Supplementary Materials for:

A Recurrent Dominant-Negative E47 Mutation Causing Agammaglobulinemia and BCR$^-$ B-cells

Bertrand Boisson, Yong-Dong Wang, Amma Bosompem,
Cindy S. Ma, Annick Lim, Tatiana Kochetkov, Stuart G.Tangye,
Jean-Laurent Casanova and Mary Ellen Conley

*Corresponding author. E-mail: mconley@uthsc.edu

This PDF file includes:

DETAILED METHODS

Whole exome sequencing:

Massively parallel sequencing: Genomic DNA (3µg) was extracted from whole blood lysates from Pt. 3 and his parents and sheared with a Covaris S2 Ultrasonicator (Covaris). An adapter-ligated library was prepared with the Paired-End Sample Prep kit V1 (Illumina). Exome capture was performed with the 50Mb SureSelect Human All Exon kit (G3370, Agilent Technologies). Paired-end sequencing was performed on an Illumina Genome Analyzer IIx, generating 100-base reads.

*Sequence alignment, variant calling and annotation***:** Image analyses and base calling were performed using the Illumina Genome Analyzer Pipeline software (GAPipeline version 1.8 or greater) with default parameters. Reads were aligned to a human reference sequence (UCSC assembly hg19, NCBI build 37) and genotypes were called at all positions where there were high quality sequence bases (Phred-like Q25 or greater) at minimum coverage of eight, using Genomics Workbench v5.1 (CLC Bio). For each sample, about 90% of at least 80 million reads were uniquely mapped to the targeted human exon regions, giving an average depth coverage of 120. Under such coverage, about 94% of targeted regions were covered by eight reads or more.

Exome analysis - filtering criteria for variants: To identify the pathogenic mutations, we performed a series of filtering steps. From Pt.3, we first discarded the variants that did not change the amino acid sequence or were not predicted in the splice-site regions. We then excluded the variations that were reported in the dbSNP132 database, that were found in other public databases including the Exome Variants Server and the 1000 Genomes project database, reasoning that the causative variant was uncommon and unlikely to be registered in the SNP database or control samples. Finally the list of variants was filtered against the variants identified from each of the parents. Three heterozygous

variants were identified: TCF3, ANKRD13D, and MSLNL. There were a low number of reads for ANKRD13D, and MSLNL. Only TCF3 was confirmed by Sanger sequencing.

Quantitative RT-PCR:

Total RNA was extracted from EBV-immortalized B cells with Trizol (Invitrogen) and reverse-transcribed (2 μg) with the SuperScript III first-strand synthesis kit (Invitrogen). *E47* and *E12* mRNAs were quantified by RT-qPCR in a TaqMan Gene Expression Assay (E47 specific probe: hs01012686; E12 specific probe hs01016246) with normalization against a beta glucuronidase probe.

Western blots:

Cells were lysed in a buffer containing 40 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5% Triton X-100 supplemented with protease and phosphatase inhibitors (Complete and PhoStop, Roche). Cytoplasmic and nuclear extracts were separated using the NE-PER nuclear and cytoplasmic extraction reagent (ThermoScientific). For immunoprecipitations, 2μg of anti-E47 (554077; BD Biosciences) was added to 1 mg of total protein extract (approximately 5 x 10^7 cells) and incubated overnight at 4 °C. Protein A agarose beads (Sigma-Aldrich) were added to the samples, which were then incubated 1 hour at 4 °C. The beads were then washed three times with lysis buffer, and resuspended in Laemmli buffer. Proteins were separated by SDS-PAGE and transferred to PVDF membranes. The membranes were incubated with antibodies to E47, E2A (sc-763; Santa Cruz) or MyoD (sc-377460; Santa Cruz), and visualized by chemiluminescence using species-specific HRP-linked antibodies (GE Healthcare).

Production of expression vectors:

Expression vectors were produced by cloning cDNAs obtained from GeneCopoeia into CMV driven N-terminal DDK-Myc-flag tagged expression vectors (pCMV6-Origene). Site directed mutagenesis was used to introduce the E555K mutation into the E47 vector using QuikChange II kit (Agilent). Expression vectors were transfected into HEK-293 cells by Lipofectamine LTX (Invitrogen). Nuclear extracts were obtained with NE-PER nuclear and cytoplasmic extraction reagent (ThermoScientific) 19 hours after transfection.

