%), and lack of 14q11.2 loss. The third cluster (cluster 3, right) revealed a mixture of up- and downregulated genes. Clusters 2 and 3 showed gain of 1q32.1-q32.3 (8/28 cases, 29%) and loss of Xp22.33 (14/28, 50%), which were absent or uncommon in cluster 1. There was no correlation of the molecular clusters with tissue sites ( $P>0.05$ ), suggesting that the clustering is not related to the signature of the tissue type or tumor microenvironment. Overall, our results showed that TNKL is molecularly heterogeneous with distinct clusters associated with specific CNA. Notably, cases with nodal presentation (N-group) clustered together (Figure 2) and showed frequent loss of 14q11.2, consistent with the proposal that EBV-positive T/NK lymphoma with nodal disease (nodal TNKL) is distinct from those with extranodal disease. We therefore proceeded to compare the clinical, pathological and molecular features of nodal *versus* extranodal TNKL.

### Comparison of the clinicopathological features of the nodal and extranodal groups

Patients in the N-group were older than those in the EN-group (average age, 61 *versus* 48 years;  $P=0.0071$ ) (Table 1). None of the patients had nasal involvement (Figure 1A). Regarding CD8 and CD56 expression, the most common phenotype in the N-group was CD8<sup>+</sup>/CD56<sup>-</sup> (9/17 cases, 52.9%) (Online Supplementary Table S3). The majority (14/16, 87%) were derived from T cells and only two were of NK-cell origin (Figure 1A). Interestingly, patients in the N-group more often had advanced stage disease ( $P=0.0003$ ) (Table 1) and poorer overall survival ( $P=0.00112$ ) (Figure 1C) compared to those in the EN-group. We also performed multivariate analysis to assess the effects of other covariates such as disease presentation, age, cell of origin, nasal involvement and stage of disease on overall survival. Our results demonstrated that only nodal presentation (HR=4.04,  $P=0.016$ ) and stage of disease (HR=2.06,

$P=0.0015$ ) were still significant after adjusting for other factors (Online Supplementary Table S4).

We further compared the above clinicopathological features between T-cell cases in the N-group (N-T) and NK-cell cases in the EN-group (EN-NK) and between T-cell cases in the N-group (N-T) and the EN-group (EN-T) to understand whether the differences between the N-group and the EN-group were simply related to lineage (Table 2A,B). Our results revealed that older age, lack of nasal involvement, CD8 expression and poor outcome remained significantly associated with N-T cases when compared to both EN-NK and EN-T cases, indicating that the N-group is distinct from the EN-group beyond cell of origin.

### Molecular signature and copy number aberrations of the nodal-group compared to extranodal-group

#### Gene expression profiling

In order to determine whether the molecular features of cases with nodal presentation are distinct, we compared the molecular signatures and copy number changes between the N-group and the EN-group. From the list of differentially expressed genes (Online Supplementary Table S5), we observed upregulation of several T-cell-related genes in the N-group, such as CD2, CD8, CD3G, CD3D, TRAC and LEF1 (Figure 3, Online Supplementary Table S5B), and downregulation of CD56 (NCAM1) in the N-group compared to the EN-group (Online Supplementary Table S5A), consistent with the CD8<sup>+</sup>/CD56<sup>-</sup> phenotype of the N-group cases. In addition, PD-L1 (CD274) is upregulated in the N-group compared to the EN-group (Figure 3, Online Supplementary Table S5B). PD-L1 has been reported to be overexpressed in EBV-associated lymphomas, including ENKTL and patients with relapsed ENKTL have shown effective responses to anti-PD1 therapy.<sup>16-18</sup> Hence, the upregulation of PD-L1 in the N-group may have potential therapeutic implications for anti-PD1 treatment.

To confirm our data on differentially expressed genes, we selected CD2, a pan-T-cell marker, and PD-L1 for validation using multiplex immunofluorescence. Consistent with the GEP results, the expression of PD-L1 and CD2 was significantly higher in the N-group than in the EN-group (Figure 4). Further analysis revealed higher expression of CD2 in the tumor cells (CD3<sup>+</sup>), but not in non-tumor cells (CD3<sup>-</sup>), of N-group cases (Figure 4), corroborating the T-cell origin of the cases. For PD-L1 expression,

**Table 2A.** Comparison of clinicopathological features, gene expression profiling clusters, and loss of TCR loci in nodal group-T cases and extranodal group-NK cases.

	Nodal-T	Extranodal-NK	P
Age (SD), <sup>a</sup> years	63.2 (15.83)	49.3 (17.6)	0.0131 <sup>b</sup>
Sex (female/male)	4/12	8/17	0.734 <sup>c</sup>
Nasal involvement (no/yes)	14/0	7/11	0.000389 <sup>c</sup>
CD8 (negative/positive)	3/12	20/3	8.55×10 <sup>-5c</sup>
CD56 (negative/positive)	13/2	2/23	7.92×10 <sup>-5c</sup>
TCR loss (negative/positive)	0/12	14/0	1.04×10 <sup>-5c</sup>
Stage (1 & 2/3 & 4)	2/13	14/4	0.000361 <sup>c</sup>
Cluster (cluster 1 / 2 or 3)	9/7	4/21	0.0144 <sup>c</sup>
Median survival (months)	2.47	24	0.0316 <sup>d</sup>

<sup>a</sup>Mean (standard deviation); <sup>b</sup>t-test; <sup>c</sup>Fisher exact test; <sup>d</sup>log-rank test.

