haematologica
Journal of the Ferrata Storti Foundation

Enteropathy-associated T-cell lymphoma: molecular clonal evolution and its clinical consequences

by Bauke Ylstra, Patricia J.T.A. Groenen, Tjitske Los-de Vries, Jurriaan Janssen, Maxine Rouvroye, Jeroen Maertzdorf, Jacqueline Egthuijsen, Liping Fu, Jeroen Luijks, Hetty J Bontkes, Yongsoo Kim, Gerd Bouma and Daphne de Jong

Received: February 27, 2026.
Accepted: July 3, 2026.

Citation: Bauke Ylstra, Patricia J.T.A. Groenen, Tjitske Los-de Vries, Jurriaan Janssen, Maxine Rouvroye, Jeroen Maertzdorf, Jacqueline Egthuijsen, Liping Fu, Jeroen Luijks, Hetty J Bontkes, Yongsoo Kim, Gerd Bouma and Daphne de Jong. Enteropathy-associated T-cell lymphoma: molecular clonal evolution and its clinical consequences.
Haematologica. 2026 July 16. doi: 10.3324/haematol.2026.300770 [Epub ahead of print]

Enteropathy-associated T-cell lymphoma: molecular clonal evolution and its clinical consequences

Bauke Ylstra[1*#], Patricia J.T.A. Groenen[2*], Tjitske Los-de Vries[1], Jurriaan Janssen[1], Maxine Rouvroye[3], Jeroen Maertzdorf[1], Jacqueline Egthuijsen[1], Liping Fu[1], Jeroen Luijks[2], Hetty J. Bontkes[4], Yongsoo Kim[1], Gerd Bouma[3**], Daphne de Jong[5**]

1 Amsterdam University Medical Center, Dept of Pathology, Amsterdam, the Netherlands
2 Radboud University Medical Center, Dept of Pathology/ Laboratory for Tumor Genetics, Nijmegen, the Netherlands
3 Amsterdam University Medical Center, Dept of Gastroenterology, Amsterdam, the Netherlands
4 Amsterdam University Medical Center, Department of Laboratory Medicine, Laboratory Specialized Diagnostics & Research, Section Medical Immunology, Amsterdam, the Netherlands
5 The Netherlands Cancer Institute, Dept. of Pathology, Amsterdam, the Netherlands

* shared first authors
** shared last authors
# corresponding author

Running title: Oncogenetic evolution from RCD to EATL

Author contributions:
Study design: BY, GB, DJ; data collection: JL, MR, JM, JE, JL; data interpretation: BY, PJTAG, TLV, JJ, YK, DJ; manuscript writing: BY, PJTAG, TLV, JJ, GB, DDJ; manuscript approval: all authors

Key words: enteropathy-associated T-cell lymphoma, refractory celiac disease, clonal evolution, T-cell receptor rearrangement, molecular oncogenesis

Grant support: This work was supported by the Netherlands MDL Foundation (SK18-27).

Conflict of interest disclosures for each author: None of the authors have any potential financial, professional or personal conflicts interests to disclose.

Data transparency and availability statement
The data underlying Figures 3, 4, 5, and Supplementary Figure S2 are available in the published article and its online Supplementary Material. The raw sWGS and WES data generated in this study was deposited at the European Genome-phenome Archive under accession code EGAS00001006669. Pipelines used for the preprocessing of NGS data can be found on Github: (I) Pipeline for CNA analysis: https://github.com/tgac-vumc/QDNAseq.snakemake (II) Pipeline for mutation calling with tumor-only mode: https://github.com/tgac-vumc/tumor-only-snakemake, and (III) Pipeline for mutation calling with tumor-normal mode: https://github.com/tgac-vumc/WES-snake. The codes associated to these repositories are accessible without restrictions.

Acknowledgements
We thank Paul P. Eijk from the Tumor Genomic Analyses Core (www.tgac.nl) of the Amsterdam UMC, Amsterdam, the Netherlands for his help and advice with the NGS experiments. We acknowledge the Hartwig Medical Foundation, Amsterdam, the Netherlands for providing access to reference whole-genome sequencing data of the Netherlands population. Santiago Bolaños Meneses and Ferdinand van den Toorn from the TGAC are acknowledged for expert data stewardship.

Corresponding author:
Prof. Dr. Bauke Ylstra
Dept. of Pathology, Tumor Genome Analysis Core – TGAC.nl
AmsterdamUMC, loc VUmc
Van der Boechorststraat 6B, Rm ZH 4J-008, 1081BT Amsterdam
M: +31 6 38306756,
E: b.ylstra@amsterdamumc.nl
ORCID: 0000-0001-9479-3010

List of abbreviations:

ASCL	adjacent small cell lymphocytosis
EATL	enteropathy-associated T-cell lymphoma
FFPE	formalin-fixed, paraffin embedded
IEL	intra-epithelial lymphocytes
NGS	next-generation sequencing
RCD	refractory celiac disease
sWGS	shallow Whole Genome Sequencing
TR	T-cell receptor
VAF	Variant allele frequency
WES	whole exome sequencing

Abstract

Refractory celiac disease (RCD) is a rare complication of celiac disease and bears a high risk for life-threatening enteropathy-associated T-cell lymphoma (EATL). While RCD is generally considered a direct EATL precursor, supporting evidence is limited and the predictive value for EATL development of molecular aberrations in RCD as risk factors unclear. We aim to dissect the clonal evolution of RCD and EATL to support molecular diagnostics-based clinical decision making.

RCD specimens of patients with (n=6) and without (n=21) development of EATL within 5 years after diagnosis, and EATL specimens (n=13) with their paired adjacent small cell lymphocytosis (ASCL) in the vicinity of the tumor were used for mutation and copy-number aberrations analysis using next-generation sequencing (NGS). In addition, NGS-based T-cell receptor (TR) clonotype analysis for 11 EATL/ASCL pairs was performed. Clonal evolutionary models were inferred and associations of genetic alterations in RCD to EATL risk assessed.

Integrated TR gene rearrangement and mutation profiles for 10/11 EATL and ASCL sample pairs shows mixed clonally shared and independent T-cell populations. Only in 2/5 EATL/RCD pairs, a related precursor population was demonstrated. Genetic aberrations were observed in 19/27 duodenal RCD samples irrespective of progression to EATL and in both RCD1 and RCD2.

EATL develops via parallel subclonal evolution and selection from a pool of locally residing jejunal T-cell populations with a restricted TR repertoire and a heterogeneous genetic makeup, but not necessarily from RCD populations in the duodenum. These findings provide leads for modifying current clinical decision-making for RCD and EATL patients.

Introduction

Celiac disease (CD) is a chronic inflammatory disorder and caused by consumption of gluten in genetically susceptible individuals with a prevalence of 1% in the European population.[1,2] Occasionally, patients present with severe persistent villous atrophy, or have recurrent malabsorption despite adhering to a strict gluten-free diet and are clinically diagnosed as having refractory CD (RCD).[3] Traditionally, a distinction is made between RCD type 1 (RCD1) and type 2 (RCD2), which is defined by the abundance of immunophenotypically abnormal intraepithelial lymphocytes (IELs) measured by flowcytometry, with an internationally accepted cut point of >20% to diagnose RCD2.[4,5] RCD2 generally shows mono- or sometimes oligoclonal T-cell receptor (TR) gene rearrangements, while RCD1 generally shows a polyclonal TR pattern.[5–7] RCD2 bears a high risk for severe and fatal CD complications, including a risk of 30 to 50% to develop enteropathy-associated T-cell lymphoma (EATL) within 5 years.[8–10] RCD1 patients have a lower risk of 3 to 14% to develop EATL within 5 years, and 5 year survival is much better.[8,9] Taken together, current diagnostic criteria within RCD are rather arbitrary.

EATL presents either as the first manifestation of CD in previously undiagnosed patients or after a period of a months to years of uncomplicated CD.[11] Only a minority presents after a period of RCD. The 5 year overall survival rate of EATL is only 10-30%.[12,13]

Current models of EATL oncogenesis typically assume a linear clonal evolution from CD, via RCD2 to EATL. Thereby, IELs with relatively subtle phenotypical aberrations (RCD2) would progress by acquiring genetic alterations to eventually present as morphologically highly atypical tumor cells with a highly aberrant immunophenotype, CD30 expression and extensive T-cell marker loss (EATL). This model is based on immunophenotypical characteristics related to the shared cell-of-origin and receptor (TR) rearrangements demonstrated in selected paired samples of RCD2/EATL and similarities in the overall genomic landscapes of RCD2 and EATL.[4,14–17] Thus far, however, actual evidence is very limited and as a consequence the impact of specific acquired alterations as (genetic) risk factors for EATL development is unclear. Such information, however, is a requirement for personalized advise and risk management of individual CD and RCD patients.

Here, we study the genetic evolutionary trajectories of 35 patients using RCD samples with and without progression to EATL, EATL samples, and populations of prominent epitheliotropic small cell T-cell infiltrates in their vicinity that are distinct from the EATL tumor mass, further designated as “adjacent small cell infiltrates” (ASCL) (Figure 1, Supplementary Figure S1A). With this patient sample series, we dissect the longitudinal evolution of EATL through its presumed low-risk and high-risk evolutionary stages.

Methods

Study cohort

Patients with RCD and EATL diagnosed and treated between 2004 and 2018 at the Gastroenterology Department of AmsterdamUMC were selected for this study. Informed consent was obtained from all patients, including for use of biopsy material for research purposes. This study was approved by the Medical Ethical Review Committee of VU University Medical Center (METC 2019.467) and Biobank Pathology Review Committee of VU University Medical Center (TcB 2017.023).

Minimum active follow-up of RCD patients who did not develop EATL was > 6 years. RCD1 and RCD2 classification was defined according to international criteria,[4,5] and progression to EATL was confirmed according to the current WHO Classification.[18]

In total 51 formalin-fixed paraffin-embedded (FFPE) tissue samples from 35 patients were collected (Figure 2; Supplementary Table 1). Thirteen EATLs were included of which two had developed de novo and 11 were previously diagnosed with RCD with an interval ranging <1 to 13 years. For all EATLs, paired samples were available: adjacent small cell lymphocytosis ASCL (n=8), RCD (n=2) and both ASCL

and RCD (n=3). A separate set of RCD samples were included of 22 patients of whom 21 had not developed EATL within 5 years after diagnosis.

Using H&E for morphological features and appropriate immunohistochemically stained sections highlighting immunophenotypical differences as a guide, representative areas of EATL and ASCL were demarcated by the hematopathologist (DDJ), carefully excluding any EATL contamination (Supplementary Methods, Supplementary Figure S1). Immunophenotypical features of EATL, ASCL and RCD are listed in Supplementary Tables 1 and 2. Demarcated regions were macro-dissected followed by DNA extraction.[19]

Shallow whole genome sequencing (sWGS), whole exome sequencing (WES) and Bio-informatic procedures.

ShallowWGS for CNAs, WES for mutation analysis and bioinformatic procedures were performed as previously described (Supplementary Methods).[20] The ACE algorithm (v0)[21] was used to quantify tumor content from sWGS CNA profiles. Variant allele frequencies (VAFs) of the mutations were corrected by copy-numbers and tumor cell percentage. All analyses were performed using R (v4.1.3).

After an extensive filtering strategy to include high confidence variants and exclude low confidence and germline variants, mutations identified by both Mutect2 (GATK4 v4.3.0) and LoFreq (v2.1.3.1) pipelines[22,23] were used for downstream analysis (Supplementary Methods). We then applied a virtual panel of 126 genes relevant to EATL and T-cell lymphomas (Supplementary Table 3) .[16,17,24]

NGS-based TR rearrangement analysis

NGS-based TR rearrangement analysis is described in detail in Supplementary Methods. In brief, a one-step protocols for detection of the TRBV-TRBD-TRBJ, TRBD-TRBJ and TRGV-TRGJ rearrangement is performed using primers as described by the EuroClonality-NGS Working Group[25] and NGS analyzed using Euroclonality ARResT/Interrogate software.[26] Clonotypes were identified according to the DEPART algorithm similar to Van de Brand et al.[27] and semi-quantified.[28]

Evolutionary modelling

Evolutionary trajectories were predicted from combined TR rearrangement, CNAs and mutation data of patient paired samples as described in (Supplementary Methods). Thereby, rearranged TRG and/or TRB alleles (clonotypes) were considered as a primary ground truth to determine clonal relations. For RCD samples, no TR data was available, and clonal relations were based on mutation and CNA data only.

Results

The genetic landscape of enteropathy-associated T-cell lymphoma, adjacent small cell lymphocytosis and refractory celiac disease

Enteropathy-associated T-cell lymphoma (EATL):
TR analysis confirmed monoclonality for 10/11 tested samples with varying addition of a polyclonal background pattern and/or the presence of up to 3 additional minor clones (Supplementary Table 4). TRG was most informative. Relative read abundance of bi-allelic TRG rearrangements was often disbalanced either due to chromosomal gains at the TRB locus (7q35) and/or TRG locus (7p15) as determined by CNAs (Supplementary Table 4 and 12). WES and sWGS analysis was performed for all EATL samples (n=13). Mutations were determined by WES to an average depth of 187 reads (range 152-245) for each EATL sample. CNAs and mutations were consistent with previously published information[16,17,24,29] and confirm that the study cohort is representative (Supplementary Results, Figure 3, Supplementary Figure 2, Supplementary Tables 5-12).

Adjacent small cell lymphocytosis (ASCL):
TR rearrangements in ASCL identical to the paired EATL were observed in 5/11 cases (Supplementary Table 3). In another 4/11 cases, the identical EATL-associated clonotypes were only detected as a minor component in the polyclonal background of the ASCL that showed a restricted repertoire. In 2/11 cases, a clonal relationship between ASCL and EATL based on TR rearrangement could not be observed. CNAs were significantly less frequent in ASCL samples compared to EATLs, only observed in 3/11 cases. Each of these CNAs was also detected in the matching EATL (Figure 3, Supplementary Figure S2).
Mutations were determined by WES to an average sequence depth of 318 reads (range 148-475x), which high depth was implemented to warrant sufficient sensitivity and comparability. A median of 11 non-synonymous mutations (range 4-83, mean 21) were identified in 11 ASCL samples (Figure 3, Supplemental Tables 5-12). A very similar mutational spectrum as observed in EATL was observed for *JAK/STAT* pathway mutations for all ASCL samples either as single alterations in *JAK1*, *STAT3, SH2B3 or SOCS1* (n=5) or combined (n=6) as well as for genes coding for epigenetic modification, histon modification and other regulating mutations. Also, in ASCL, multiple different mutations in the same gene within a single sample were seen at similar rates compared to EATL.

Refractory celiac disease (RCD):
Of 27 RCD samples (6 RCD1, 21 RCD2), 2 samples showed CNAs exceeding the standard cut point for statistical significance, both including 1q loss.
Mutations were determined by WES to an average depth of 280 reads (range 68-462). In 8/27 cases, no mutations were observed (Supplementary Table 13). In 19 cases, a median of 2 mutations per sample (range 1-10) were seen, including 8 JAK/STAT pathway mutations in 7 cases (25.9%). Other mutations were predominantly observed in epigenetic modification genes (37.0%). It should be noted that 3/6 samples that classified as RCD1 according to standard criteria harbored mutations and/or CNAs (Supplementary Tables 5-13).

Clonal evolution in ASCL and EATL

We constructed evolutionary trees for 11 ASCL/EATL pairs using shared and non-shared TR clonotypes with their relative NGS read abundance as a starting point, followed by shared and non-shared mutations with CNA-corrected VAFs, and CNAs as markers for their clonal and subclonal nature. Results of all cases are documented in Supplementary Figure S3 and selected cases are illustrated in Figure 4.
The EATL-associated monoclonal TR clonotypes could be traced back in the paired ASCL sample in 9/11 cases. The monoclonal clonotype was not the major product in ASCL in 4/9 pairs. In one case in which no clonal TR rearrangements in the EATL sample were observed, the EATL clone could be traced back in the ASCL sample based on a set of 5 shared identical mutations with high VAFs. Based on similar

VAFs within the sample, mutations were grouped and related to T-cell populations based on relative TR allele abundance subclonal evolution was recognized in both ASCL and EATL

The majority of mutations as well as secondary dominant TR clonotypes in the ASCL sample could not be assigned to the EATL precursor population, but rather to independent T-cell populations. Similarly, independent admixed aberrant T-cell populations were sometimes also present in the EATL sample. These findings support the notion that EATL develops in a background pool of aberrant small T-cells with varying levels of clonal outgrowth and clonal selection.

In 6 patients (#29, #34, #37, #52, #67, #82), JAK/STAT pathway gene mutations (*JAK1*, *STAT3, SOCS1, SH2B3*) that were observed in EATL were already present in the paired ASCL sample. In 7 patients, JAK/STAT pathway mutations (*JAK1*, *STAT3,* and regulators) were not an early event but only acquired at transformation to EATL or even later (#35, #40, #43, #52, #53, #72, #73). In various patients JAK/STAT pathway mutations were observed in ASCL cells that were not clonally related to EATL and/or in not clonally related T-cells admixed with tumor cells of the EATL samples (#4,#29, #40, #44, #72). This indicates that JAK/STAT pathway activation is an important but not uniformly initiating event in EATL lymphomagenesis.

Examples of clonal evolution

Case 37 (Figure 4A), illustrates evolution of an ASCL T-cell clone with bi-allelic, unproductive TRG rearrangements (with clonotypes TRG V3 -0/3/-6 J1=J2 and TRG V3 -0/7/-10) without a polyclonal background and *JAK1, CXCR4, CIITA* and *TP73* mutations (VAF at clonal level) together with a set of CNAs that was propagated to the EATL as clonal genetic alterations. At transformation, additional CNAs were acquired with subsequent subclonal evolution. Also in the ASCL sample, independent subclonal evolution was observed at the mutational level with at least two recognizable subclones with a larger subclone bearing additional *NCOR1, KMT2C* and *IKZF1* mutations (VAF >0.30) and a smaller one with *NOTCH1, BRD4* and *CIITA* mutations (VAF around 0.18). These alterations were not propagated to the EATL.

Case 44 (Figure 4B), illustrates clonal evolution from a clonal population in the ASCL sample with bi-allelic unproductive TRG rearrangements (clonotypes TRG V4 -0/2/-5 J1=J2 and TRG V5 -6/7/-1 J1=J2) and a set of mutations and CNAs, which based on VAF levels were possibly acquired at two stages of clonal evolution before transformation to EATL. Moreover, in the ASCL sample, a clonal TRB-incomplete rearrangement (TRB D2 -13/9/-6 J2-2) was also detected most probably representing an independent T-cell clone in the polyclonal background the ASCL sample since this rearrangement could not be identified in the EATL. In the ASCL sample also a large set of mutations with various VAF levels (ranging from 0.36 to 0.01) and including *JAK1* and various epigenetic modifiers, that were also not part of the EATL clone were observed. This illustrates the presence of multiple genetically aberrant T-cell clones and an active subclonal evolution process in the tumor region that also includes a direct EATL precursor.