EMSAs:

Nuclear extracts were incubated 15 min at room temperature in 10mM Tris pH7.5, 40mM NaCl, 1mM EDTA, 10% glycerol with 100nM final ^{32}P-labeled probes in presence of poly-dIdC (100ng/µl). DNA-protein complexes were separated on 5% acrylamide gel with 0.5x TBE. The probes were derived from the regulatory regions of murine immunoglobulin heavy chain enhancer (µE5 gatcccagaacacctgcagcag), and murine creatinine kinase (gtcacccccaacacctgctgccta).

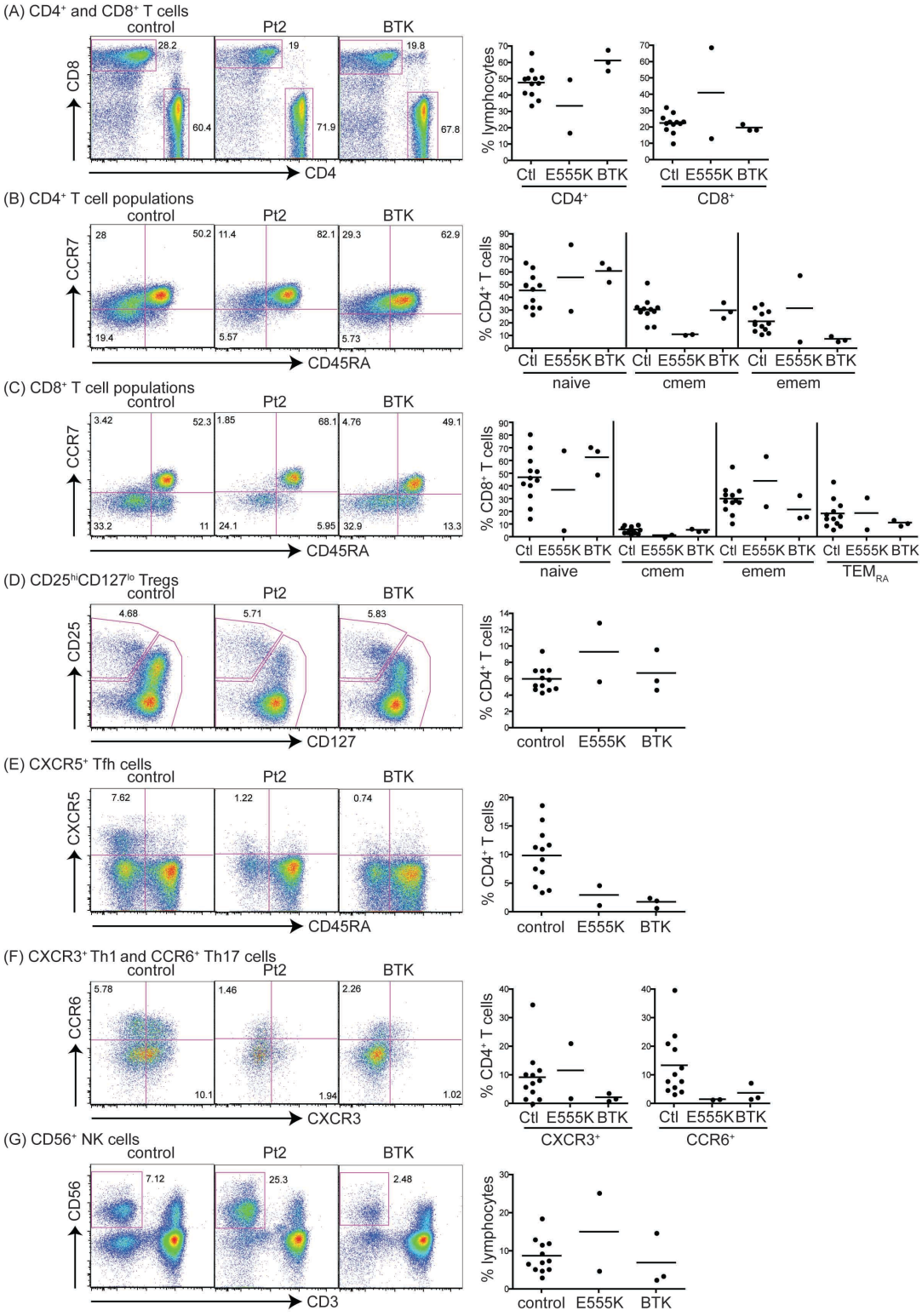
Luciferase assays:

The HEK-293 cells were co-transfected with the plasmid expressing E47, the firefly luciferase pGL4-µE5-µE2(x2) and the *Renilla* luciferase (pRL-TK; Promega) in the presence of Lipofectamine LTX reagent (Invitrogen). After 24 h, luciferase activities were assessed with Dual-Glo Luciferase assay kit (Promega). pGL4-µE5-µE2(x2) were obtained by cloning the DNA fragment (µE5-µE2(x2): 5'-CTAGCagaaCACCTGcagCAGCTGgcaggaaCACCTGcagCAGCTGgcaG in NheI and BglII-digested pGL4 plasmid. The E2A binding sites are underlined.