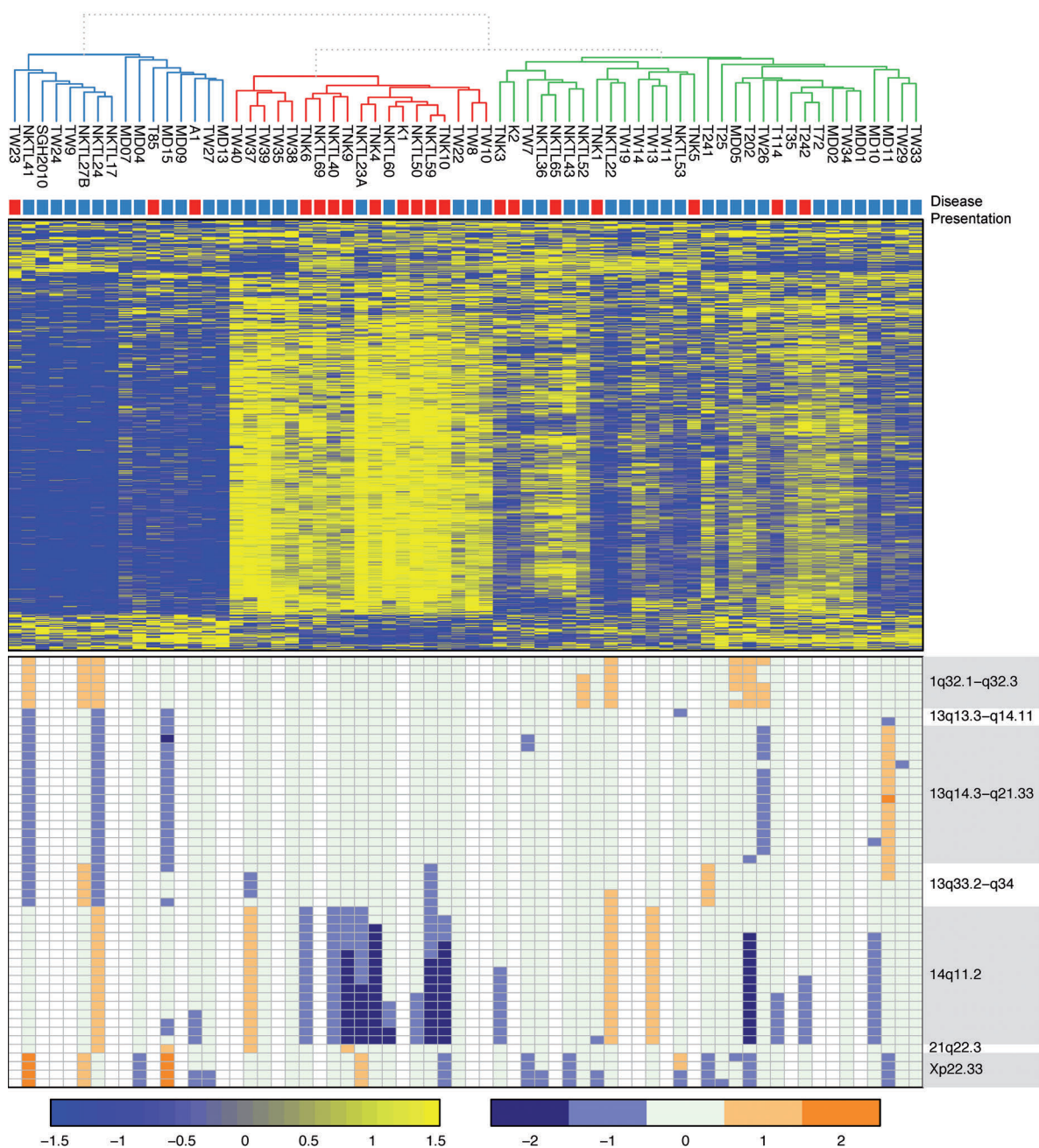
**Table 2B.** Comparison of clinicopathological features, gene expression profiling clusters, and loss of TCR loci in nodal group-T cases and extranodal group-T cases.