In *case 40* (Figure 4C), based on shared bi-allelic unproductive TRG rearrangements (clonotypes TRG V5P -3/7/-5 JP and TRG V4 -1/0/-4 J1=J2) and *TENM2, TET2* and *KMT2D* mutations an ASCL clone is observed that is directly related as a precursor to the EATL. Additionally, a polyclonal irregular TR pattern is observed. This is consistent with an oligoclonal/restricted TR repertoire derived from independent admixed T-cell populations. These T-cell populations contain a set of mutations that is not found in the EATL sample. In the EATL sample, next to the tumor clone two populations were identified by their independent TRG rearrangements, suggesting admixture of various independent populations. Based on relative VAF levels, a set of private mutations was not likely to be present in the EATL clone, but rather in these independent T-cell populations. These are most likely similar to ASCL and would then not be morphologically distinguished from normal T-cells as part of the tumor

microenvironment (as illustrated in Supplementary Figure S1B-E). In the EATL clone, various CNAs and mutations, including two *JAK1* mutations were acquired during further subclonal evolution.

In *case 73* (Figure 4D), in the EATL sample, multiple dominant rearranged TRB and TRG alleles could be detected, which represent at least 2 T-cell population in a polyclonal background. Of these rearranged TR alleles, none could be traced back in the ASCL sample. A direct EATL precursor population could therefore not be identified in the ASCL population. In the corresponding RCD sample, only 2 CNAs were observed. Also, these bear no relation to ASCL or EATL clones.

Is RCD a step in clonal evolution to EATL?

In 5 patients, we could study the relation between EATL and a preceding RCD1 or RCD2 sample, of which 3 patients also had an ASCL sample available. The interval between the RCD duodenal biopsy sample and the jejunum-located EATL was 6-13 months. In all RCD samples genetic alterations could be demonstrated, with only mutations in 2 cases, 1 case with CNAs only and in 2 cases with both. In RCD34 a clonal relation with the subsequent EATL could be demonstrated based on shared identical mutations (*RASA2, STK10, SH2B3* VAF 0.136-0.066), identifying this population as a direct EATL precursor.

To verify if mutations could be missed due to insufficient sensitivity four RCD samples (RCD35, RCD67, RCD72, RCD73) were subjected to deeper sequencing (average depth 239-650). Thereby previous results were confirmed in 3 cases (RCD1: RCD35, RCD72; RCD2: RCD67) lacking any indications for clonal relations. For RCD72, an initially identified *PIK3CA* mutation (VAF 0.041), which was shared at identical base-pair level in the subsequent EATL was confirmed and an additional *GNAS* mutation was identified (VAF 0.031) that was also present in the background population of the ASCL sample.

These results indicate that RCD proliferations are genetically highly heterogeneous and in some cases harbor very small populations that can proceed to malignancy, but not uniformly so.

Examples of relations between RCD, ASCL and EATL

In *case 35* (Figure 5A), the RCD1 population in a duodenal biopsy sample shows more extensive CNAs and only a single PTPRS mutation (VAF 0.022). These genetic alterations are not propagated to the subsequent EATL at the jejunal site. A direct clonal relation is therefore not likely. No TR data are available for this case. On top of the clonal *JAK1* mutation (VAF 0.4) in the complete EATL population, clonal evolution has given rise subclones with two additional *STAT3* mutations (based on similar VAF levels of 0.12 likely bi-allelic in the same subclone) and various additional smaller subclones (VAFs 0.031-0.018)

In *case 72* (Figure 5B), an ASCL clone can be identified as a direct clonal precursor to EATL (TRB D1 -7/4/-4 J1-5 and a TRG V3 -0/9/-7 J1=J2 and/or a TRG V10 -0/9/-2 J1=J2), which based on an identical *PIK3CA* mutation can be linked to a population in RCD at the duodenal site. Next to this clonal evolutionary line, a population in ASCL outside the EATL tumor mass (characterized by the other clonal TRG rearrangement) as well as intermixed with the EATL tumor cells can be observed. Moreover, A second population within RCD1 in the duodenum characterized by a *GNAS* mutation can be traced to the ASCL population in the jejunum but is not part of the EATL clone.

Molecular features of RCD as risk factors for EATL development

We evaluated if genetic features in RCD can predict development to malignancy within 5 years (Supplementary Table 13). Five of 6 RCD progressing cases showed genetic mutations (median 3.8, range 1-10), while non-progressors showed mutations aberrations in 14/21 cases with a median mutation level of 1.5 (range 1-6) and CNAs in 2/21 cases. Mutations in the JAK-STAT pathway as well

as characteristic 1q loss were observed in 5/21 (23.8%) in non-progressor cases of which one with simultaneous *JAK1* and *STAT3* mutation and in 2/6 (33.3%) progressors. The majority of *JAK1* mutations were nonsense mutations indicating their pathogenic nature. Within the limitations of the small numbers of cases and genetic alterations, no unique predictive genetic features could be identified. Notably, treatment of RCD was rather heterogeneous, but distribution of limited (topical corticosteroids) and intensified treatments (amongst others cladribine, chemotherapy and stem cell transplantation) did not differ between the groups (Supplementary Table 14).
In summary, we did not find evidence to apply genetic profiling of RCD samples in the clinical setting to serve as a prognostic feature to guide clinical decision making.

Discussion

We demonstrate the complex clonal evolution of EATL from a regional pool of T-cell populations that have a small lymphocytic morphology, varying clonal outgrowth and a heterogeneous genetic makeup. This T-cell population also harbors the direct (clonally related) EATL precursor. Both ASCL populations and EATLs are subject to subclonal evolution, adding to the genetic complexity. In RCD biopsy samples diagnosed prior to EATL, only occasionally small cell populations that share identical mutations to ASCL or EATL are observed. Overall, the presence and rate of mutations and other genetic alterations are not specific for RCDs that progress to EATL.

Considering the current differentiation between RCD1 and RCD2 and resulting suboptimal risk predictions, the pertinent clinical question for patients with RCD is to provide an individual risk prediction for severe disease and transformation to EATL. In this study, we approached this issue by dissecting the longitudinal evolution of EATL though its presumed low-risk and high-risk evolution stages. In contrast to traditional methods that rely solely on PCR fragment length comparison, we applied NGS-based TR rearrangement analysis that generates a complete fingerprint of T-cell clones[25] allowed to recognize a direct clonally-related precursor population of the EATL in a mostly heterogeneous pool of T-cell populations in the ASCL samples in which clonal precursors were not always the dominant population. Complementary results were obtained by dissecting mutational spectra in ASCL and EATL phases. Parallel clonal evolution resulted in a remarkable pattern of multiple unique mutations in genes directly involved in the same signaling pathways ("redundancy"), predominantly in *JAK1, STAT3*, JAK/STAT regulators and epigenetic regulators. These finding explain the frequently reported multiple simultaneous *JAK1/STAT3* mutations in EATL samples[16,17,24,29] that were unexplained and seemed of no biological advantage. Moreover, oncogenesis did not follow an ordered step-wise process from RCD to EATL with characteristic early and late events. JAK/STAT dysregulation was not uniformly identified as an early driver event, but was regularly acquired only late in clonal evolution.

We observed mutations and CNAs in RCD samples of patients who did and did not develop EATL, confirmative of earlier observations[16] In line with recent observations by Singh and co-workers[30] but in contrast to previous studies[16,17], we demonstrate that such alterations are not restricted to RCD2 classified samples but are also observed in RCD1 classified samples. This further underpins that the distinction between RCD1 and RCD2 may not be a strict biological one, but rather a matter of sensitivity. In only 2 of 5 cases, identical genetic alterations present in EATL were also observed in paired RCD samples. This is in line with few published sample pairs in which an EATL clone was traced back to preceding RCD using PCR-based TR rearrangements[4,15], but also shows that this is not a common scenario and indicates that RCD biopsied at the duodenal site does not necessarily contain the precursor population that accounts for transformation to an EATL at the jejunal site. RCD can therefore be considered a risk factor for EATL, but not necessarily as a precursor lesion[30]. Interpretation of RCD2 as a low-grade lymphoproliferative disorder[18,31] or even low-grade lymphoma is therefore not justified.[32]

This study has various limitations. Like other orphan conditions, also our study is relatively small, but with these 35 patients and paired samples, we were able to cover the complete spectrum of EATL oncogenesis. With few RCDs evolving into EATL, conclusions on the predictive value of genetic alterations for EATL risk remains limited. Using the novel technique of TR-NGS clonotyping, we were able to monitor dominant and secondary TR clonality to determine clonal relations in greater detail, but with only limited access to RCD biopsy samples, we were unable to include TR rearrangement information for RCD and consequentially cannot exclude of a direct clonal relation between RCD and EATL.

In conclusion, we propose a model of oncogenesis of EATL in which under immunological pressure in the context of CD, a complex pool of small T-cell clones develops which is the target of genetic alterations.[33] This notion is strongly supported by Singh and co-workers[30] who applied single cell multi-omic technologies in samples of 10 RCD1 and 2 RCD2 individuals to show evolution of independent, parallel T-cell populations in both RCD1 and RCD2 that are heterogeneous at the immunophenotypical level (sCD3, CD4, CD8, CD103 status) and bear multiple independent pathogenic mutations (*JAK1, STAT3, DDX3X, DNMT3A* and others). While such T-cell clones may be biopsied at the duodenal site, the immunological pressure in CD patients is present throughout the small intestine and oligoclonal T-cell pools can be expected to be present throughout the small intestine. Moreover, RCD clones have been shown to be prone to recirculation and can be observed in the peripheral blood and at various organ sites.[15] As the immunological pressure persists, aberrant T-cell clones remain at risk of acquiring further oncogenic events and in a prolonged process of clonal evolution and selection may eventually lead to transformation to EATL.[7] Thereby, our findings directly extent the recently proposed model of parallel clonal evolution for RCD[30] to the subsequent ASCL and EATL stages. This oncogenetic scenario is not unique and parallels the oncogenesis of extranodal marginal zone lymphoma of mucosa-associated lymphoid tissue type (EMZL) in which (microorganism-initiated) autoimmunity is the primary driver of a B-cell deregulation which are also at continuous risk for accumulating additional oncogenic events with subsequent prolonged clonal evolution and selection.[34] The clinical consequence of these findings is that RCD remains an important risk factor for the development of EATL, but challenges the notion of duodenal biopsies to be used for identifying precursor lesions for EATL. The presence of genetic alterations, even JAK/STAT pathway alterations, may not provide sufficient additional evidence for risk assessment and hence would probably be insufficient to guide treatment choices in the light of EATL prevention leading to both over- and undertreatment. Similarly, when considering JAK/STAT as a treatment target such mutation should be specific for the malignant EATL clone, distinct from alterations in admixed pre-malignant populations.

References

1. Lindfors K, Ciacci C, Kurppa K, et al. Coeliac disease. Nat Rev Dis Prim. 2019;5(1):3.
2. Burger JPW, Roovers EA, Drenth JPH, Meijer JWR, Wahab PJ. Rising incidence of celiac disease in the Netherlands; an analysis of temporal trends from 1995 to 2010. Scand J Gastroenterol. 2014;49(8):933-941.
3. van Wanrooij RLJ, Bouma G, Bontkes HJ, et al. Outcome of referrals for non-responsive celiac disease in a tertiary center: low incidence of refractory celiac disease in the Netherlands. Clin Transl Gastroenterol. 2017;8(1):e218.
4. Verbeek WHM, Goerres MS, von Blomberg BME, et al. Flow cytometric determination of aberrant intra-epithelial lymphocytes predicts T-cell lymphoma development more accurately than T-cell clonality analysis in refractory celiac disease. Clin Immunol. 2008;126(1):48-56.
5. Green PHR, Paski S, Ko CW, Rubio-Tapia A. AGA clinical practice update on management of refractory celiac disease: expert review. Gastroenterology. 2022;163(5):1461-1469.
6. Hussein S, Gindin T, Lagana SM, et al. Clonal T cell receptor gene rearrangements in coeliac disease: implications for diagnosing refractory coeliac disease. J Clin Pathol. 2018;71(9):825-831.
7. Liu H, Brais R, Lavergne-Slove A, et al. Continual monitoring of intraepithelial lymphocyte immunophenotype and clonality is more important than snapshot analysis in the surveillance of refractory coeliac disease. Gut. 2010;59(4):452-460.
8. Rubio-Tapia A, Kelly DG, Lahr BD, et al. Clinical staging and survival in refractory celiac disease: a single center experience. Gastroenterology. 2009;136(1):93-99.
9. Malamut G, Afchain P, Verkarre V, et al. Presentation and long-term follow-up of refractory celiac disease: comparison of type I with type II. Gastroenterology. 2009;136(1):81-90.
10. Malamut G, Cellier C. Refractory celiac disease. Expert Rev Gastroenterol Hepatol. 2014;8(3):323-328.
11. van Gils T, Nijeboer P, Overbeek LI, et al. Risks for lymphoma and gastrointestinal carcinoma in patients with newly diagnosed adult-onset celiac disease: Consequences for follow-up: Celiac disease, lymphoma and GI carcinoma. United Eur Gastroenterol J. 2018;6(10):1485-1495.
12. West J, Jepsen P, Card TR, Crooks CJ, Bishton M. Incidence and survival in patients with enteropathy-associated t-cell lymphoma: Nationwide Registry studies from England and Denmark. Gastroenterology. 2023;165(4):1064-1066.e3.
13. Meeuwes FO, Brink M, Plattel WJ, et al. enteropathy-associated t-cell lymphoma: a population-based cohort study on incidence, treatment, and outcome in the netherlands. EJHaem. 2024;5(6):1215-1222.
14. Schmitz F, Tjon JML, Lai Y, et al. Identification of a potential physiological precursor of aberrant cells in refractory coeliac disease type II. Gut. 2013;62(4):509-519.
15. Cellier C, Delabesse E, Helmer C, et al. Refractory sprue, coeliac disease, and enteropathy-associated T-cell lymphoma. French Coeliac Disease Study Group. Lancet. 2000;356(9225):203-208.
16. Cording S, Lhermitte L, Malamut G, et al. Oncogenetic landscape of lymphomagenesis in coeliac disease. Gut. 2022;71(3):497-508.
17. Soderquist CR, Lewis SK, Gru AA, et al. Immunophenotypic spectrum and genomic landscape of refractory celiac disease type II. Am J Surg Pathol. 2021;45(7):905-916.

18. Bhagat G, Cerf-Bensussan N, Dave SS, Naresh KN. Enteropathy-associated T-cell lymphoma. In: WHO Classification of Tumours Editorial Board. Haematolymphoid tumours. Lyon (France): International Agency for Research on Cancer. WHO classification of tumours series, 5th ed. 2024, vol. 11, pp. 717-721.

19. van Essen HF, Ylstra B. High-resolution copy number profiling by array CGH using DNA isolated from formalin-fixed, paraffin-embedded tissues bt - genomic structural variants: methods and protocols. Methods Mol Biol. 2012;838:329-341.

20. Mendeville MS, Janssen J, Los-de Vries GT, et al. Integrating genetic subtypes with PET scan monitoring to predict outcome in diffuse large B-cell lymphoma. Nat Commun. 2025;16(1):109.

21. Poell JB, Mendeville M, Sie D, et al. ACE: absolute copy number estimation from low-coverage whole-genome sequencing data. Bioinformatics. 2019;35(16):2847-2849.

22. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotech. 2013;31(3):213-219.

23. Wilm A, Aw PPK, Bertrand D, et al. LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 2012;40(22):11189-11201.

24. Roberti A, Dobay MP, Bisig B, et al. Type II enteropathy-associated T-cell lymphoma features a unique genomic profile with highly recurrent SETD2 alterations. Nat Commun. 2016;7:12602.

25. Brüggemann M, Kotrová M, Knecht H, et al. Standardized next-generation sequencing of immunoglobulin and T-cell receptor gene recombinations for MRD marker identification in acute lymphoblastic leukaemia; a EuroClonality-NGS validation study. Leukemia. 2019;33(9):2241-2253.

26. Bystry V, Reigl T, Krejci A, et al. ARResT/Interrogate: an interactive immunoprofiler for IG/TR NGS data. Bioinformatics. 2017;33(9):435-437.

27. van den Brand M, Möbs M, Otto F, et al. EuroClonality-NGS recommendations for evaluation of b-cell clonality analysis by next-generation sequencing: a structured approach with the DEPART algorithm. J Mol Diagn. 2023;25(10):729-739.

28. Langerak AW, Groenen PJTA, Brüggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. Leukemia. 2012;26(10):2159-2171.

29. Moffitt AB, Ondrejka SL, McKinney M, et al. Enteropathy-associated T cell lymphoma subtypes are characterized by loss of function of SETD2. J Exp Med. 2017;214(5):1371-1386.

30. Singh M, Louie RHY, Samir J, et al. Expanded T cell clones with lymphoma driver somatic mutations accumulate in refractory celiac disease. Sci Transl Med. 2025;17(798):eadp6812.

31. Campo E, Jaffe ES, Cook JR, et al. The international consensus classification of mature lymphoid neoplasms: a report from the clinical advisory committee. Blood. 2022;140(11):1229-1253.

32. Malamut G, Soderquist CR, Bhagat G, Cerf-Bensussan N. Advances in nonresponsive and refractory celiac disease. Gastroenterology. 2024;167(1):132-147.

33. Ettersperger J, Montcuquet N, Malamut G, et al. Interleukin-15-dependent t-cell-like innate intraepithelial lymphocytes develop in the intestine and transform into lymphomas in celiac disease. Immunity. 2016;45(3):610-625.

34. Tzioni M-M, Watanabe N, Chen Z, et al. Primary thyroid B-cell lymphoma: molecular insights into its clonal evolution and relapse. J Pathol. 2025;265(2):123-131.

Figure legends

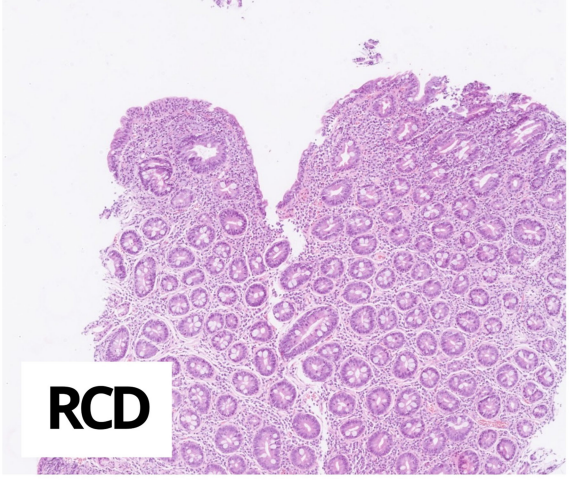
Figure 1: Histomorphological and immunohistochemical features. RCD (row 1), ASCL (row 2) and EATL (row 3) tissue samples. Left column Hematoxyline en Eosine (H&E) staining; middle column CD3 immunohistochemistry (IHC) staining; Right column CD30 IHC.
In total 51 formalin-fixed paraffin-embedded (FFPE) tissue samples from 35 patients

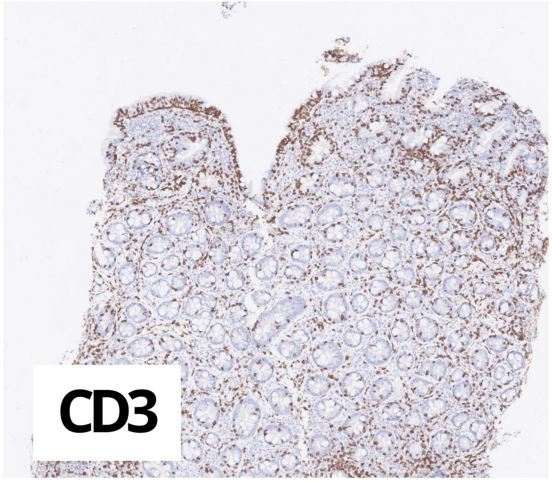
Figure 2: Schematic of the 51 FFPE samples from 35 patients included in the study. Upper box RCD patients and samples without progression (blue), middle box RCD patient and sample with progression (green). The lower boxes show paired and triplet samples of patients with RCD (green), ASCL (yellow) and EATL samples (red). pt(s), patient(s).