T Cell Phenotyping:

Frozen peripheral blood lymphocytes were thawed, washed three times and stained with the following mAbs: CD3 Pacific Blue, CD56 Brilliant Violet 605, CD127 Brilliant Violet 650, CXCR3 Brilliant Violet 421, and CCR6 Brilliant Violet 605 (BioLegend); CD45RA PerCP Cy5.5, CD25 PE, and CD45RA FITC (eBioscience); anti-human CD8 PE-Cy7, CD4 APC-Cy7, and CXCR5 Alexa 647 (Becton Dickinson); anti-human CCR7 FITC (R&D Systems). All samples were acquired using a SORP LSRII (Becton Dickinson) and analysed using FlowJo (Tree Star).

Fig. S1


(A) CD4+ and CD8+ T cells
control
Pt2
BTK
CD8
CD4
28.2
60.4
19
71.9
19.8
67.8
% lymphocytes
Ctl E555K BTK
CD4+
CD8+
(B) CD4+ T cell populations
CCR7
CD45RA
28
50.2
19.4
11.4
82.1
5.57
29.3
62.9
5.73
% CD4+ T cells
naive
cmem
emem
(C) CD8+ T cell populations
3.42
52.3
33.2
11
1.85
68.1
24.1
5.95
4.76
49.1
32.9
13.3
% CD8+ T cells
TEMRA
(D) CD25hiCD127lo Tregs
CD25
CD127
4.68
5.71
5.83
control E555K BTK
(E) CXCR5+ Tfh cells
CXCR5
7.62
1.22
0.74
(F) CXCR3+ Th1 and CCR6+ Th17 cells
CCR6
CXCR3
5.78
10.1
1.46
1.94
2.26
1.02
CXCR3+
CCR6+
(G) CD56+ NK cells
CD56
CD3
7.12
25.3
2.48

Legend: **Supplementary Fig. 1**

The T cell phenotype of two of the patients with the E555K mutation in E47, Pt. 2 and Pt. 3, was compared with that of 12 healthy controls and 3 patients with mutations in BTK. The dot plots for Pt. 2 are shown on the left. Pt. 3 has had vaccine associated polio and hepatitis and has a reversed CD4/CD8 ratio as is seen in some patients with immunodeficiency and chronic infection. Total peripheral blood lymphocytes were stained with mAb to CD3, CD4, CD8, CCR7 and CD45RA to determine the proportions of $CD4^+$ and $CD8^+$ cells (**A**), the percent of $CD4^+$ cells that were naïve ($CD3^+CD4^+CD8^-CCR7^+CD45RA^+$), central memory (cmem; $CD3^+CD4^+CD8^-CCR7^+CD45RA^-$), and effector memory (emem; $CD3^+CD4^+CD8^-CCR7^-CD45RA^-$) (**B**); and the percent of $CD8^+$ cells that were naïve ($CD3^+CD4^-CD8^+CCR7^+CD45RA^+$), central memory (cmem; $CD3^+CD4^-CD8^+CCR7^+CD45RA^-$), effector memory (emem; $CD3^+CD4^-CD8^+CCR7^-CD45RA^-$) and revertant CD45RA effector memory (TEM_{RA}; $CD3^+CD4^-CD8^+CCR7^-CD45RA^+$) (**C**). To identify T cell helper populations, cells were stained with mAb to CD4, CD25, CD127, CD45RA, CXCR5, CXCR3 and CCR6. Regulatory T cells (Tregs) were defined as $CD4^+CD25^{hi}CD127^{lo}$ (**D**). From the non-Tregs (ie CD25 low to intermediate population in panel D), the following T helper populations were then defined as T follicular helper cells (Tfh; $CD4^+CD45RA^-CXCR5^+$) (**E**), T helper 1 cells (Th1; $CD4^+CD45RA^-CXCR5^-CXCR3^+CCR6^-$), and T helper 17 cells (Th17; $CD4^+CD45RA^-CXCR5^-CXCR3^-CCR6^+$) (**F**). For NK cell phenotyping, total PBMCs were stained with mAbs to CD3 and CD56 and NK cells were defined as $CD3^-CD56^+$(**G**).