	Nodal-T	Extranodal-T	P
Age (SD), <sup>a</sup> years	63.2 (15.83)	50.5 (13.5)	0.0281 <sup>b</sup>
Sex (female/male)	4/12	6/7	0.27 <sup>c</sup>
Nasal involvement (no/yes)	14/0	6/6	0.00401 <sup>c</sup>
CD8 (negative/positive)	3/12	9/2	0.00431 <sup>c</sup>
CD56 (negative/positive)	13/2	7/6	0.0957 <sup>c</sup>
TCR loss (negative/positive)	0/12	3/7	0.0779 <sup>c</sup>
Stage (1 & 2/3 & 4)	2/13	5/7	0.185 <sup>c</sup>
Cluster (cluster 1 / 2 or 3)	9/7	6/7	0.715 <sup>c</sup>
Median survival (months)	2.47	25.7	0.0253 <sup>d</sup>

<sup>a</sup>Mean (standard deviation); <sup>b</sup>t-test; <sup>c</sup>Fisher exact test; <sup>d</sup>log-rank test.

both tumor and non-tumor cells of the N-group cases revealed higher expression compared to their counterparts in the EN-group (Figure 4). Based on morphology, we affirmed that most of the PD-L1<sup>+</sup> non-tumor cells were macrophages indicating a higher number of macrophages in the tumor microenvironment of N-group cases. This is also in line with our gene expression results showing upregulation of CD68, a macrophage marker, in the N-group compared to the EN-group (Figure 3).

We further assessed the 9p24.1/PD-L1/PD-L2 locus for CNA since this was reported to be associated with upregulation of PD-L1 in the majority of cases of classical Hodgkin lymphoma.<sup>19</sup> However, only two samples in the EN-group (NKTL24, NKTL27A) and none in the N-group showed gain of 9p24.1, indicating that the upregulation of PD-L1 in the N-group is likely due to mechanisms other than CNA. In order to understand whether PD-L1 overexpression may be driven by EBV,<sup>16,17</sup> we also correlated the



**Figure 2.** Composite heatmap of gene expression profiling clusters with (top) dendrogram highlighting three clusters in blue, red and green, (upper-middle) disease presentation (nodal in red and extranodal in blue), (lower middle) matrix of the top 500 most highly variable genes, and (bottom) cluster-specific copy number alterations of 66 samples of T/NK-cell lymphoma. Each column represents a case. Genes (rows) in the GEP matrix are ordered based on clustering, but chromosomal segments (rows) in cluster-specific CNA were ordered simply based on their genomic location. Yellow in the GEP matrix indicates upregulation and blue represents downregulation. For the cluster-specific CNA, orange indicates copy number gain and blue represents copy number loss. The GEP heatmap shows three distinct clusters. Cluster 1 (middle, red) shows enrichment for nodal presentation, upregulation of genes and distinctive loss of 14q11.2. Cluster 2 (left, blue) reveals downregulation of genes and loss of 13q14.3-q21.33. Cluster 3 (right, green) shows a mixture of upregulation and downregulation.

expression of EBER and PD-L1 in the tumor population but did not observe any significant association between the two (*Online Supplementary Figure S3*). This suggests that the upregulation of PD-L1 in this group of EBV-positive T/NK lymphomas may either be a consequence of indirect mechanisms driven by EBV or attributed to other novel pathways which will require additional functional studies for further validation.

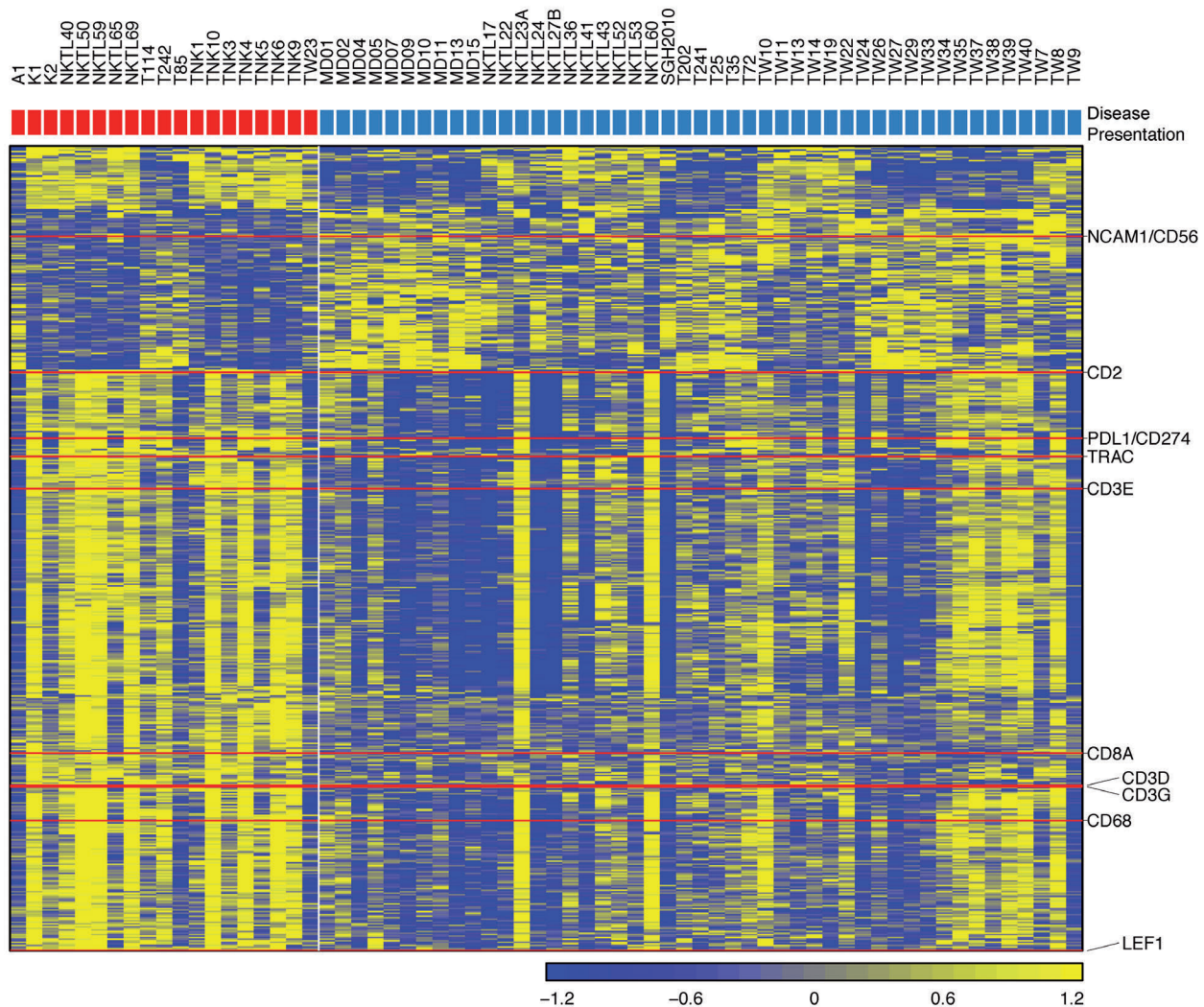
We performed GSEA and compared the N-group to the EN-group, as well as N-T cases to both EN-NK and EN-T cases. However, we did not identify any significant signaling pathways which are distinctly enriched in either group. This may be related to small sample size, and/or measurement variability from heterogeneity in sample processing and preparation.