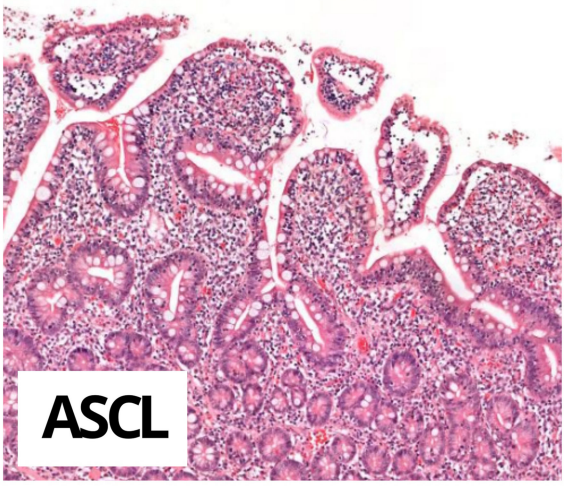
Figure 3: Oncoprint depicting unique and shared genetic alterations and T-cell clonality of paired and triplet patient samples. Genetic alterations (rows) in sample pairs/triplets (columns) of RCD progressed < 5 years (yellow, n=5), ASCL (orange, n=11) and EATL (red, n=13). The left axis indicates genes, grouped by pathways or chromosomal arms (CNA) on the far left. The right y-axis indicate the relative frequency of genetic alterations for each of the 3 histological groups. Top x-axis from top to bottom are the frequencies of each mutation and CNAs, the composition of the pairs and triplets (Cohort), type of T-cell rearrangements (TR) and degree of clonal relationship (related). The nature of the mutations is indicated by colour coding, missense, nonsense, nonstop, splice site, frame-shift deletion, in-frame deletion (legend far right, alterations). Private (non-shared) mutations are indicated by symbols in the boxes, and lettered from -a, to -f. In case multiple mutations in a single gene are found in an individual sample, these symbols are combined. Shared mutations are depicted by boxes without further symbols (or shared symbols in case of multiple alterations in single samples). Information on shared and private alterations can also be found graphically depicted in Supplementary Figure 3 (circos plots) and in Supplementary Table 7.

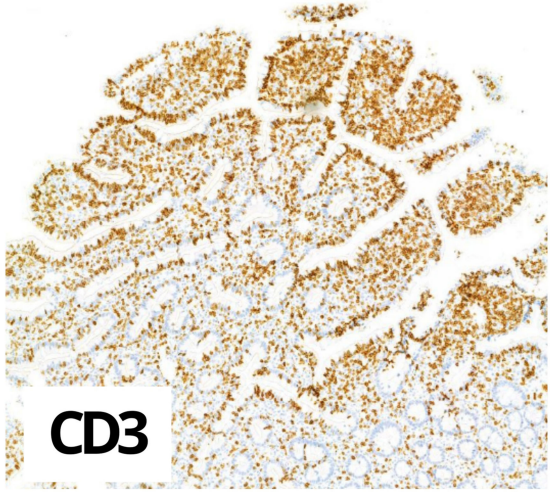
Figure 4: Variants of clonal evolution in ASCL and EATL pairs as inferred from TR clonotypes, shared and private clonal and non-clonal mutations and CNAs. A, patient #37, B. Patient #44, C patient #40, D patient #73. The EATL population is depicted in red, and the admixed independent T-cell populations that are not part of the EATL clone are indicated in orange. ASCL populations are depicted in yellow and RCD in green. Relative population sizes with the samples are estimated based on TR clonotype abundance and VAF levels of mutations. X-CNAs, multiple additional chromosomal copy number aberrations; LOH, Loss of heterozygosity

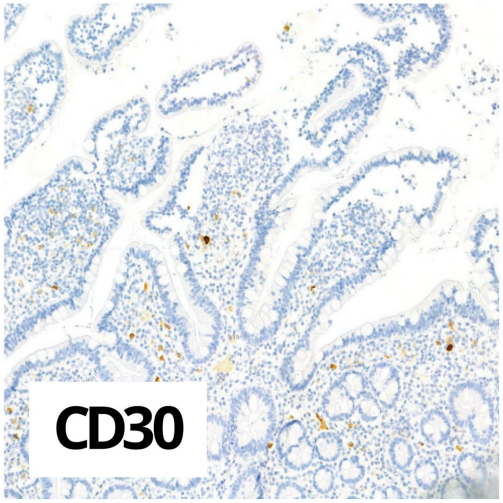
Figure 5: Variants of clonal evolution in RCD, ASCL and EATL as inferred from TR clonotypes, shared and private clonal and non-clonal mutations and CNAs. A, patient #35, B. Patient #72.

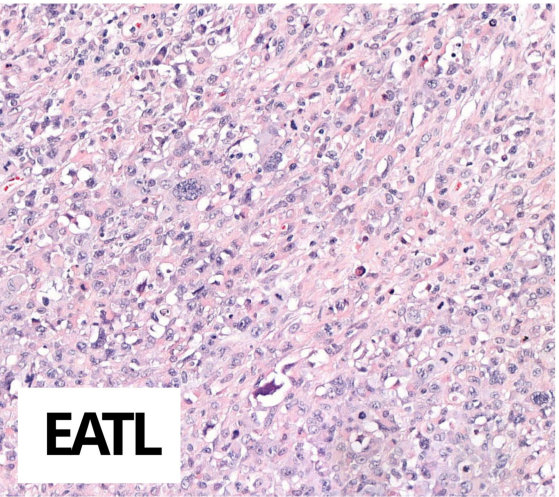

RCD

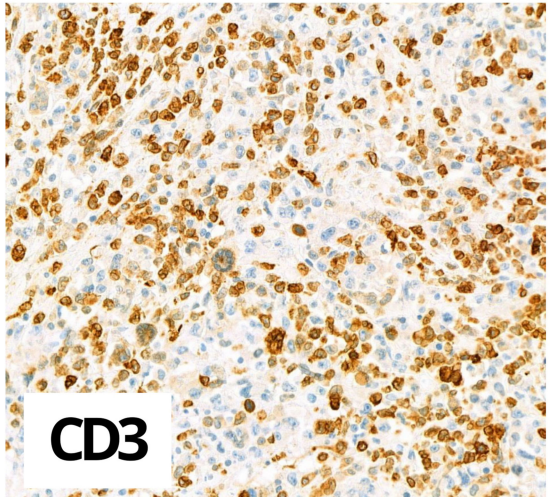

CD3

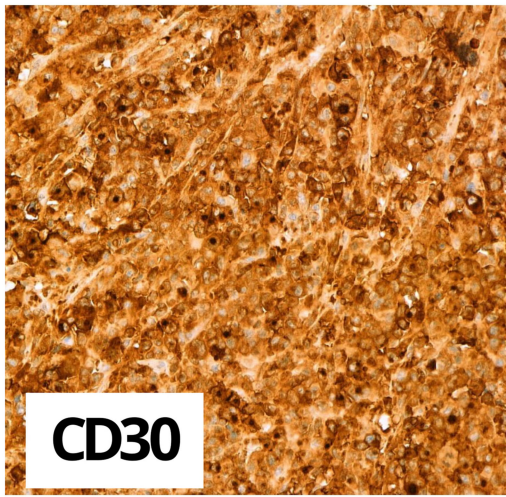

ASCL


CD3


CD30


EATL


CD3


CD30

Single:
RCD-

n=21 pts & samples

Single: RCD, n=1 pt & sample

Pairs:
-
ASCL
EATL

n=8 pts &
16 samples

Triplets:
RCD
ASCL
EATL

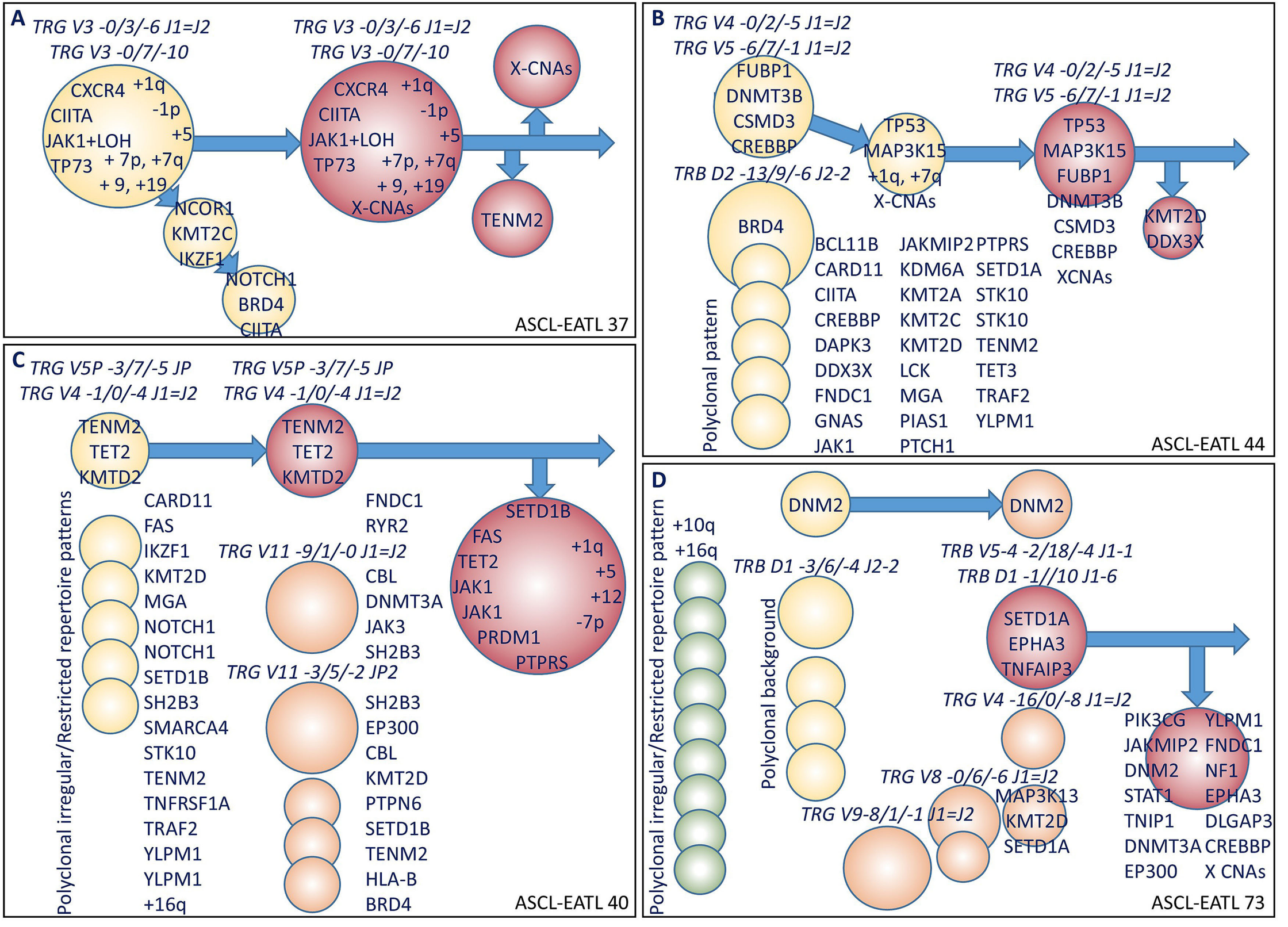
n=3 pts &
9 samples

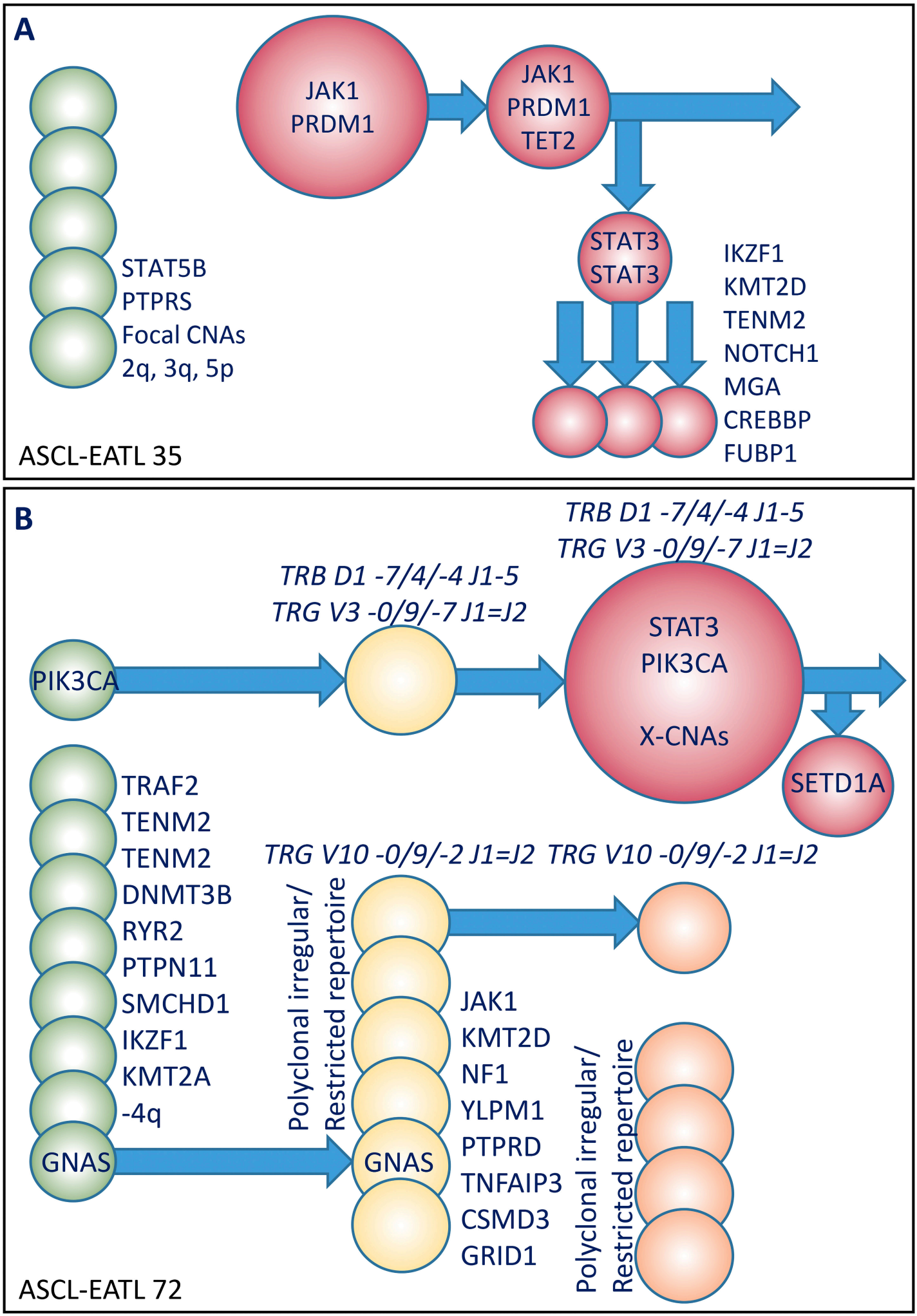
Pairs:
RCD
-
EATL

n=2 pts &
4 samples

= RCD- no progression to EATL n = 21
= RCD with progression to EATL n= 6
= ASCL n = 11 samples
= EATL n = 13 samples


JAK/STAT signalling
Epigenetic modification
Transcription regulation
Receptor tyrosine kinase
Other
RAS/MAP kinase signalling
DNA damage repair
NFkB
Immune regulation
PI3K/AKT signalling
CNA
JAK1
STAT3
SH2B3
PTPRD
KMT2D
KMT2C
SETD1A
TET2
CREBBP
SMARCA4
DNMT3B
SETD1B
DNMT3A
KMT2A
TET3
DOT1L
EP300
PRDM1
TENM2
BRD4
DDX3X
TP73
YLPM1
BCL11A
BCL11B
IKZF1
TDRD6
PTPRS
MGA
RYR2
CSMD3
FNDC1
GNAS
NOTCH1
FUBP1
STK10
EPHA3
PTCH1
MAP3K13
MAP3K15
NF1
RASA2
TERT
TP53
TRAF2
CARD11
CIITA
CXCR4
HLA−B
PIK3CA
1p
1q
5p
5q
6p
6q
7p
7q
8q
9q
10p
10q
12p
12q
16q
80
60
40
20
0
Cohort
TR
Related
ASCL−4
EATL−4
ASCL−52
EATL−52
ASCL−37
EATL−37
RCD−35
EATL−35
ASCL−40
EATL−40
ASCL−44
EALT−44
RCD−72
ASCL−72
EATL−72
RCD−73
ASCL−73
EATL−73
ASCL−29
EATL−29
RCD−67
ASCL−67
EATL−67
RCD−34
EATL−34
ASCL−43
EATL−43
ASCL−82
EATL−82
RCD progressed
ASCL
EATL
0
0.5
1
Cohort
RCD progressed <5yr
ASCL
EATL
TR
Restricted pattern
Clonal in restricted background
Clonal in polyclonal background
Bi−clonal
Clonal
Related
Not related
Related and not−related
Related, small component
Related
Alterations
MISSENSE
MISSENSE−a
MISSENSE−b
MISSENSE−c
MISSENSE−d
MISSENSE−e
MISSENSE−f
NONSENSE
NONSENSE−a
NONSTOP
SPLICE_SITE
SPLICE_SITE−a
FRAME_SHIFT_DEL
FRAME_SHIFT_DEL−a
IN_FRAME_DEL
IN_FRAME_DEL−a
Loss
Gain