**Copy number analysis**

The copy number profile revealed similarities and differences between the two groups (*Figure 5*). The recurrent copy number losses present in both groups included loss of 3q26.1 and 22q11.23 (*Online Supplementary Table S6*). Interestingly, loss of 14q11.2 was present in the majority

of N-group cases but identified only in a minority of EN-group cases ( $P < 0.05$ ). In contrast, the EN-group was characterized by recurrent loss of 6q22.3-q26 and 9p21.3 as well as gains in 1q22-q44, 2q22.2-q33.1, 6p21.3, 7q21.1-q36.3, 11q24.3, 13q14.2, and 17q21.2-q25.3 (*Online Supplementary Table S6B,C*). In support of the validity of our copy number results, the recurrent CNA identified in the EN-group in this study have also been reported to be frequently observed in ENKTL in previous studies.<sup>20,21</sup>

Since the 14q11.2 locus contains the T-cell receptor alpha constant (TRAC or TCRA locus, *Online Supplementary Table S6A*), we hypothesized that the loss of 14q11.2 in the N-group might be an indication of T-cell lineage as focal loss within the TCR loci (commonly bi-allelic) can be a reflection of physiological rearrangement of the TCR loci via VDJ recombination. We correlated the loss of TCR loci (TCRA, TCRB, TCRG, TCRD) in our samples with cell of origin. Only cases with lineage assigned based on complete TCRB/TCRG immunohistochemical and clonality data were used for this analysis (cases denoted with an asterisk in *Online Supplementary Table S1* were excluded). Indeed 13 out of 16 (81%) TNKL

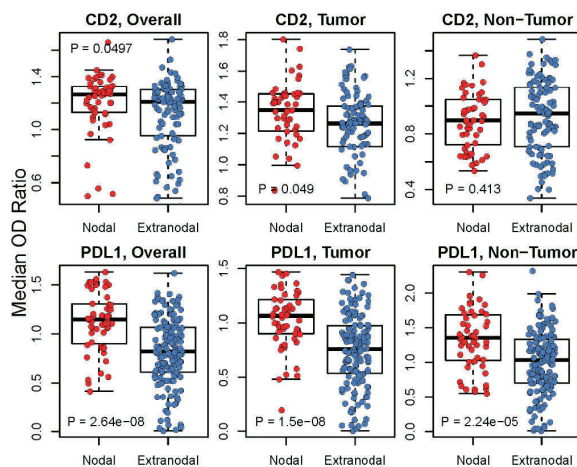


**Figure 3.** Heatmap of differentially expressed genes between nodal- and extranodal-groups (fold change > 1.5,  $P < 0.05$ , false discovery rate < 0.15). Selected genes related to T cells are highlighted. Upregulation and downregulation in the heatmap are marked by yellow and blue, respectively.

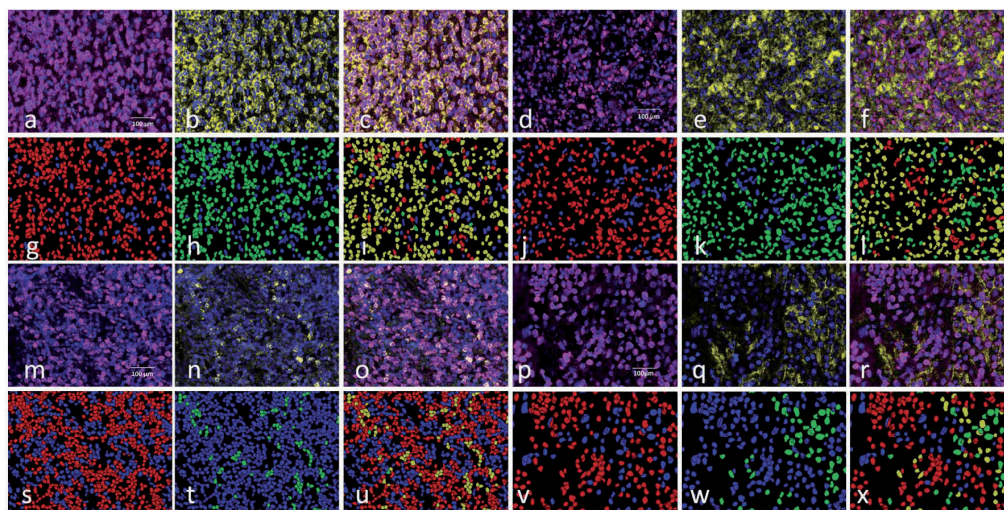
of T-cell origin (as detected by clonality and TCRB/TCRG immunohistochemistry) showed deletion of at least two TCR loci (Figure 1A, *Online Supplementary Figure S4*, *Online Supplementary Table S7*). TCRA and TCRD were the most commonly deleted loci, possibly due to positional proximity. None of the five TNKL of NK-cell lineage displayed loss of TCR loci (Figure 1A). To verify the specificity of 14q11.2 loss as a marker for T-cell lineage, we performed OncoScan profiling of additional T-cell lymphomas, benign lymphoid tissues, and normal lymphocytes including T cells, B cells, and NK cells (*Online Supplementary*

*Methods*). The loss of one or more TCR loci was detected in five out of six (83%) peripheral T-cell lymphomas and two normal T-cell samples, but not observed in benign lymphoid tissues, normal B cells, or NK cells (*Online Supplementary Figure S5*; *Online Supplementary Table S7*). Taken together, 20 of 24 samples of T-cell origin showed loss of TCR loci but not any of the nine non-T-cell samples ( $P < 0.0001$ ) (*Online Supplementary Table S8*, *Online Supplementary Figure S6*). Our findings suggest that loss of TCR loci, especially the 14q11.2 region, detected via the OncoScan platform may be an indicator of T-cell origin.

A



B



**Figure 4. Protein expression of CD2 and PD-L1 in the nodal and extranodal groups.** (A). Plots of CD2 and PD-L1 expression in all cells, tumor and non-tumor cells. Each dot represents the median optical density (OD) ratio of a specific marker in each image analyzed. At least four images containing at least 10,000 cells were quantified per case. The expression of CD2 and PD-L1 across all cell types (both tumor and non-tumor cells) in the cases is illustrated in the left plots indicated 'Overall', while that for tumor and non-tumor cells is represented in the middle and right plots, respectively. Our results showed that CD2 expression is significantly upregulated in tumor cells of the N-group compared to the EN-group. Interestingly, PD-L1 expression is significantly higher in both tumor and non-tumor cells in the N-group than in the EN-group. (B). Expression of CD2 and PD-L1 in the N-group and the EN-group cases using multiplexed immunofluorescence (MIF, a-f and m-r) and corresponding multispectral analysis (g-l and s-x). CD3 stained cell membrane of tumor cells (magenta) in single unmixed images (a, d, m, and p). Single CD2 (b and n, cell membrane stain) and PD-L1 (e and q, cell membrane) expression are stained yellow. Images a-l represent a case in the N-group (TNK4) showing high CD2 and PD-L1 expression. CD3<sup>+</sup> tumor cells accounted for 82.6% of cells (a and g, positive cells marked magenta and red, respectively) and 67.75% of cells (d and j, positive cells marked magenta and green, respectively). CD2 was present in 81.84% of cells (b and h, positive cells marked yellow and green, respectively). Composite CD3 and CD2 image (c) illustrated 75.9% of double positive cells (l, yellow cells), which represented tumor cells with positive CD2 expression. This case also showed high PD-L1 expression in 87.0% of cells (e and k, positive cells marked yellow and green, respectively). Composite CD3 and PD-L1 image (f) revealed 55.9% of tumor cells showing PD-L1 expression (l, yellow cells). Images m-o, s-u represent a case in the EN-group (TW9) with comparatively low expression of CD2. CD3<sup>+</sup> tumor cells accounted for 66.22% of cells (m and s, positive cells marked magenta and red, respectively). Single CD2 expression quantified as 6.75% of cells (n and t, positive cells marked yellow and green, respectively) and the composite image (o) showed 6.51% of CD3<sup>+</sup>/CD2<sup>+</sup> double positive cells (u, double positive cells marked yellow). Images p-r, v-x represent a case in the EN-group (NKTL43) with relatively low expression for PD-L1. CD3<sup>+</sup> tumor cells accounted for 67.12% of cells (p and v, positive cells marked magenta and red, respectively). PD-L1 expression is low in 17.53% of cells (q and w, positive cells marked yellow and green, respectively). Composite CD3 and PD-L1 image (r) revealed 5.55% of CD3<sup>+</sup>/PD-L1<sup>+</sup> cells (x, yellow cells).