A
TRG V3 -0/3/-6 J1=J2
TRG V3 -0/7/-10
CXCR4 +1q
CIITA -1p
JAK1+LOH +5
TP73 + 7p, +7q
+ 9, +19
NCOR1
KMT2C
IKZF1
NOTCH1
BRD4
CIITA
TRG V3 -0/3/-6 J1=J2
TRG V3 -0/7/-10
CXCR4 +1q
CIITA -1p
JAK1+LOH +5
TP73 +7p, +7q
+ 9, +19
X-CNAs
X-CNAs
TENM2
ASCL-EATL 37
B
TRG V4 -0/2/-5 J1=J2
TRG V5 -6/7/-1 J1=J2
FUBP1
DNMT3B
CSMD3
CREBBP
TP53
MAP3K15
+1q, +7q
X-CNAs
TRG V4 -0/2/-5 J1=J2
TRG V5 -6/7/-1 J1=J2
TP53
MAP3K15
FUBP1
DNMT3B
CSMD3
CREBBP
XCNAs
KMT2D
DDX3X
TRB D2 -13/9/-6 J2-2
BRD4
Polyclonal pattern
BCL11B
CARD11
CIITA
CREBBP
DAPK3
DDX3X
FNDC1
GNAS
JAK1
JAKMIP2
KDM6A
KMT2A
KMT2C
KMT2D
LCK
MGA
PIAS1
PTCH1
PTPRS
SETD1A
STK10
STK10
TENM2
TET3
TRAF2
YLPM1
ASCL-EATL 44
C
TRG V5P -3/7/-5 JP
TRG V4 -1/0/-4 J1=J2
TENM2
TET2
KMTD2
TENM2
TET2
KMTD2
SETD1B
FAS
TET2
JAK1
JAK1
PRDM1
PTPRS
+1q
+5
+12
-7p
Polyclonal irregular/Restricted repertoire patterns
CARD11
FAS
IKZF1
KMT2D
MGA
NOTCH1
NOTCH1
SETD1B
SH2B3
SMARCA4
STK10
TENM2
TNFRSF1A
TRAF2
YLPM1
YLPM1
+16q
TRG V11 -9/1/-0 J1=J2
FNDC1
RYR2
CBL
DNMT3A
JAK3
SH2B3
TRG V11 -3/5/-2 JP2
SH2B3
EP300
CBL
KMT2D
PTPN6
SETD1B
TENM2
HLA-B
BRD4
ASCL-EATL 40
D
DNM2
DNM2
TRB V5-4 -2/18/-4 J1-1
TRB D1 -1//10 J1-6
+10q
+16q
TRB D1 -3/6/-4 J2-2
Polyclonal irregular/Restricted repertoire pattern
Polyclonal background
SETD1A
EPHA3
TNFAIP3
TRG V4 -16/0/-8 J1=J2
PIK3CG
JAKMIP2
DNM2
STAT1
TNIP1
DNMT3A
EP300
YLPM1
FNDC1
NF1
EPHA3
DLGAP3
CREBBP
X CNAs
TRG V8 -0/6/-6 J1=J2
MAP3K13
KMT2D
SETD1A
TRG V9-8/1/-1 J1=J2
ASCL-EATL 73


A
JAK1
PRDM1
JAK1
PRDM1
TET2
STAT3
STAT3
STAT5B
PTPRS
Focal CNAs
2q, 3q, 5p
IKZF1
KMT2D
TENM2
NOTCH1
MGA
CREBBP
FUBP1
ASCL-EATL 35
B
TRB D1 -7/4/-4 J1-5
TRG V3 -0/9/-7 J1=J2
TRB D1 -7/4/-4 J1-5
TRG V3 -0/9/-7 J1=J2
PIK3CA
STAT3
PIK3CA
X-CNAs
SETD1A
TRAF2
TENM2
TENM2
DNMT3B
RYR2
PTPN11
SMCHD1
IKZF1
KMT2A
-4q
GNAS
TRG V10 -0/9/-2 J1=J2
TRG V10 -0/9/-2 J1=J2
Polyclonal irregular/
Restricted repertoire
GNAS
JAK1
KMT2D
NF1
YLPM1
PTPRD
TNFAIP3
CSMD3
GRID1
Polyclonal irregular/
Restricted repertoire
ASCL-EATL 72

Enteropathy-Associated T-cell lymphoma: molecular clonal evolution and its clinical consequences. Ylstra *et al.*

Supplementary Methods

Study cohort

Pathology and sampling: Histopathologically, regions of extensive T-cell infiltrates with small cell morphology and with prominent epitheliotropic features are commonly observed around EATL tumor sites in small intestinal resection specimens (Figure 1, Supplementary Figure S1A). These infiltrates lack overt aberrant morphological features, clearly differentiating these from the tumor cells, but often show some aberrant immunophenotypical features (varying T-cell marker loss, aberrant marker expressions). CD30 expression is generally negative. For this study, we have designated these infiltrates "adjacent small cell infiltrates" (ASCL). In the present series, in 11 jejunal resection specimens ASCL areas were present in a 1-1.5 cm margin adjacent to the EATL tumor mass. Immunophenotypically, these infiltrates were positive for CD3 (10/11) and CD7 (9/9) and showed marker loss for CD5 (7/7) and CD2 (3/9) and were negative for CD30 with only weak and partial (20%) expression in 2/11 cases. Aberrant CD79a was observed in one case (Supplementary Table 2, Figure 1 and Supplementary Figure S1).

Shallow whole genome sequencing (sWGS) and whole exome sequencing (WES)

Shallow WGS for copy number aberration (CNA) analysis and whole exome sequencing (WES) for mutation analysis were performed as previously described[1]. Genomic DNA extraction was performed according to the manufacturer's protocols (GeneRead DNA FFPE kit, Qiagen, Hilden, Germany) and genomic libraries were prepared from 100ng of DNA according to the manufacturer's protocols. DNA sequencing libraries were prepared using KAPA HyperPlus Library Preparation. For copy number analysis, libraries were subjected to sWGS using a 50bp single read modus on the NextSeq2000 machine (Illumina San Diego, CA).

For mutation analysis, exome target enrichment was performed from the same libraries using the KAPA HyperCap workflow v3.0 and sequenced on an Illumina NovaSeq 6000. After filtering duplicate and off-target reads, the mean coverage depth of exonic regions was approximately 50x for samples from normal tissue, approximately 190x for tumor samples and approximately 285x for RCD and ASCL samples to allow for optimal comparability of all samples across the cohort according to tumor/aberrant cell content. Additionally, 4 RCD samples were further sequenced to a mean coverage depth of approximately 500x.

Bioinformatic procedures for mutations and copy number aberrations

After genomic DNA extraction, libraries were prepared (100ng DNA) and subjected to sWGS using a 50bp single read modus on NextSeq2000 (Illumina, San Diego, CA, USA). For mutation analysis, exome targeted enrichment was performed from the same DNA libraries and sequenced on the Illumina NovaSeq 6000 to approximately 190x depth for tumor DNA samples and approximately 285x for RCD- and ASCL- DNA samples to allow for optimal comparability of all samples across the cohort. Additionally, 4 RCD DNA samples were further sequenced to a mean coverage of approximately 500x. We performed genome-wide copy number alteration (CNA) analysis on shallow whole-genome sequencing (sWGS) data using the QDNAseq.snakemake pipeline (available at https://github.com/tgac-vumc/QDNAseq.snakemake)[2]. The workflow first aligns raw reads with BWA, followed by duplicate marking with Picard. CNA profiles were then generated using the R package QDNAseq (v3.21) at a 100

kbp resolution and segmented using CGHcall (v3.21)[23]. To quantify tumor content and compare clonal sizes, we integrated the CNA profiles using the ACE algorithm (v0)[4].
Somatic mutations were identified using two comparable pipelines: a tumor-only workflow (https://github.com/tgac-vumc/tumor-only-snakemake) and a tumor-normal workflow (https://github.com/tgac-vumc/WES-snake). Both pipelines employ Mutect2 (GATK4 v4.3.0) and LoFreq (v2.1.3.1) for variant calling[5,6]. From the samples with available matched normal samples, we constructed a preservation list using the tumor-normal workflow. High-confidence variants were identified as those detected by both callers that met the following criteria: minimum coverage of 10x, minimum allele frequency of 0.01, at least 2 reads per strand, a ROQ score of ≥20, and a TLOD score of ≥10 intersected with BCFtools 1.9. These mutations were annotated by snpSift (v4.3) and funcotator (v4.1.9)18 using the databases of COSMIC (v84), Gnomad (v2.0.2), gencode(v34), dbsnp (build 151), clinvar (20180401), and Hartwig Medical Foundation panel of normals (v2.0). Mutations labeled as common in dbsnp or present >3 times in HMF panel of normals were excluded. Intronic, silent, UTR or flanking mutations were removed for further analysis. Next, we applied the tumor-only workflow for the entire cohort, to be consistent across the samples. This pipeline uses panel of normal (PON) to remove known germline mutations. Finally, mutations identified by both LoFreq and Mutect2, or those listed in preservation list and called by only one of LoFreq/Mutect2 were used for the downstream analysis. We then applied a virtual panel of 126 genes to study only genes relevant to EATL and T-cell lymphomas (Supplementary Table 3)[7–9]. Variant allele frequencies (VAFs) of the mutations are corrected by copy-numbers and tumor cell percentage for the subsequent clonal evolution analysis. Oncoprints were created with Complexheatmap (2.21.2). All analyses were performed using R (v4.1.3).

NGS-based TR rearrangement analysis
NGS-based TR rearrangement analysis was performed using the primers as described by the EuroClonality-NGS Working Group[10]. For this study, the one-step protocols for detection of the TRBV-TRBD-TRBJ, TRBD-TRBJ and TRGV-TRGJ rearrangement were performed. Libraries were sequenced on an Illumina NextSeq-500 system (Illumina, Inc. San Diego, CA, USA) and analyzed using Euroclonality ARResT/Interrogate software, developed within the EuroClonality NGS Working Group[11]. The nucleotide sequences were assessed for the presence of V, (D) and J genes and the junctional region containing the complementary determining region 3, which is represented as an amino acid sequence. With this information the sequences are attributed to clonotypes, which are characterized by the same V and J gene and junctional region. Targets were evaluated according to the DEPART algorithm similar to Van de Brand *et al.*[12]. Abundance of clonotypes was calculated based on the number of reads of a clonotype in relation to all clonotypes of the specific type of rearrangement. The molecular conclusion for the clonality status was based on the integration of the results of the individual clonality targets, i.e: TRBV-TRBD-TRBJ, TRBD-TRBJ and TRGV-TRGJ rearrangements, according to standard EuroClonality strategies for the conventional GeneScan-based assays[13].

Evolutionary modelling
Evolutionary trajectories were predicted from combined TR rearrangement, CNAs and mutation data in paired samples. For ASCL-EATL pairs, shared rearranged TRG and/or TRB alleles (clonotypes) were used to determine a direct clonal relation (ASCL-EATL clone). Additional rearranged TR alleles or an irregular polyclonal/restricted repertoire were observed in the ASCL or EATL sample, these were considered as independent admixed T-cell populations. Abundance levels of rearranged TR alleles were considered as the relative proportion of T-cell populations within the sample. Following rearrangements, shared mutations were assigned to the ASCL-EATL clone. For those samples that, based on TR results, lacked independent admixed T-cell populations in the background, non-shared (private) mutations were

considered subclonal evolution. VAF levels suggested the relative population size of the subclone(s) (dominant, significant, minor population). In case TR results indicated the presence of independent T-cell populations next to the ASCL-EATL clone, private mutations could not always be unequivocally assigned. Based on similar VAFs within the sample, mutations were grouped and related to common T-cell populations based on relative secondary TR allele abundance to infer most probable evolutionary models. For RCD samples, no TR data was available, and clonal relations were based on CNA and mutations data only.

Supplementary results

The genetic landscape of enteropathy-associated T-cell lymphoma

WES and sWGS analysis was performed in all EATL samples (n=13). Mutations were determined by WES to an average depth of 187 reads (range 152-245) for each EATL sample.

The genomic landscape of CNAs for the 15 most commonly gained or lost chromosomal arms and mutational spectrum is visualized in Figure 2 and Supplementary Figure S2). The most frequent CNAs in 13 EATL samples were gains of chromosome 1q (9/13, 69%), 7q with or without chromosome 7p (8/13, 61.5%), 9q (8/13, 61.5%), chromosome 5 (7/13, 53.8%) and to a lesser extent chromosome 12 and 8q (both 4/13, 30.7%). Loss of chromosome 10p (7/13, 53.8%) with and without 10q (4/13) and 6q (4/13, 30.7%) were seen and commensurate with what is reported in the literature[8]. A median of 13 non-synonymous mutations (range 3-54, mean 16) were identified in the 13 EATL samples in the 126 genes of the virtual panel (Figure 2, Supplementary Tables 2-10). In this cohort, we confirmed the high frequency of JAK/STAT pathway deregulation in 12/13 (92.3%) EATLs, either by single mutations in STAT3 (n=3) or their regulator SOCS1 (n=1), but more often (8/13,61.5%) by multiple mutations both in JAK1, STAT3 and STAT1 as well as various regulating genes (SH2B3, JAKMIP2, PTPRD, PTPRC) within a single sample. A high level of mutations in genes affecting DNA methylation and histon modification was observed in 12/13 EATLs with most frequent synchronous mutations in multiple genes including TET3, TET2, KMT2D, KMT2C, DOTL1, SETD1A, EP300 and CREBBP. Also previously reported genes involved in DNA damage repair (TP53, TERT, POT1), NFkB pathway signaling (TNFAIP3, CARD11, TRAF2) and other regulatory genes were confirmed.

Taken together, these data are consistent with previously published information and confirm that this cohort is representative.

Supplementary References

1. Mendeville MS, Janssen J, Los-de Vries GT, et al. Integrating genetic subtypes with PET scan monitoring to predict outcome in diffuse large B-cell lymphoma. Nat. Commun. 2025;16(1):109.
2. Scheinin I, Sie D, Bengtsson H, et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. Genome Res. 2014;24(12):2022–2032.
3. van de Wiel MA, Kim KI, Vosse SJ, et al. CGHcall: Calling aberrations for array CGH tumor profiles. Bioinformatics. 2007;23(7):892–894.
4. Poell JB, Mendeville M, Sie D, et al. ACE: absolute copy number estimation from low-coverage whole-genome sequencing data. Bioinformatics. 2019;35(16):2847–2849.
5. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotech. 2013;31(3):213–219.
6. Wilm A, Aw PPK, Bertrand D, et al. LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 2012;40(22):11189–11201.
7. Soderquist CR, Lewis SK, Gru AA, et al. Immunophenotypic Spectrum and Genomic Landscape of Refractory Celiac Disease Type II. Am. J. Surg. Pathol. 2021;45(7):905–916.
8. Cording S, Lhermitte L, Malamut G, et al. Oncogenetic landscape of lymphomagenesis in coeliac disease. Gut. 2022;71(3):497–508.
9. Roberti A, Dobay MP, Bisig B, et al. Type II enteropathy-associated T-cell lymphoma features a unique genomic profile with highly recurrent SETD2 alterations. Nat. Commun. 2016;7:12602.
10. Brüggemann M, Kotrová M, Knecht H, et al. Standardized next-generation sequencing of immunoglobulin and T-cell receptor gene recombinations for MRD marker identification in acute lymphoblastic leukaemia; a EuroClonality-NGS validation study. Leukemia. 2019;33(9):2241–2253.
11. Bystry V, Reigl T, Krejci A, et al. ARResT/Interrogate: an interactive immunoprofiler for IG/TR NGS data. Bioinformatics. 2017;33(3):435–437.
12. van den Brand M, Möbs M, Otto F, et al. EuroClonality-NGS Recommendations for Evaluation of B-Cell Clonality Analysis by Next-Generation Sequencing: A Structured Approach with the DEPART Algorithm. J. Mol. Diagn. 2023;25(10):729–739.
13. Langerak AW, Groenen PJTA, Brüggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. Leukemia. 2012;26(10):2159–2171.

Online Supplementary Tables 1-14 provided as Excel files only.

Supplementary Table legends

Supplementary Table 1: Summary of clinical features of cases investigated

Dx, diagnosis; DOD, *Dead of Disease*; CR – *Complete Remission*; A(W)D – *Alive (With) Disease*; Lost to FU *Lost to Follow-Up*

Supplementary Table 2: Immunophenotypical features of ASCL and EATL as determined by immunohistochemistry
Pos, positive; neg, negative; n.a. not available

Supplementary Table 3: Gene panel of EATL associated genes
Table of all genes selected with the ir main selecion source from literature, yes

Supplementary Table 4: NGS-based TR clonality analysis in ASCL and EATL
TRB, T-cell receptor beta; TRG, T-cell receptor gamma. * TRB-DJ are unproductive rearrangements. ** Bi-allelic TRG and/or TRB rearrangements are frequently occurring in T-cell lymphomas, however, from the data obtained it can not be excluded that the two clonal rearrangements are present in two different clones (bi-clonal)

Supplementary Table 5: Sequencing metrics

PCT_PF_UQ_READS, Percent Unique Reads Passing Filter; PCT_EXC_DUPE, Percent Excluded as Duplicates; PCT_EXC_OVERLAP, Percent Excluded Due to Overlap; PCT_EXC_OFF_TARGET Percent Excluded as Off-Target.

Supplementary Table 6: Cellularity as fitted by ACE,
Visually verified and most likely fit set. A fit of 1 means that no copy number alterations were present or very limited.

Supplementary Table 7: Mutation overview
NA, no mutation data available, all other denominatoirs as described in Supplementary Table 8

Supplementary Table 8: Description of mutation overview

Supplementary Table 9: Number of mutations detected in the gene panel per sample.

Supplementary Table 10: Frequency of mutations per gene per group.