To understand the mechanism of loss of the TCRA locus in 14q11.2 better, we performed fluorescence *in situ* hybridization using breakapart probes labeling the TCRA locus to look for translocation of TCRA. Our results showed an intact TCRA locus, indicating that the 14q11.2 loss was not a result of translocation of the TCRA gene and, therefore, more likely a result of the V(D)J recombination that occurs during early stages of T-lymphocyte maturation when segments between the VDJ gene loci are deleted (*Online Supplementary Figure S7; Online Supplementary Methods*).

The N-group also demonstrated significantly higher expression of a previously published T-cell signature<sup>13</sup> compared to the EN-group, reinforcing our hypothesis that loss of TCR loci in the former is associated with T-cell lineage ( $P < 0.0001$ , *Online Supplementary Figure S8; Online Supplementary Methods*).

### Molecular signature and copy number aberrations of the nodal group compared to normal controls

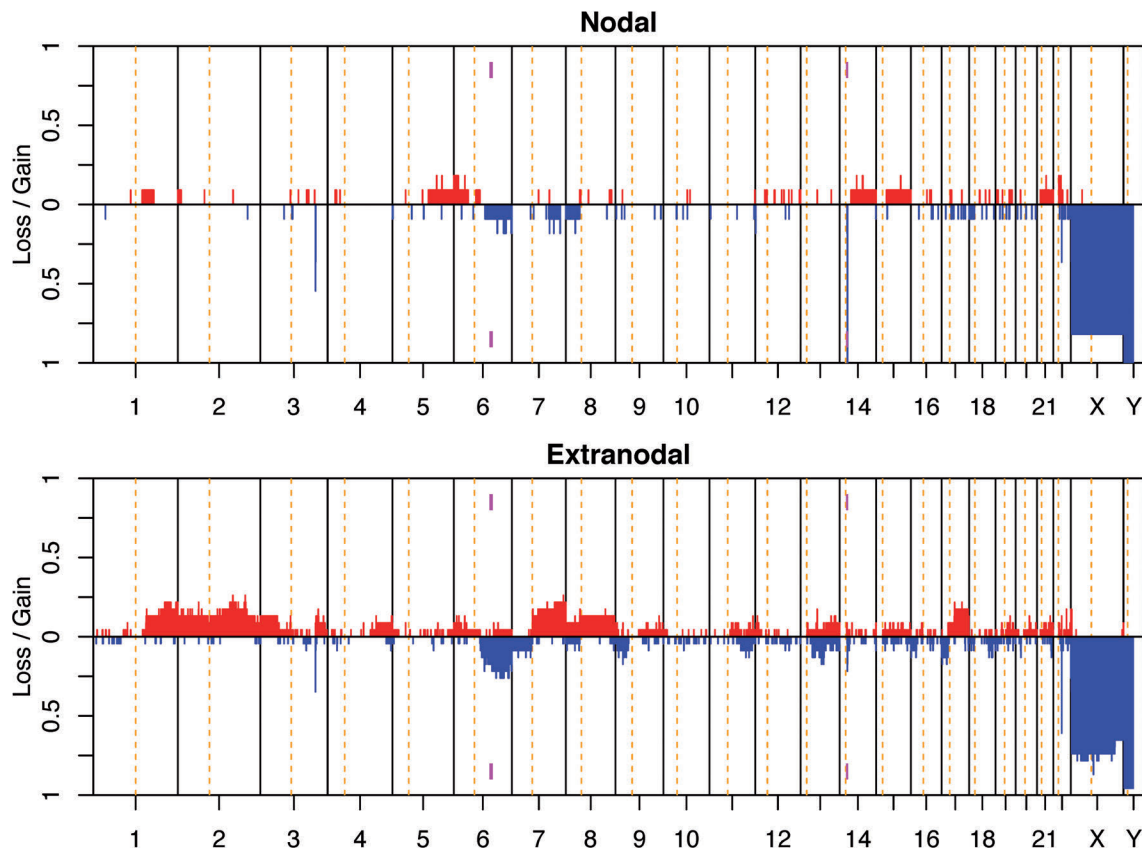
Since very little is known about the molecular biology of nodal TNKL as a disease entity, we generated the differentially expressed genes (*Online Supplementary Table S9*) of the 19 N-group cases compared to normal lymph node controls and performed GSEA. Our results revealed significant enrichment for hallmark gene sets such as MTORC1\_SIGNALING, IL6\_JAK\_STAT3\_SIGNALING as well as several gene sets related to the cell cycle and genomic instability,

including G2M\_CHECKPOINT, E2F\_TARGETS, MYC\_TARGETS, and APOPTOSIS (*Online Supplementary Table S10*).

Twelve out of 19 N-group cases were analyzed for CNA. The recurrent CNA occurring in at least 20% of samples include loss of chr14q11.2 (12/12, 100%), chr3q26.1 (8/12, 67%), and chr22q11.23 (4/12, 33%) (*Online Supplementary Table S6A*). Loss of chr14q11.2 was the most frequent CNA, consistent with this aberration as a marker of T-cell lineage. The genes residing in chr22q11.23 include LOC768328, GSTTP2, LOC391322, LOC100420769, GSTTP1 and GSTT1. No known genes reside in the chr3q26.1 locus. There were no recurrent gains observed in this group.

### Discussion

This is the first large-scale, high-throughput genomics study to interrogate the molecular heterogeneity of TNKL using an unbiased and unsupervised approach. As expected, TNKL is not a uniform entity and molecular subsets exist. Interestingly, we identified one molecular subtype to be associated with distinct CNA, including the loss of 14q11.2, and this group was enriched for nodal presentation and T-cell lineage. In line with recent reports, we verified that nodal TNKL is associated with older age, lack of nasal involvement and the CD8<sup>+</sup>/CD56<sup>-</sup> phenotype, in contrast to ENKTL. We also discovered that patients with



**Figure 5. Penetrance plots showing the frequency of gains and losses of genomic regions of cases in the nodal and extranodal groups.** Each chromosome is represented on the x-axis, and the y-axis indicates the proportion of gain or loss of the corresponding genomic region within the corresponding population. Gains are shown in red and losses in blue.



nodal TNKL often presented with an advanced stage of disease and had a much poorer survival than those with ENKTL. These differences remained significant even when we compared only the T-cell cases in both groups, indicating that there are underlying biological differences, other than cell of origin, between the two groups (Tables 1 and 2). GSEA failed to identify distinct signaling pathways in the two groups although this may be due to limited sample size and/or measurement variability. Alternatively, there may be mutational and epigenetic differences between the two groups which were not investigated in this study.

The diagnosis and classification of TNKL is difficult because of considerable heterogeneity and overlap in morphology, phenotype, cell lineage, and clinical presentation due to involvement of multiple anatomic sites.<sup>22</sup> Our data revealed that TNKL presenting with nodal disease is characterized by unique clinical and phenotypic features as well as molecular and copy number signatures, thereby suggesting that it has a distinct molecular biology and deserves a separate classification from ENKTL. Our results also highlight the importance of anatomic location in defining the disease and that localization of the disease to nodal sites at presentation can be a valuable clue to distinguish between nodal TNKL and ENKTL.<sup>22</sup> This distinction is especially important considering the significantly different clinical outcome between the two entities and hence a pressing need to treat nodal TNKL with more effective therapy. However, whether nodal TNKL represents a distinctive entity or should be classified under the broad category of peripheral T-cell lymphoma, not otherwise specified, as recommended by the 2016 revised WHO classification of lymphoid neoplasms, requires further study.<sup>1,2,4</sup> Interestingly, two cases in the N-group (TW23 and TNK5) were of NK-cell lineage with the CD8<sup>+</sup>/CD56<sup>+</sup> phenotype and both clustered with the EN-group rather than with the N-group (Figure 2) despite their nodal presentation. Whether such cases are biologically akin to EN-group cases remains uncertain and warrants additional study with more numerous samples.

The molecular biology of nodal TNKL is poorly understood and, to date, there is only one report describing the molecular signature of three cases of nodal TNKL.<sup>25</sup> The study showed significant enrichment of immune response genes with upregulation of genes associated with cytotoxic activation and downregulation of genes associated with T- and B-cell activation. Pathways generally known to be involved in tumorigenesis, such as apoptosis, proliferation, or cell adhesion, were not significantly enriched. In this study, 19 cases of nodal TNKL were profiled for gene expression and 12 were subjected to copy number testing. Our study demonstrated enrichment of gene sets related

to the JAK/STAT3 signaling pathway, similar to ENKTL in which activating mutations of JAK3 and phosphorylation of STAT3 result in activation of this oncogenic pathway. Interestingly, gene sets related to the cell cycle and genomic instability were also enriched in nodal TNKL. Genomic instability is an evolving hallmark of cancer and it can be a result of many different pathways.<sup>24</sup> Understanding the mechanisms leading to genomic instability will lay the foundation for novel therapeutic approaches that exploit the selective vulnerability of cancers conferred by their unstable genomes.<sup>25,26</sup> Interestingly, we have verified that loss of 14q11.2 correlated well with loss of TCR loci and T-cell lineage. Considering the importance of accurate diagnosis of the cell of origin in lymphoid malignancies, it is worth noting that detection of aberrations in TCR loci using the OncoScan platform, which works reliably on FFPE materials, may be a novel and potentially useful diagnostic tool for determining T-cell lineage.