Supplementary Table 11: Mutation matrix used to create oncoprints

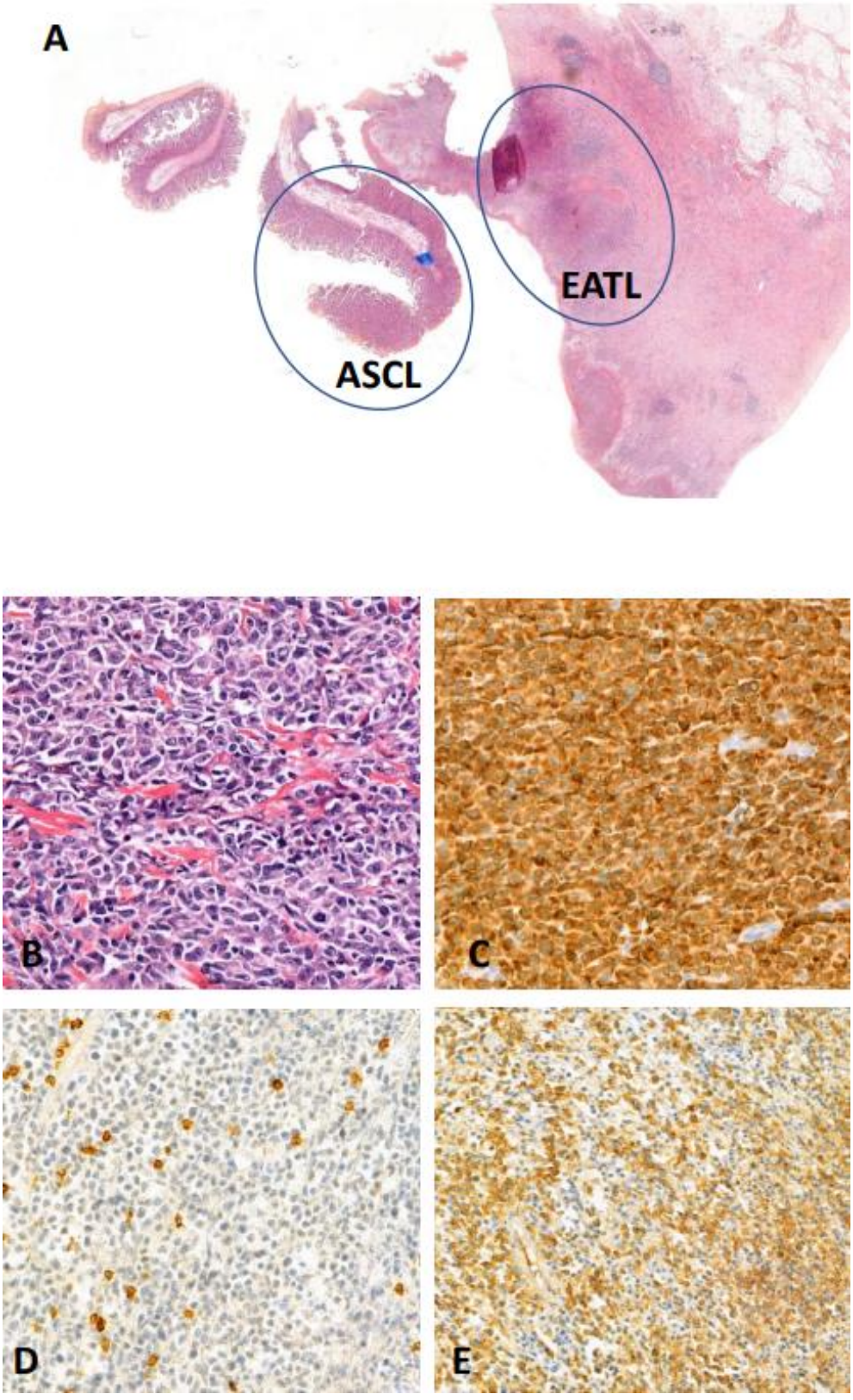
Supplementary Table 12: Copy number aberration matrix per chrosomal arm, per sample, used to create oncoprint
0, normal copy number, -1 loss, 1 gain

Supplementary Table 13: Genetic aberrations in RCD with and without progression to EATL
n.a., no mutation data available

Supplementary Table 14: Summary of clinical features of RCD cases, including treatment.
DOD: dead of disease, A(W)D: alive (with) disease, DOOC: dead of other causes, SCT: stem cell transplantation, n.a.: not available. Marsh: classification of histological changes associated with celiac disease

Supplementary Figures

Supplementary Figure S1:

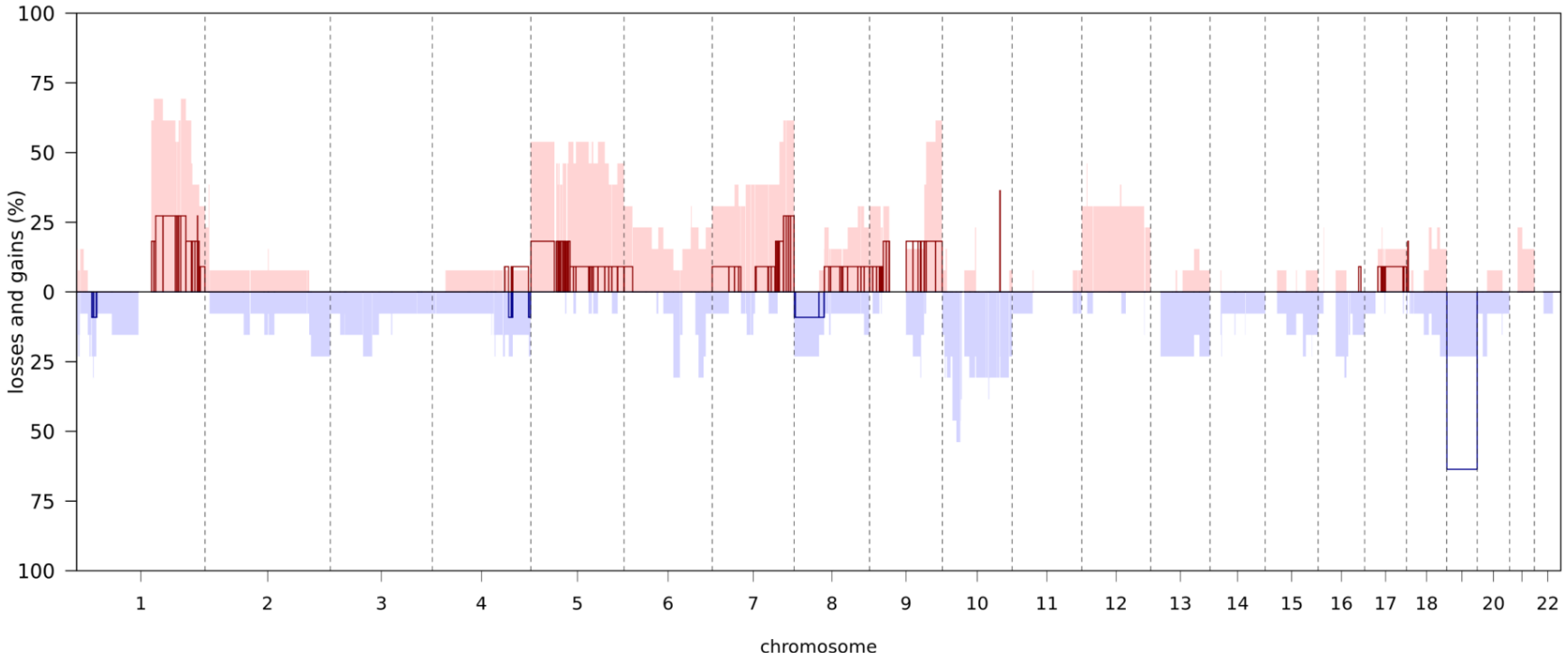


Legend Supplementary Figure S1: Histological features of ASCL and EATL

A: Areas of ASCL are distinct from EATL and macrodissected for separate DNA isolation and sequencing (H&E).

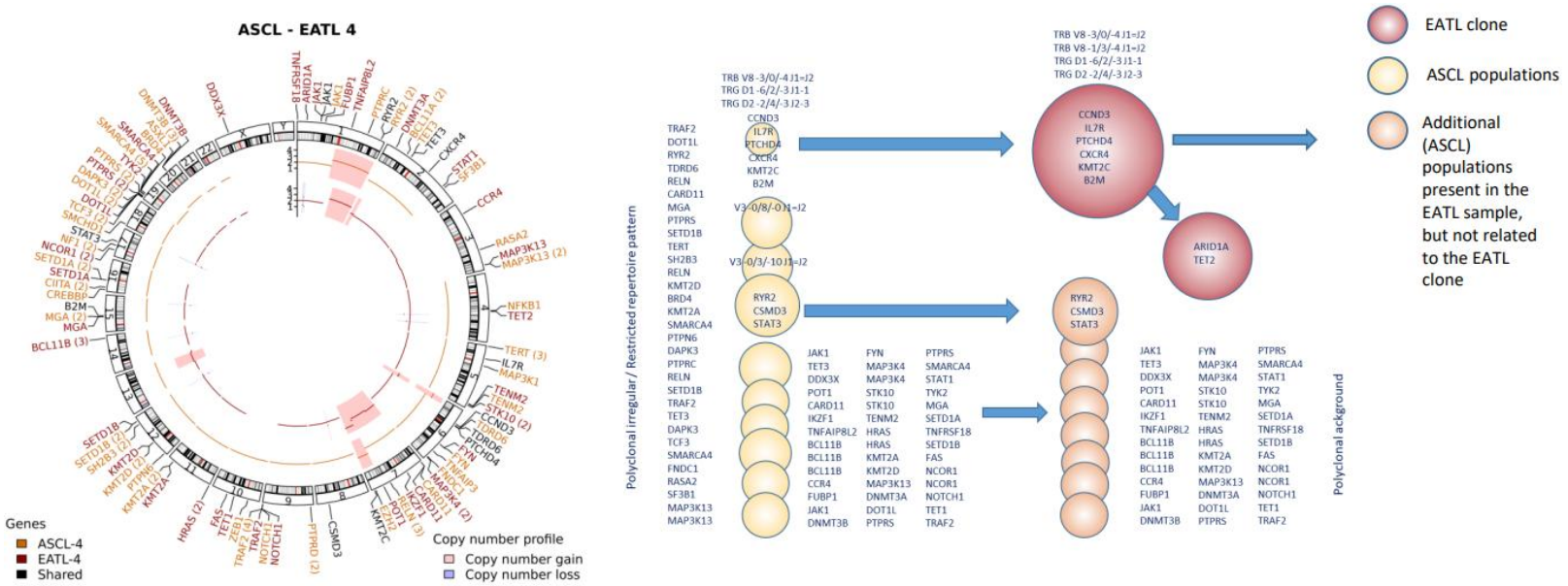
B: a case of EATL shows an extensive admixture of small T-cells (H&E). C: The large tumor cells and the small admixed T-cells are positive for CD3. D: both populations are negative for CD5, supporting the aberrant nature of the small T-cell population. E: the tumor cells are positive for CD30, while the small T-cells are negative.

Supplementary Figure S2:



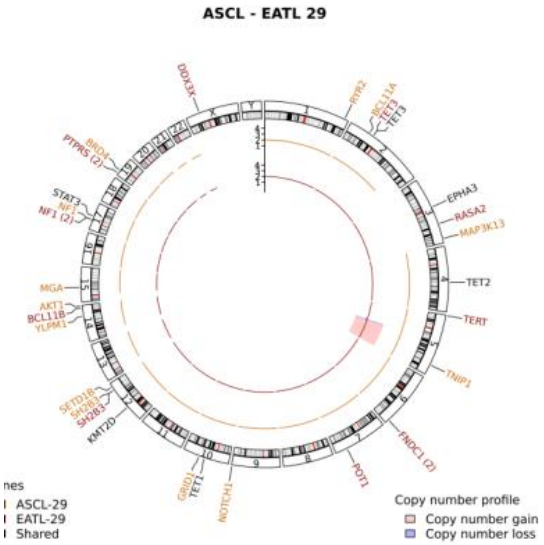
Legend Supplementary Figure S2: Copy number alteration (CNAs) landscape of EATL and ASCL
Comparison plots for CNAs. Indicated are the percentages of cases with gains (positive value, red) and losses (negative value, blue), sorted by chromosomal position (x-axis) of the EATL (filled areas) and ASCL (lines) samples. Although ASCL showed alterations in only few cases, the overall landscape of ASCL and EATL are highly similar, underpinning the notion that ASCL is subject to the same molecular oncogenic events and can be considered related to EATL.

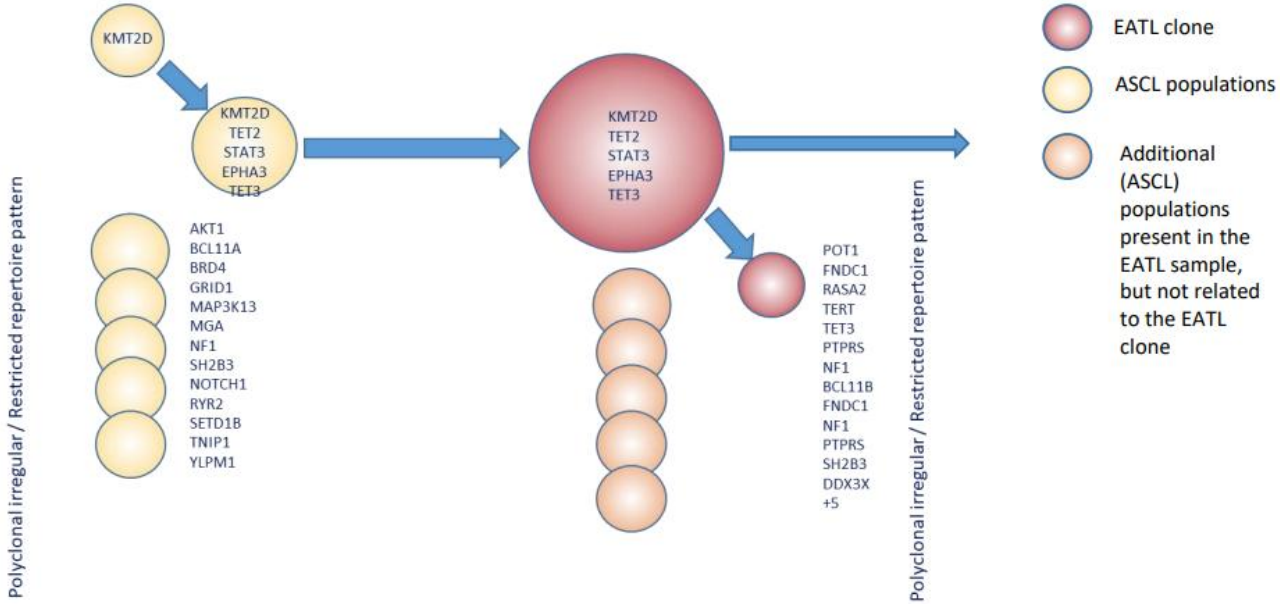
Supplementary Figure S3:



Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
4	EATL	no specific product		clonal (C) plus polyclonal irregular / restricted repertoire background	D1 -6/2/-3 J1-1, 73% and D2 -2/4/-3 J2-3, 17% Gain, multiple copies TRB D1- -6/3/-1 J1)	clonal (C/C)	V8 -3/0/-4 J1=J2, 57%, productive R and V8 -1/3/-4 J1=J2, 23%, unproductive R 7p gain, dus gain van TRG	clonal; bi-allelic TRG, TRB-DJ clonal rearrangements
	ASCL	no specific product		cclonal (C) plus polyclonal irregular / restricted repertoire background	D1 -6/2/-3 J1-1, 35% *note: D2 -2/4/-3 J2-3 (as 4th abundant clonotype in the background)*	clonal (C/C) + PCB	VV3 -0/3/-10 J1=J2, 11%, productive R and V3 -0/8/-0 J1=J2, 8%, unproductive R *note 1: V8 -3/0/-4 J1=J2 as 5th abundant clonotype in the background (2%) note 2: the clonotypes 1-4 of ASCL were not detected in the top*	clonal (biallelic TRG) in addition to (restricted repertoire) background

Integated interpretation:
A mixture of independent T-cell proliferations with subclonal evolution in both the ASCL and EATL samples are observed of which some are unique (TRG clonotypes V3 -0/3/-10 J1=J2 and V3 -0/8/-0 J1=J2 in the ASCL)and others are shared (clonotype TRG V8 -3/0/-4 J1=J2 and TRB D1 -6/2/-3 J1-1,). It should be noted that the EATL precursor clone in the ASCL sample nor the EATL carried a *JAK/STAT* mutation, while other populations in the background did.

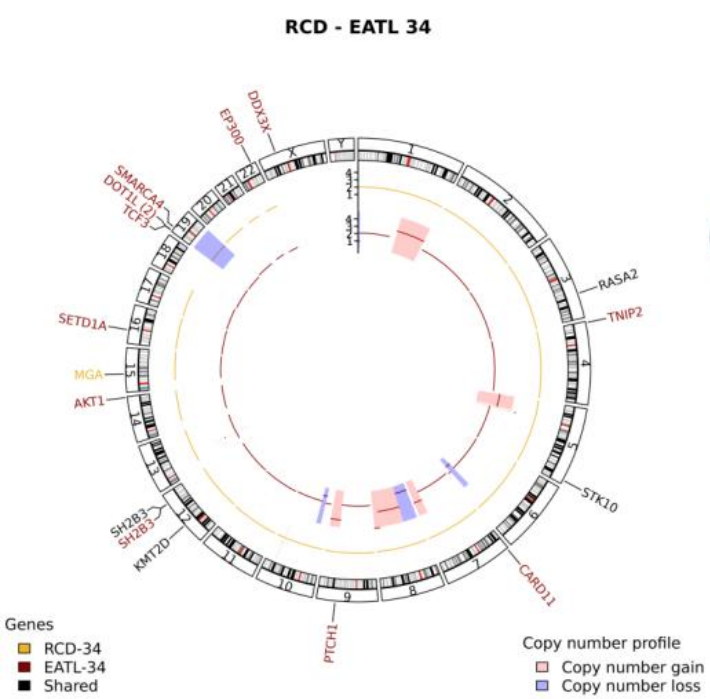


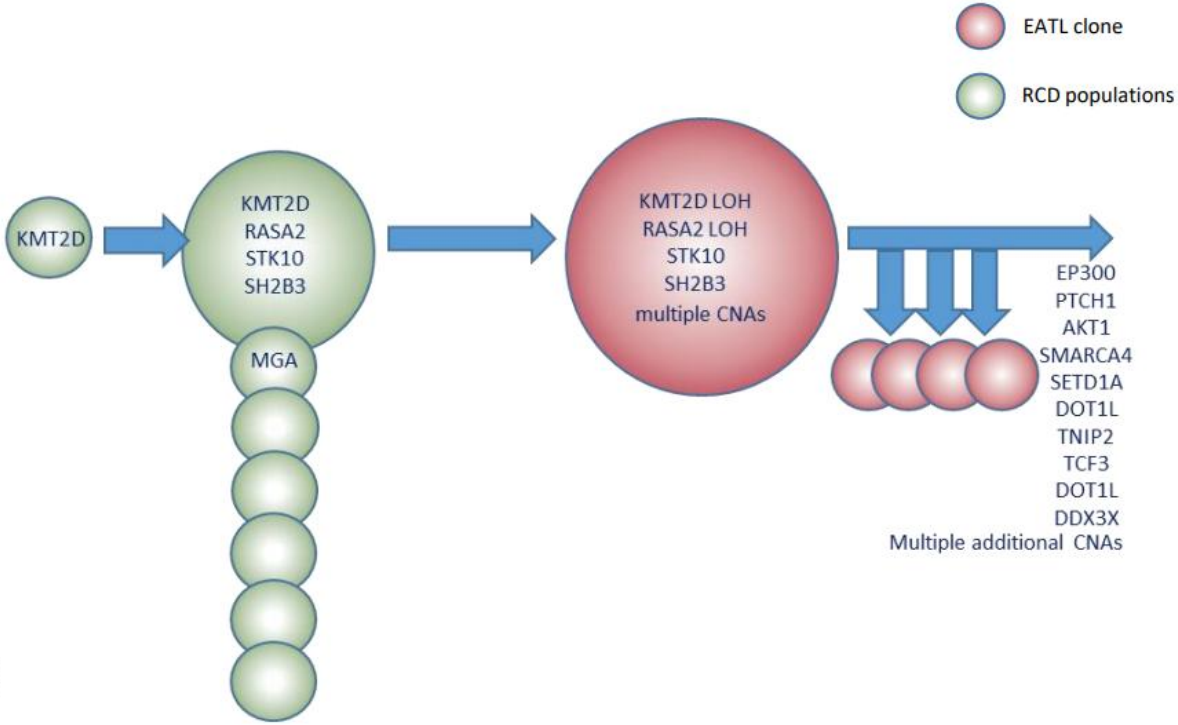


Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
29	EATL	polyclonal irregular / restricted repertoire background		polyclonal irregular		polyclonal irregular		polyclonal (irregular)
	ASCL	no specific product		no specific product		polyclonal irregular / restricted repertoire *note:* *the top 15 clonotypes of ASCL were not detected in the top 15 clonotypes of EATL*		polyclonal (irregular)

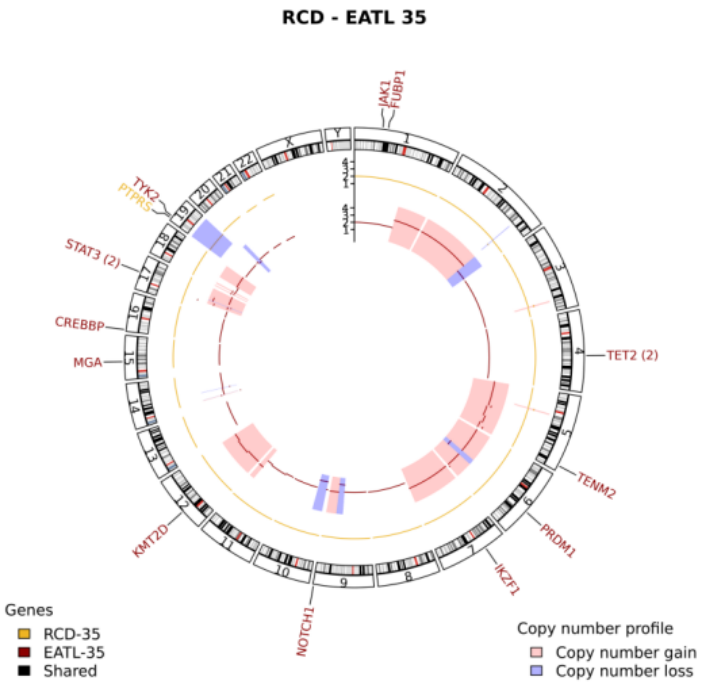
Integated interpretation:

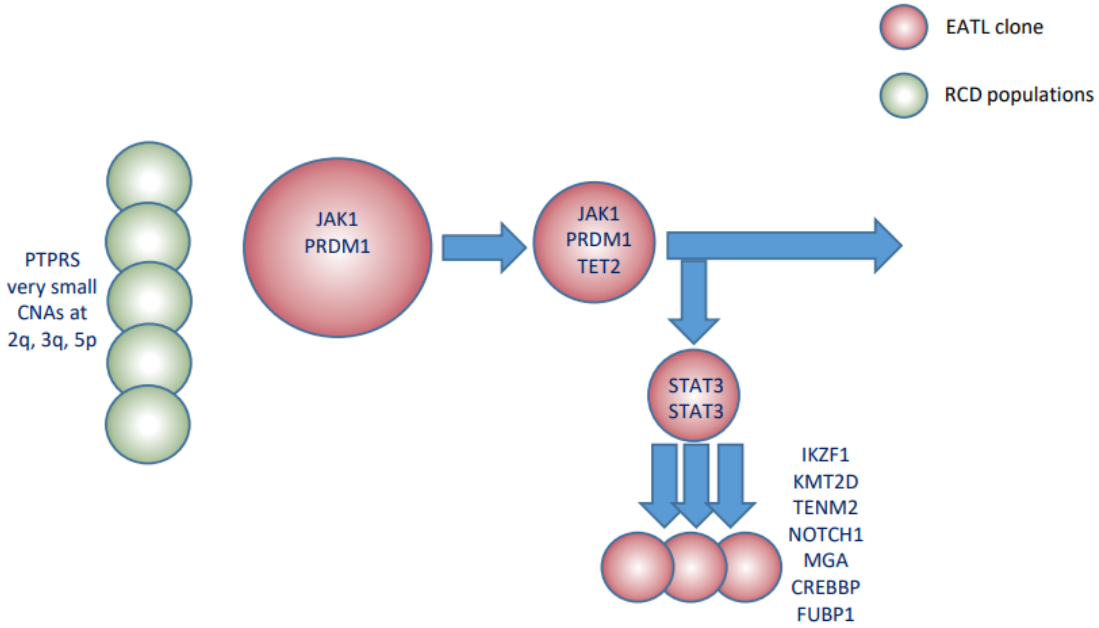
TR rearrangement provides no information. Based on 5 identically shared mutations, within the ASCL sample, a precursor population is identified. Based on VAF differences, KMT2D likely precedes the other mutations in ASCL,. Since all other mutations in ASCL are not propagated to EATL, these do not make part of the EATL precursor population. Similarly, the private mutations in EATL are likely either subclonally acquired in the tumor cells or are present in independent ASCL-like clones in the polyclonal background. Since TR rearrangements do not show shared clones, these models cannot be further differentiated.



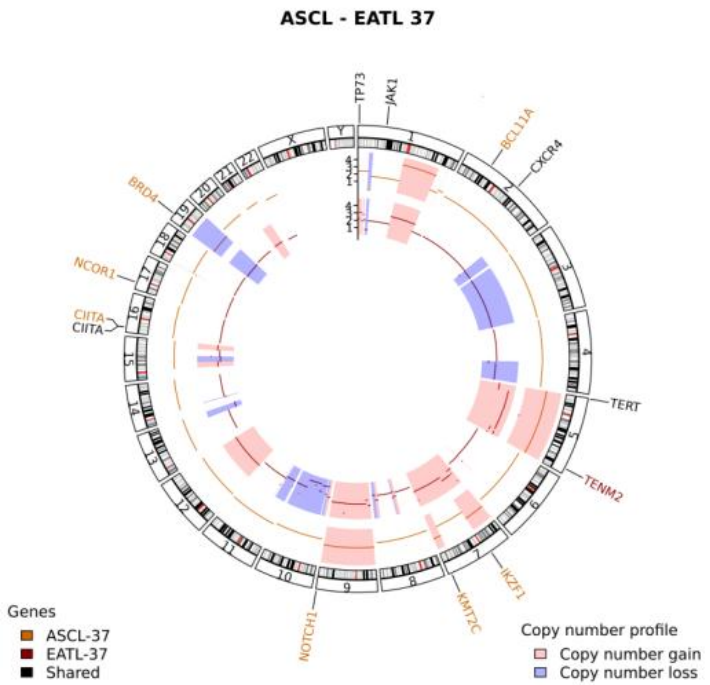


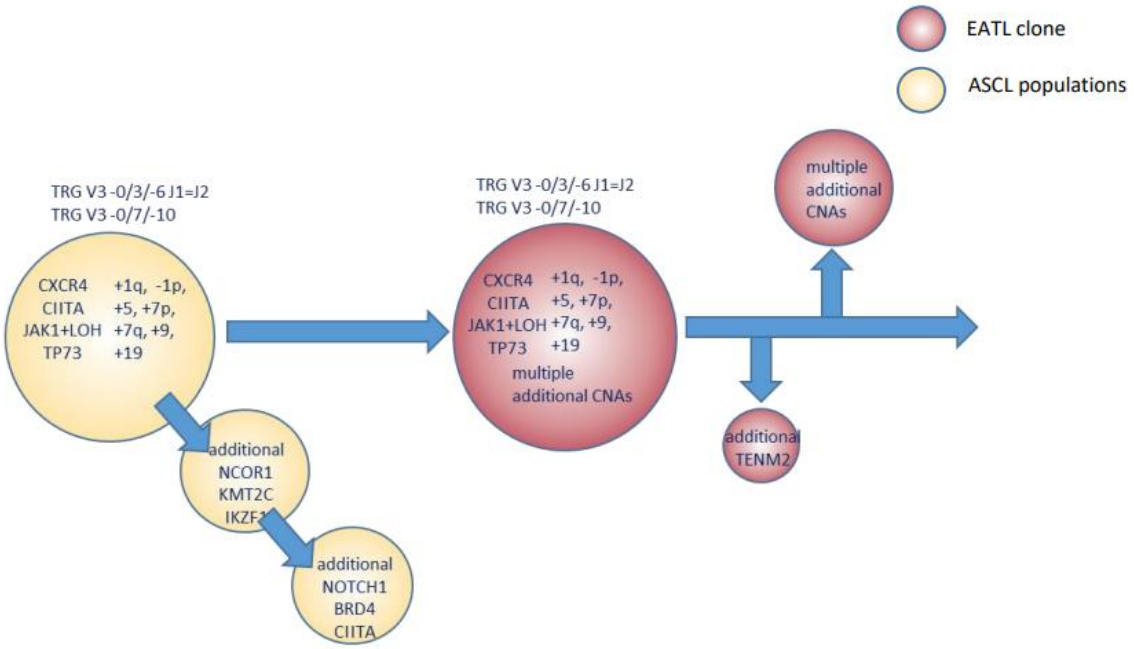
Integated interpretation:
Based on mutation patterns, a set of clonal mutations in the EATL component can be traced back in the preceding RCD sample, supportive of a direct clonal relation. This population is, however already derived from clonal evolution from a population bearing only a KMT2D mutation. It should be noted, that in this case, JAK/STAT activation does not play a major role in early phases of oncogenesis.
No TR data are available for this case.





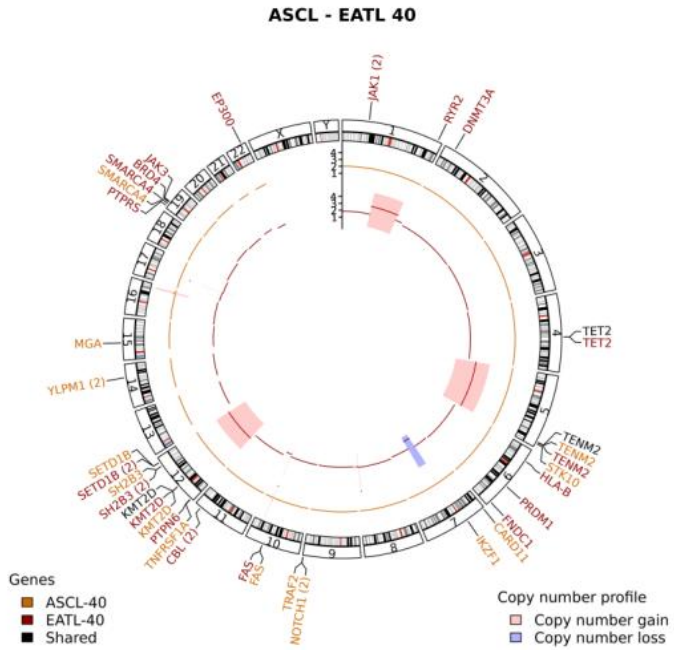
Integated interpretation:
The RCD population in a duodenal biopsy sample shows rather extensive CNAs and only a single PTPRS mutation. These alterations are not propagated to the subsequent EATL at the jejunal site. A direct clonal relation is therefore not likely. No TR data are available for this case. On top of the clonal JAK1 mutation (VAF 0.4) in the complete EATL population, clonal evolution has given rise subclones with two additional STAT3 mutations (based on similar VAF levels of 0.12 likely bi-allelic in the same subclone) and various additional smaller subclones (VAFs 0.031-0.018)

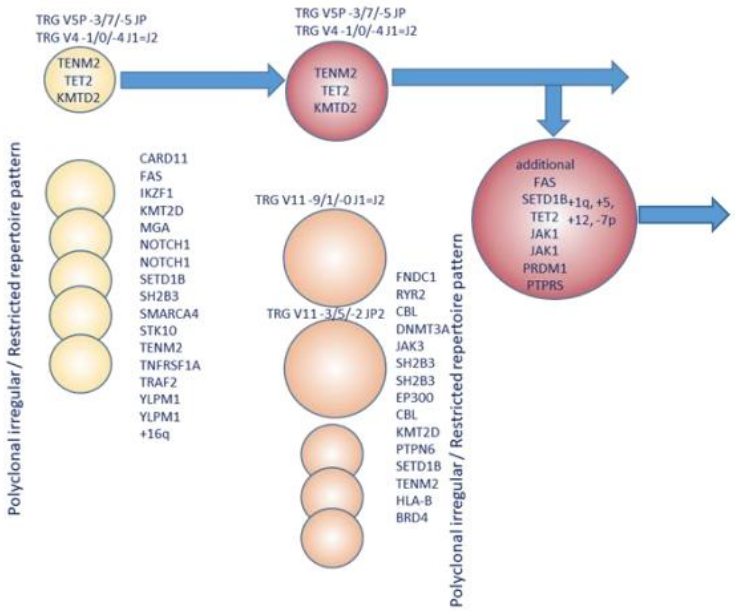


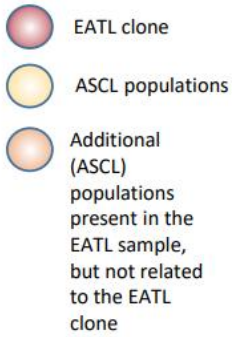


Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
37	EATL	no specific product		no specific product		clonal (C/C)	V3 -0/3/-6 J1=J2, 71%, unproductive R and V3 -0/7/-10 JP2, 24%, unproductive R	clonal (bi-allelic TRG)
	ASCL	no specific product		no data		clonal (C/C)	V3 -0/3/-6 J1=J2, 64%, unproductive R and V3 -0/7/-10 JP2, 18%, unproductive R	clonal (bi-allelic TRG)

Disbalance by +7, no indication for biclonal nature
Clean clonal products

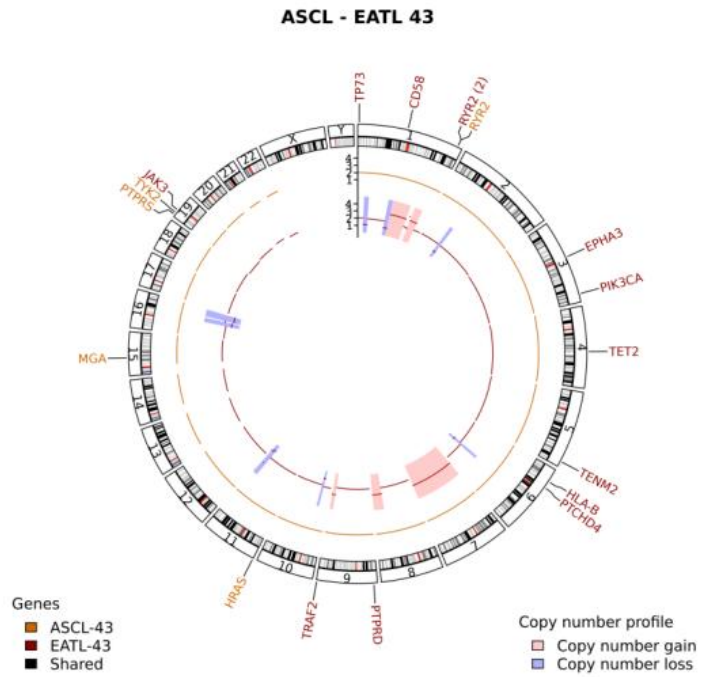


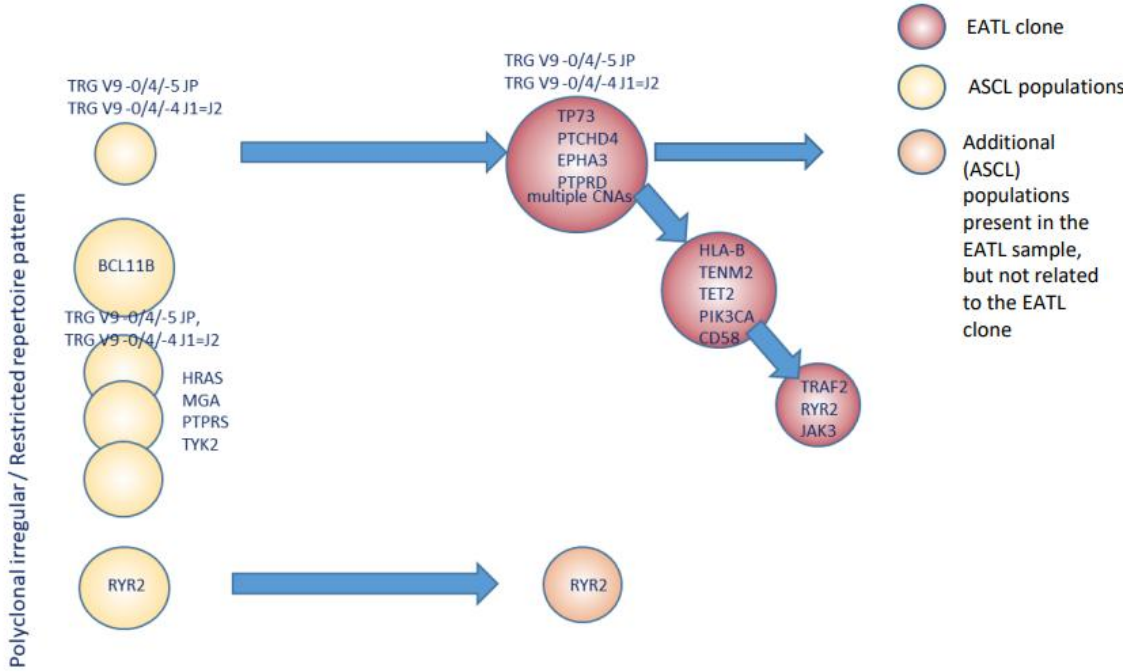




Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
40	EATL	no specific product		no specific product		"multiple dominant products" in a polyclonal irregular background	V5P -3/7/-5 JP, 10%, unproductive R and V4 -1/0/-4 J1=J2, 5%, unproductive R and two other dominant products as 3th and 4th clonotype detected (4,5% and 4% abundance)	bi-clonal or multiple clones (based on TRG) in addition to background
	ASCL	no data for this junction class		no data for this junction class		clonal (C/C) +PCB	V5P -3/7/-5 JP, 16%, unproductive R and V4 -1/0/-4 J1=J2, 5%, unproductive R	clonal (biallelic TRG) in addition to polyclonal background

Integated interpretation:
Next to the ASCL EATL precursor clone with shared bi-allelic unproductive TRG rearrangements (TRG V5P -3/7/-5 JP and TRG V4 -1/0/-4 J1=J2) and TENM2, TET2 and KMTD2 mutations, a polyclonal irregular TR pattern is observed, containiung a set of mutations not found in the EATL sample. In the EATL sample, a restricted repertoire next to the TRG V5P -3/7/-5 JP/TRG V4 -1/0/-4 J1=J2 tumor clone with two populations identified by their TRG rearrangements as well as a large set of mutations with various VAF levels, suggesting various independent clones and/or subclones. In the EATL clone, various CNAs and mutations, including two JAK1 mutations are acquired during further subclonal evolution