Successful treatment of patients with advanced cancer using antibodies against PD-1 and its ligand (PD-L1) has highlighted the critical importance of PD-1/PD-L1-mediated immune escape in cancer development. Targeting the PD-1/PD-L1 pathway has shown clinical efficacy not only in solid tumors but also in Hodgkin and non-Hodgkin lymphoma.<sup>27</sup> Overexpression of PD-L1 and/or PD-L2 has been shown to play a critical role in immune evasion by lymphoma cells.<sup>19,28</sup> Overexpression of PD-L1 has been reported in ENKTL<sup>16,17</sup> and anti-PD1 immunotherapy has been shown to be effective in patients with relapsed and refractory ENKTL.<sup>18</sup> Our findings have shown that nodal TNKL is a much more aggressive disease than ENKTL with a median survival of less than 3 months, highlighting a pressing need for more effective treatment modalities. Importantly, the upregulation of PD-L1 in nodal TNKL elucidated in this study may have potential therapeutic implications and definitely warrants further study on the possibility of PD-L1 immunotherapy in this group of tumors. The mechanism of PD-L1 upregulation in TNKL remains unclear as it did not correlate with EBV expression nor was it a result of gain in the 9p24.1/PD-L1/PD-L2 locus. Finally, we observed high PD-L1 expression on the tumor cells as well as tumor-infiltrating macrophages in the N-group cases. The mechanisms responsible for the recruitment of macrophages to tumors and the control of PD-L1 upregulation on these cells also require further evaluation.

#### Acknowledgments

*S.-B.N was supported by the Singapore Ministry of Health's National Medical Research Council Transition Award (NMRC/TA/0020/2013) and Translational & Clinical Research (TCR) Flagship Program (TCR12dec005).*

#### References

- Attygalle AD, Cabecadas J, Gaulard P, et al. Peripheral T-cell and NK-cell lymphomas and their mimics; taking a step forward - report on the lymphoma workshop of the XVth meeting of the European Association for Haematopathology and the Society for Hematopathology. *Histopathology*. 2014;64(2):171-199.
- Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
- 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.
- 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.
- 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.
6. Kato S, Takahashi E, Asano N, et al. Nodal cytotoxic molecule (CM)-positive Epstein-Barr virus (EBV)-associated peripheral T cell lymphoma (PTCL): a clinicopathological study of 26 cases. *Histopathology*. 2012;61(2):186-199.
  7. 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.
  8. 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.
  9. 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.
  10. Harms PW, Fullen DR, Patel RM, et al. Cutaneous basal cell carcinomas: evidence of clonality and recurrent chromosomal losses. *Hum Pathol*. 2015;46(5):690-697.
  11. McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). *Biostatistics*. 2010;11(2):242-253.
  12. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA*. 2001;98(9):5116-5121.
  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. 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 USA*. 2005;102(43):15545-15550.
  16. Han L, Liu F, Li R, et al. Role of programmed death ligands in effective T-cell interactions in extranodal natural killer/T-cell lymphoma. *Oncol Lett*. 2014;8(4):1461-1469.
  17. Chen BJ, Chapuy B, Ouyang J, et al. PD-L1 expression is characteristic of a subset of aggressive B-cell lymphomas and virus-associated malignancies. *Clin Cancer Res*. 2013;19(13):3462-3473.
  18. Kwong YL, Chan TS, Tan D, et al. PD1 blockade with pembrolizumab is highly effective in relapsed or refractory NK/T-cell lymphoma failing L-asparaginase. *Blood*. 2017;129(17):2437-2442.
  19. Roemer MG, Advani RH, Ligon AH, et al. PD-L1 and PD-L2 genetic alterations define classical Hodgkin lymphoma and predict outcome. *J Clin Oncol*. 2016;34(23):2690-2697.
  20. Iqbal J, Kucuk C, Deleeuw RJ, et al. Genomic analyses reveal global functional alterations that promote tumor growth and novel tumor suppressor genes in natural killer-cell malignancies. *Leukemia*. 2009;23(6):1139-1151.
  21. Huang Y, de Reynies A, de Leval L, et al. Gene expression profiling identifies emerging oncogenic pathways operating in extranodal NK/T-cell lymphoma, nasal type. *Blood*. 2010;115(6):1226-1237.
  22. Swerdlow SH, Jaffe ES, Brousset P, et al. Cytotoxic T-cell and NK-cell lymphomas: current questions and controversies. *Am J Surg Pathol*. 2014;38(10):e60-71.
  23. Ha SY, Sung J, Ju H, et al. Epstein-Barr virus-positive nodal peripheral T cell lymphomas: clinicopathologic and gene expression profiling study. *Pathol Res Pract*. 2013;209(7):448-454.
  24. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol*. 2010;11(3):220-228.
  25. Ferguson LR, Chen H, Collins AR, et al. Genomic instability in human cancer: molecular insights and opportunities for therapeutic attack and prevention through diet and nutrition. *Semin Cancer Biol*. 2015;35(Suppl):S5-24.
  26. Pawitan Y, Michiels S, Koscielny S, Gusnanto A, Ploner A. False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics*. 2005;21(13):3017-3024.
  27. Bachy E, Coiffier B. Anti-PD1 antibody: a new approach to treatment of lymphomas. *Lancet Oncol*. 2014;15(1):7-8.
  28. Green MR, Monti S, Rodig SJ, et al. Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood*. 2010;116(17):3268-3277.