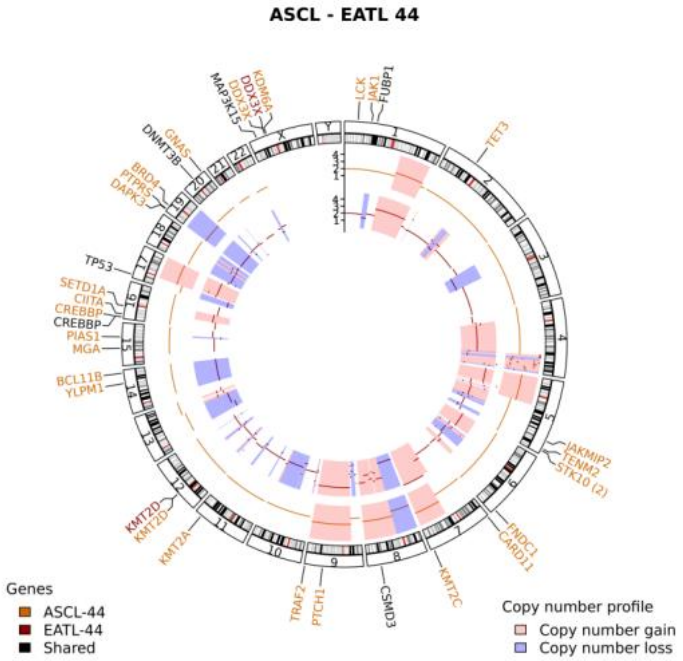


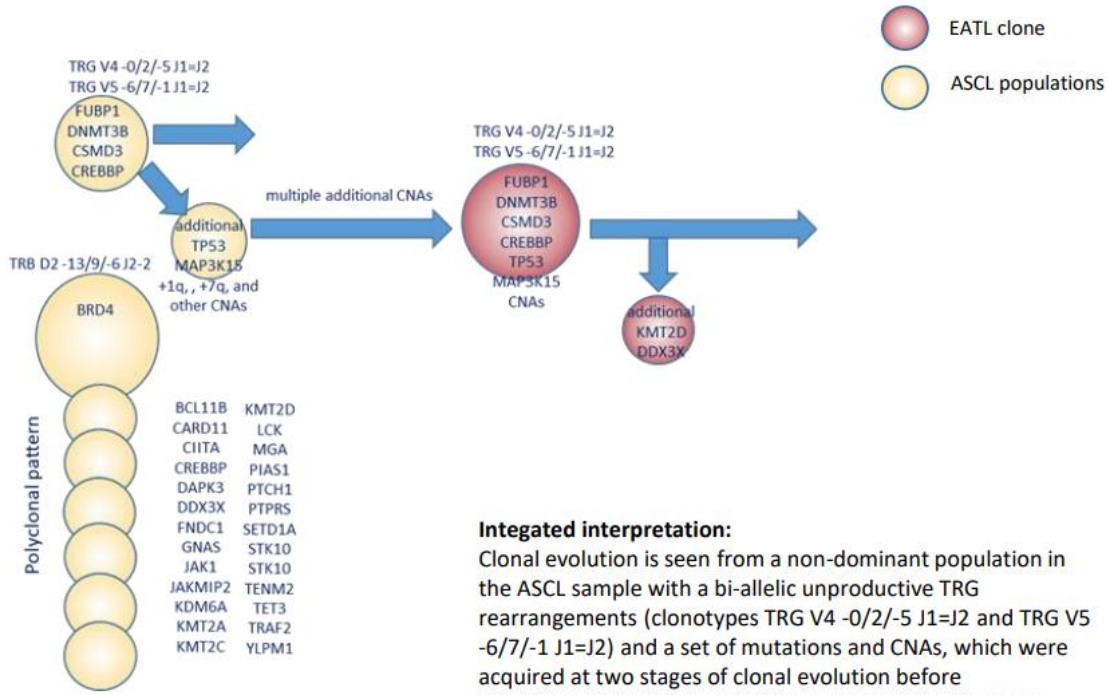


Disbalance by +7

Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
43	EATL	no specific product		no specific product		clonal (C/C)	V9 -0/4/-5 JP, 65%, productive R and V9 -0/4/-4 J1=J2, 16%, unproductive R	clonal (bi-allelic TRG)
	ASCL	no specific product		no specific product		polyclonal irregular	*note 1: V9 -0/4/-5 JP, detected as 5th abundant clonotype in the background and V9 -0/4/-4 J1=J2, detected as 6th abundant clonotype in the background note 2: the clonotypes 1-4 of ASCL were not detected in the top 15 clonotypes of EATL*	polyclonal irregular

Integated interpretation:
ASCL mutations in a polyclonal TR pattern in which the TR is present, but not the EATL founder mutations. Mutations in ASCL are therefore assigned to other resident clones
In EATL, JAK3 is only a late event
CNAs seem clonal events

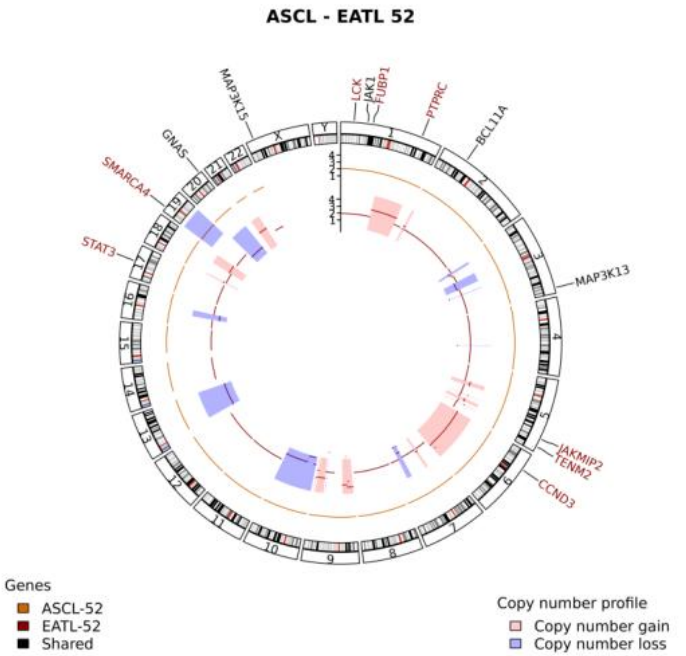


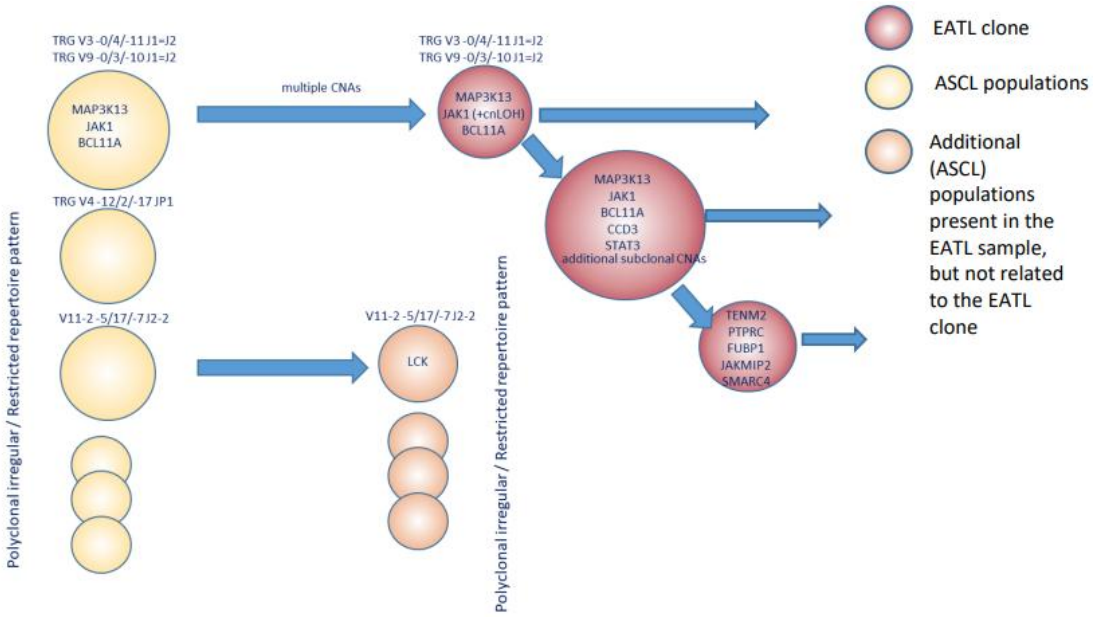


Integated interpretation:
Clonal evolution is seen from a non-dominant population in the ASCL sample with a bi-allelic unproductive TRG rearrangements (clonotypes TRG V4 -0/2/-5 J1=J2 and TRG V5 -6/7/-1 J1=J2) and a set of mutations and CNAs, which were acquired at two stages of clonal evolution before transformation to EATL. Moreover, in the ASCL sample, a dominant T-cell clone is observed with mono-allelic TRB incomplete rearrangement (TRB D2 -13/9/-6 J2-2) that is not part of the EATL clone as well as a polyclonal background bearing a large set of mutations with various VAF levels and including JAK1 and various epigenetic modifiers, that are also not part of the EATL clone.

Disbalance by +7

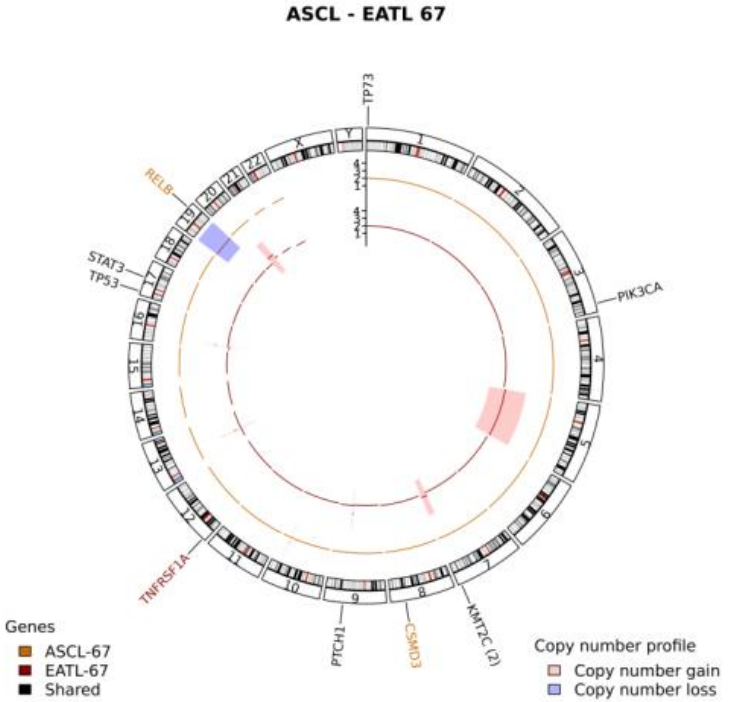
Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
44	EATL	no specific product		no specific product		clonal (C/C) +PCB	V4 -0/2/-5 J1=J2, 77%,unproductive R and V5 -6/7/-1 J1=J2, 20%, unproductive R	clonal (biallelic TRG)
	ASCL	polyclonal irregular		clonal + PCB	D2 -13/9/-6 J2-2, 27%	clonal (C/C) +PCB	V4 -0/2/-5 J1=J2, 17%, unproductive R and V5 -6/7/-1 J1=J2, 35%, unproductive R	clonal (biallelic TRG) in addition to polyclonal background

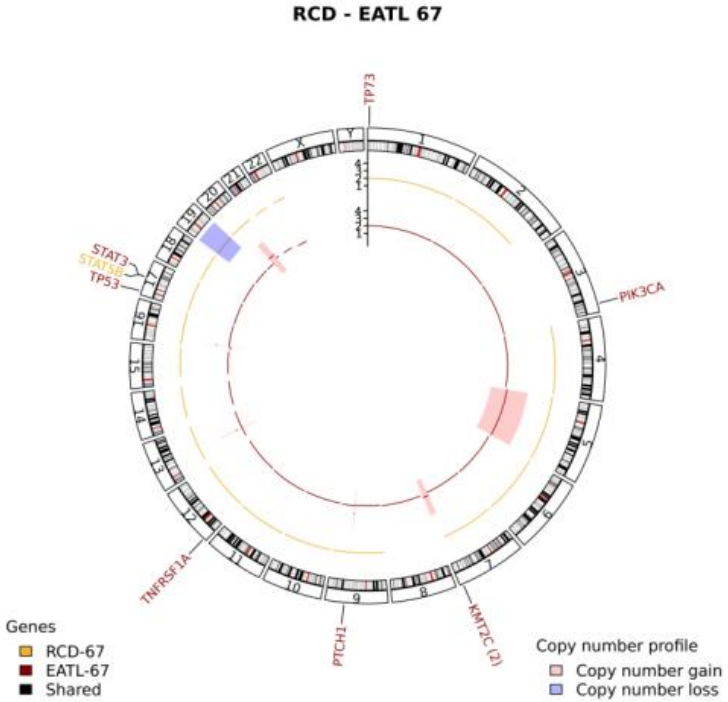




Integated interpretation:
EATL develops from an ASCL precursor and shows subsequent subclonal evolution. In the ASCL population, several independent clones are observed, of which one is also admixed in the EATL population. Morphologically this population cannot be distinguished from normal T-cells in the tumor microenvironment

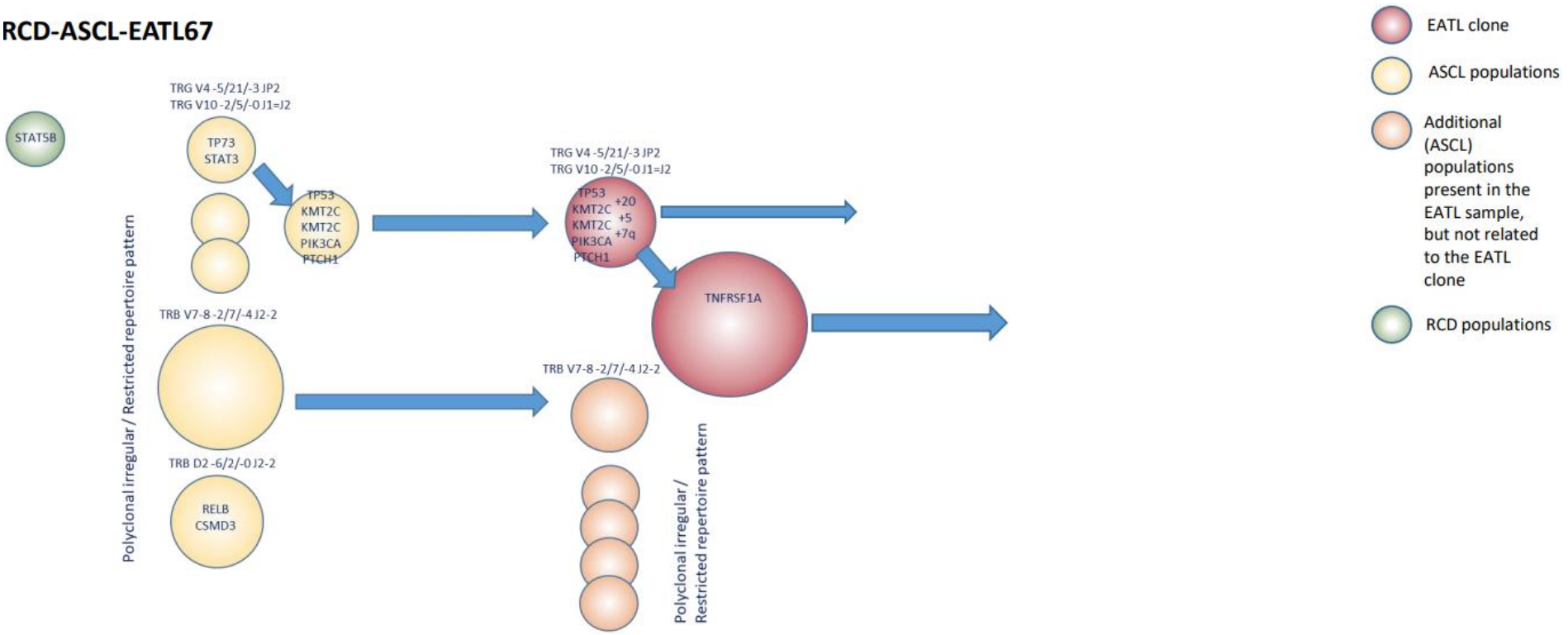
Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
52	EATL	polyclonal irregular	note: V11-2 -5/17/-7 J2-2, productive R, detected as 2nd most dominant clonotype in the background (6%)	nno data for this junction class		clonal (C/C)	V3 -0/4/-11 J1=J2, 43%, productive R and V9 -0/3/-10 J1=J2, 42%, productive R	clonal; bi-allelic TRG or bi-clonal
	ASCL	polyclonal irregular	note: V11-2 -5/17/-7 J2-2, detected as 1st dominant clonotype in the background (6%)	no data for this junction class		multiple products + PCB	V9 -0/3/-10 J1=J2, 12%, productive R and V3 -0/4/-11 J1=J2, 7%, productive R and V4 -12/2/-17 JP1, 7%, unproductive	bi-clonal or multiple clones (based on TRG) in addition to background



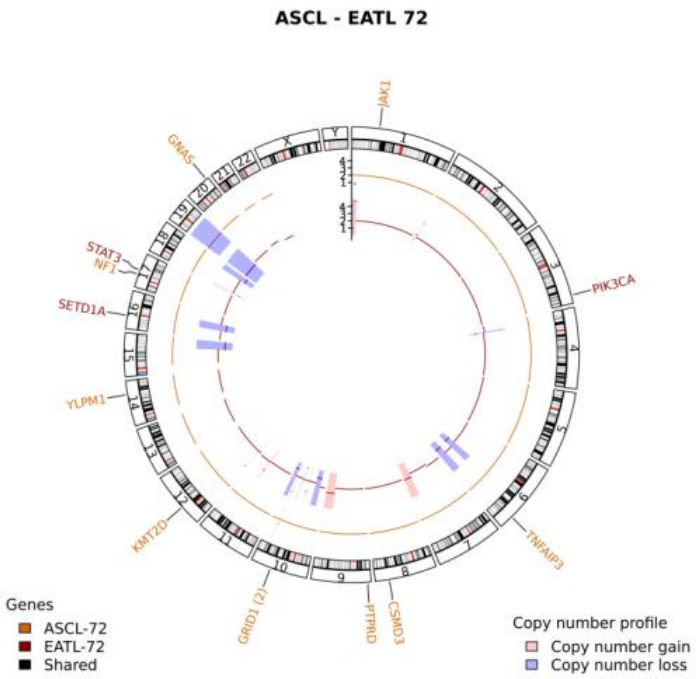


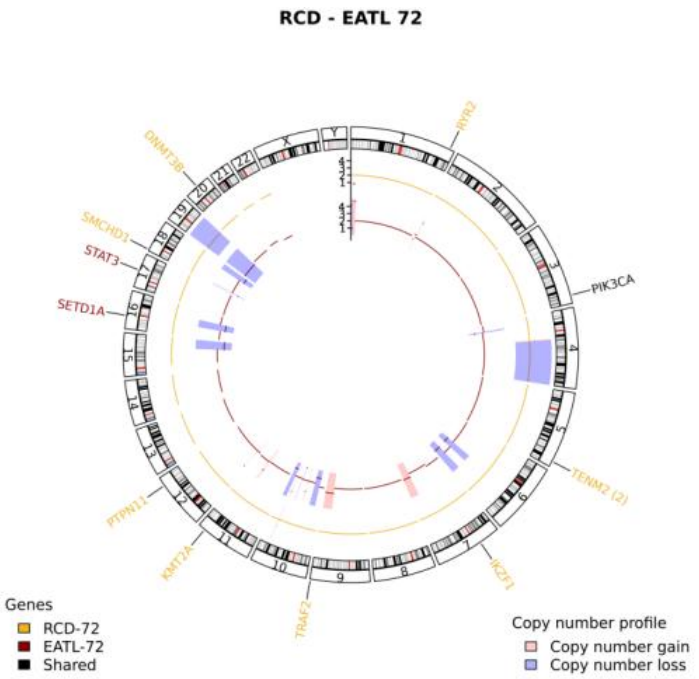
Case no.	EATL/ ASCL	Target score	Rearrangement(s): clonotype, abundance percentage	Target score	Rearrangement(s): clonotype, abundance percentage	Target score	Rearrangement(s): clonotype, abundance percentage	
67	EATL	polyclonal irregular	Note: V7-8 -2/7/-4 J2-2, productive R, is detected in the polyclonal background	polyclonal irregular		clonal (C/C) in polyclonal background	V4 -5/21/-3 JP2, 11%, productive R and V10 -2/5/-0 J1≈J2, 13%, unproductive R	clonal (biallelic TRG) in addition to polyclonal background
	ASCL	clonal	V7-8 -2/7/-4 J2-2, 41%, productive R	clonal	D2 -6/2/-0 J2-2, 37%	clonal (C/C) in polyclonal irregular / restricted repertoire background	V4 -5/21/-3 JP2, 14% productive R and V10 -2/5/-0 J1≈J2, 7%, unproductive R, + polyclonaal irregular, blokkerig beeld, restricted repertoire	clonal (biallelic TRG and TRB) in addition to polyclonal background ,

RCD-ASCL-EATL67



Integated interpretation:
Based on TR and mutation data, EATL develops from an ASCL precursor which is the product of subclonal evolution within an ASCL population. Additional independent populations with other TR clonotypes and various mutations that cannot be assigned to the EATL precursor clone are observed. Of these, One TR-identified clones is also found admixed with the EATL tumor population. It should be noted that this ASCL clone cannot be morphologically distinguished from normal T-cells in the tumor microenvironment.

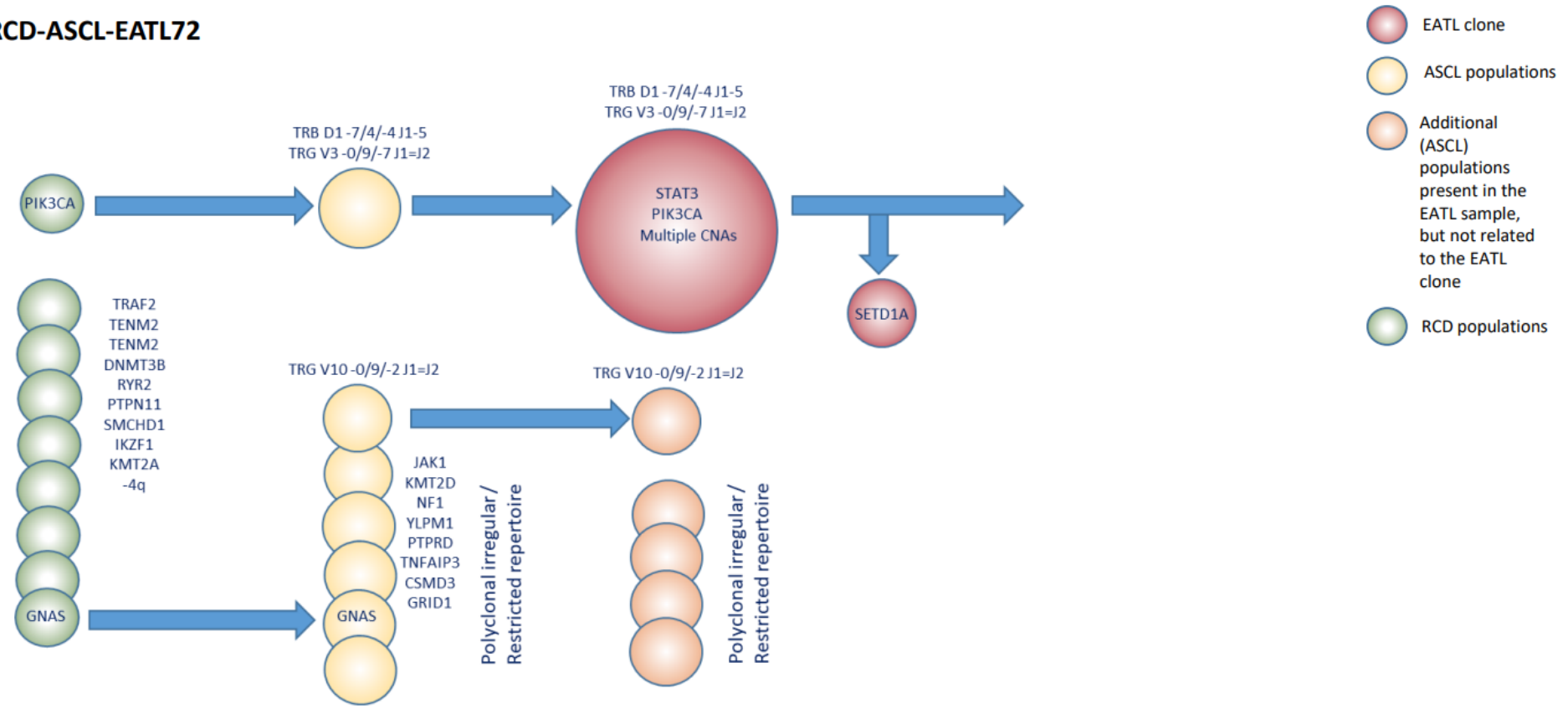




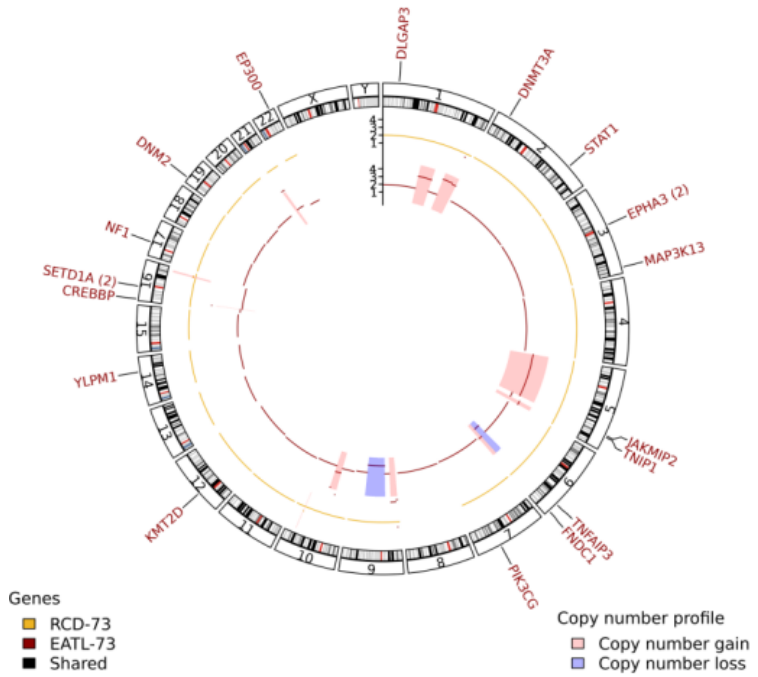
Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
72	EATL	polyclonal irregular / restricted repertoire (*note; low reads*)		clonal + PCB	D1 -7/4/-4 J1-5, 38%	clonal (C/C) +PCB	V3 -0/9/-7 J1=J2, 40%, productive R and V10 -0/9/-2 J1=J2, 21%, unproductive R	clonal (biallelic TRG) in addition to polyclonal background
	ASCL	polyclonal irregular		clonal in polyclonal irregular / restricted repertoire background	D1 -7/4/-4 J1-5, 11%	clonal (C/C) +PCB	V3 -0/9/-7 J1=J2, 13%, productive R and V10 -0/9/-2 J1=J2, 6%, unproductive R	clonal (biallelic TRG) in addition to polyclonal background

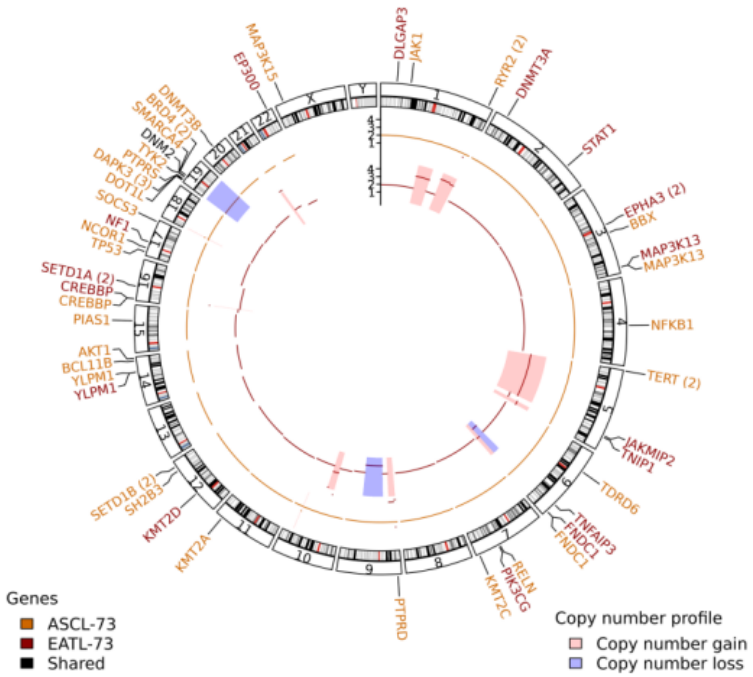
Disbalance without +7

RCD-ASCL-EATL72



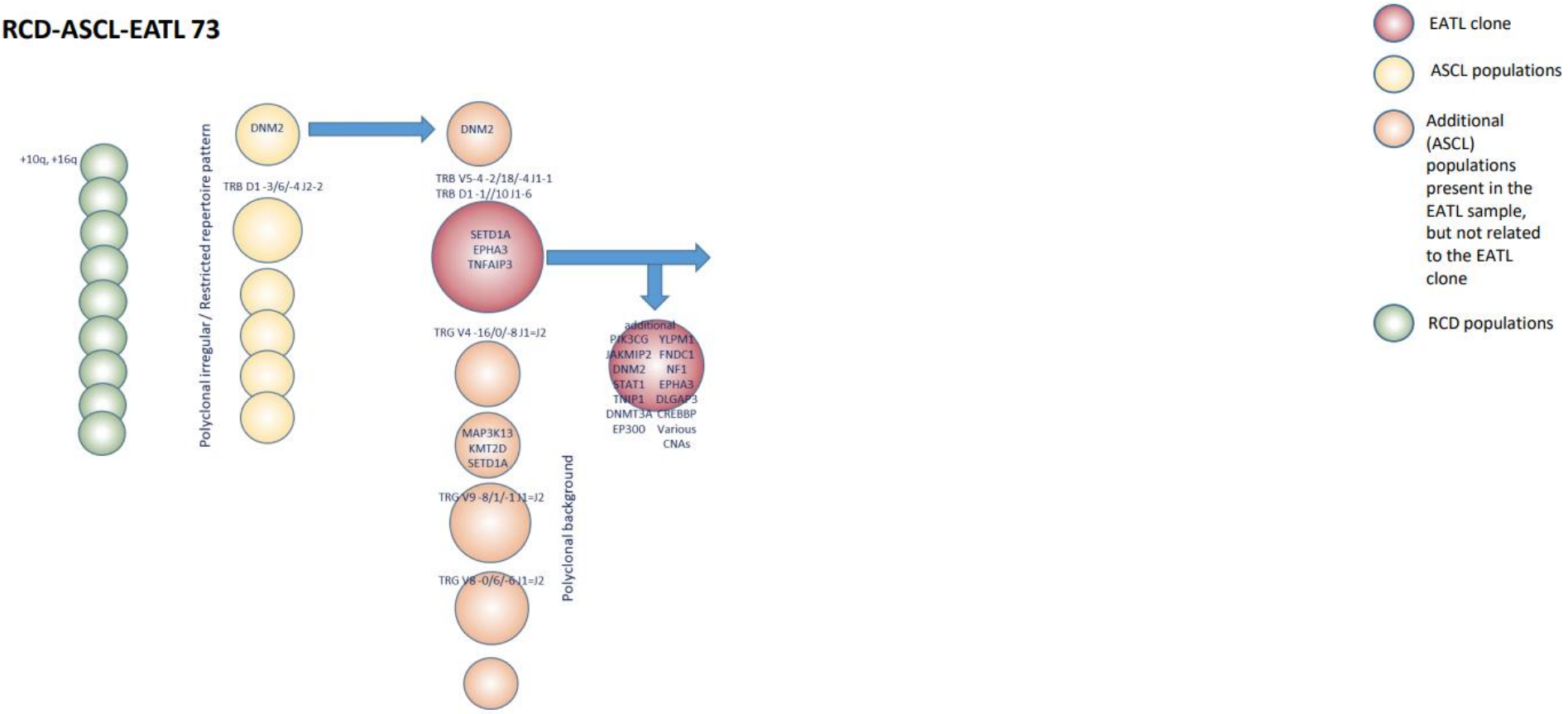
Integated interpretation:
An ASCL clone can be identified as a direct clonal precursor to EATL (TRB D1 -7/4/-4 J1-5/TRG V3 -0/9/-7 J1=J2), which based on an identical PIK3CA mutation can be linked to a population in RCD at the duodenal site. Next to this clonal evolutionary line, a population in ASCL outside the EATL tumor mass as well as intermixed with the EATL tumor cells can be observed (TRG V10 -0/9/-2 J1=J2). Moreover, A second population within RCD in the duodenum characterized by a *GNAS* mutation can be traced to the ASCL population in the jejunum but is not part of the EATL clone.



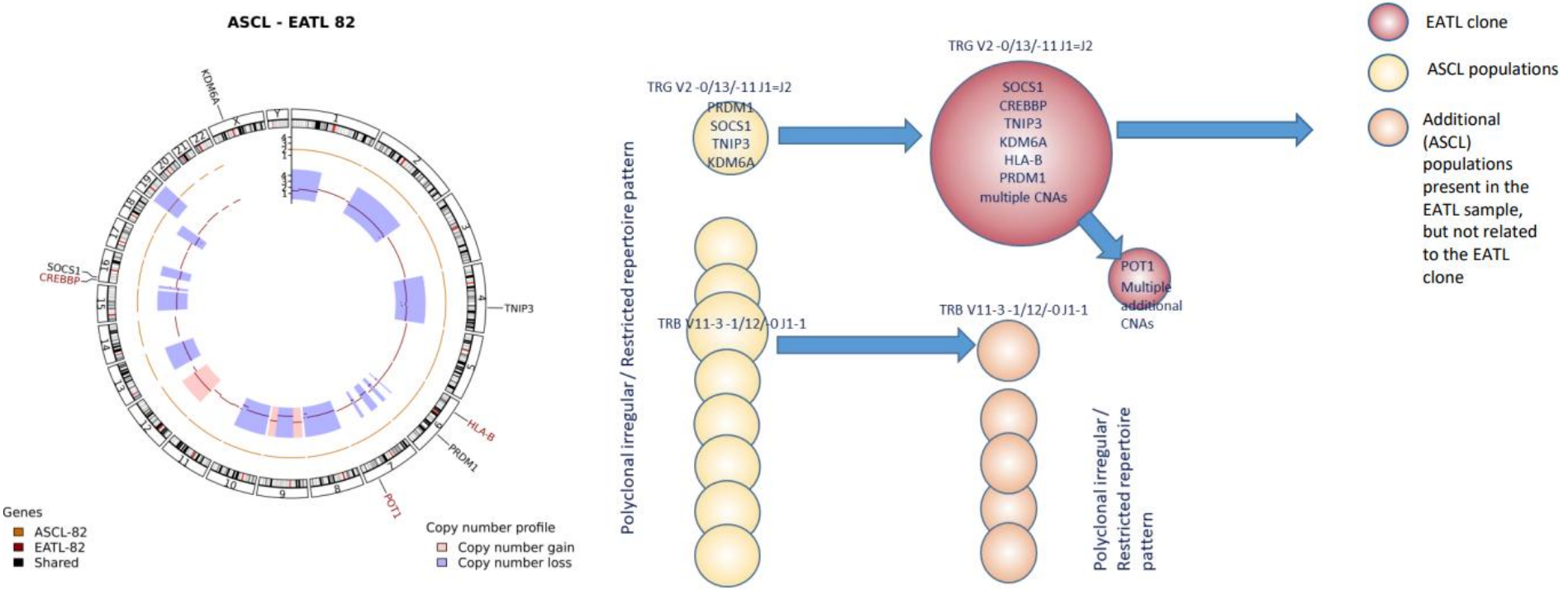


Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
73	EATL	clonal + PCB	V5-4 -2/18/-4 J1-1, 73%, productive R	clonal + PCB	D1 -1/5/-10 J1-6, 48%	multiple products + PCB	V4 -16/0/-8 J1=J2, 18%, unproductive R and V8 -0/6/-6 J1=J2, 9%, unproductive R and V9 -8/1/-1 J1=J2, 7%, unproductive R	clonal (biallelic TRB) + at least bi-clonal for TRG, or 3 clones based in addition to polyclonal background
	ASCL	no specific product		clonal + PCB	D1 -3/6/-4 J2-2, 45% in polyclonal irregular / restricted repertoire background	Polyclonal irr (top 15 clonotypes of ASCL were not detected in the top 15 clonotypes of EATL)		(based on TRB)

RCD-ASCL-EATL 73



Integrated interpretation:
From the data, derived from the EATL DNA sample, based on VAFs two clonal TRB rearrangements can be assigned to the EATL tumor clone, while the various other TR products (three clonal TRG rearrangements and two (one with bi-allelic TRG rearrangement) or three (all with a single TRG rearrangement) T-cell clones should be assigned to the polyclonal background. Of these, none could be traced back in the ASCL sample. A direct ASCL EATL precursor population could therefore not be identified in a population with features of a restricted repertoire. In the corresponding RCD sample, only 2 CNAs were observed. Also these bear no relation to ASCL or EATL clones.



Case no.	EATL/ ASCL	Target score TRBV-TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBV-TRBD-TRBJ	Target score TRBD-TRBJ	Rearrangement(s): clonotype, abundance percentage TRBD-TRBJ	Target score TRGV-TRGJ	Rearrangement(s): clonotype, abundance percentage TRGV-TRGJ	molecular conclusion
82	EATL	polpolyclonal irregular / restricted repertoire	note: V11-3 -1/12/-0 J1-1, detected as 3th abundant detected clonotype in the background (3%)	no data for this junction class		clonal in polyclonal background	V2 -0/13/-11 J1=J2, 37%, productive R	clonal in addition to polyclonal background
	ASCL	polyclonal irregular / restricted repertoire	note: V11-3 -1/12/-0 J1-1 detected as 2nd abundant detected clonotype in the background (2%)	no data for this junction class		polyclonal irregular / restricted repertoire	*note: V2 -0/13/-11 J1=J2, detected as 2nd clonotype in the background (3%)*	polyclonal irregular / restricted repertoire

Legend Supplementary Figure S3: Circos plots, TR rearrangement analysis and predicted models of clonal evolution of 13 paired EATL, ASCL and RCD cases. For all cases for which paired samples of EATL and ASCL with or without RCD samples were available, evolutionary trees were inferred. Mutations and CNAs are depicted as circosplots and TR clonotypes of TRB and TRG assays are depicted. Underlying VAF information is provided in Supplementary tables 9 and 10. In the models, the EATL population is depicted in red, admixed independent T-cell populations that are not part of the EATL clone are indicated in orange. ASCL populations are depicted in yellow and RCD in green. Relative population sizes with the samples are estimated based on TR clonotype abundance and VAF levels of mutations.