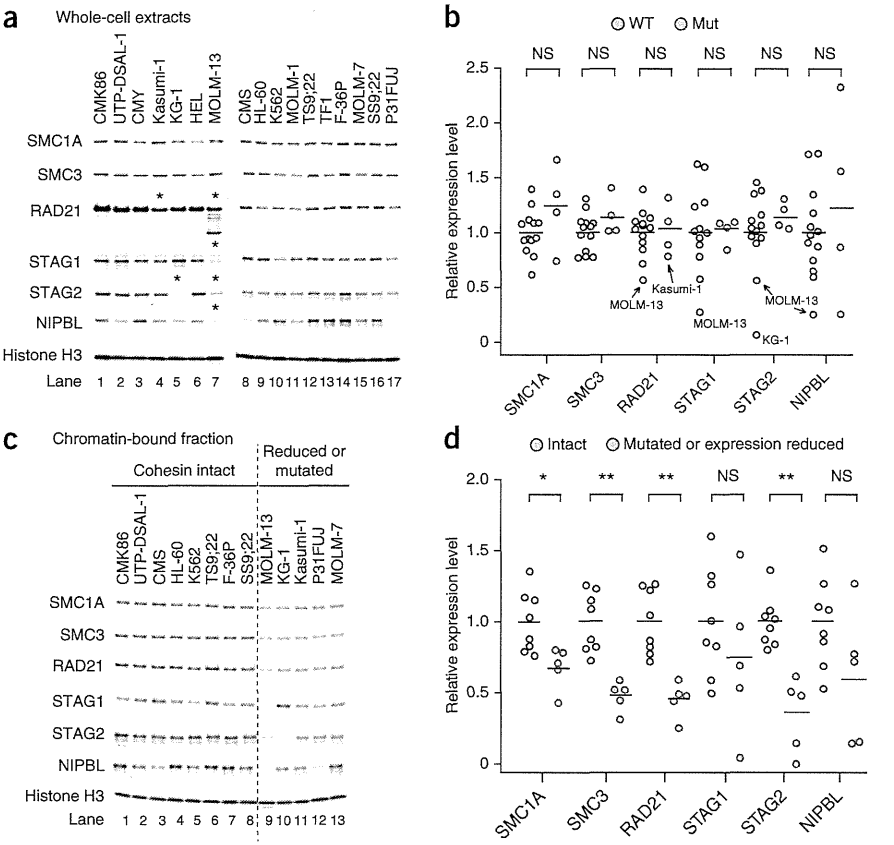

Figure 3 Abnormal cohesin expression and chromatin binding of various cohesin components in myeloid leukemic cell lines. (**a**) Protein blot analysis of the expression of various cohesin components in whole-cell extracts in 17 myeloid leukemia cell lines. Cohesin components showing evaluable reduction in expression are indicated by asterisks, which were reproducible in two independent experiments. (**b**) Expression levels of each cohesin component measured by densitometry after normalization for the mean value across all non-mutated cell lines, with histone H3 signals serving as controls. Evaluably reduced RAD21 expression in Kasumi-1 cells and severely reduced expression of cohesin components in MOLM-13 and KG-1 cells are indicated within the plots. No significant differences (NS) in the expression of the cohesin components were observed between cohesin-mutated and non-mutated cell lines (Mann-Whitney *U* test). Each circle represents a single cell line. (**c**) Protein blot analysis of cohesin components in the chromatin-bound fractions of 13 myeloid leukemia cell lines having intact cohesin (lanes 1–8), cohesin mutations and/or reduced expression of cohesin in whole-cell extracts (lanes 9–13). A representative result of two independent experiments reproducibly showing reduced chromatin-bound cohesin fractions in the cell lines in lanes 9–13 is presented. (**d**) Expression levels of cohesin components in the chromatin-bound fractions measured by densitometry after normalization for the mean value across cell lines with intact cohesin components, with histone H3 signals serving as controls. $*P < 0.05$, $**P < 0.005$ (Mann-Whitney *U* test). Horizontal bars in **b** and **d** indicate the mean values. The densitometric data are presented in **Supplementary Table 14**.

transcriptional changes induced by forced RAD21 expression were generally small[14,16,20]. However, 63 genes reproducibly and significantly showed a more than 1.2-fold increase ($n = 35$) or decrease ($n = 28$) in gene expression ($P < 0.05$), which was validated by quantitative PCR and/or RNA sequencing for 59 of the 63 genes (**Supplementary Fig. 13a–c** and **Supplementary Tables 15** and **16**).

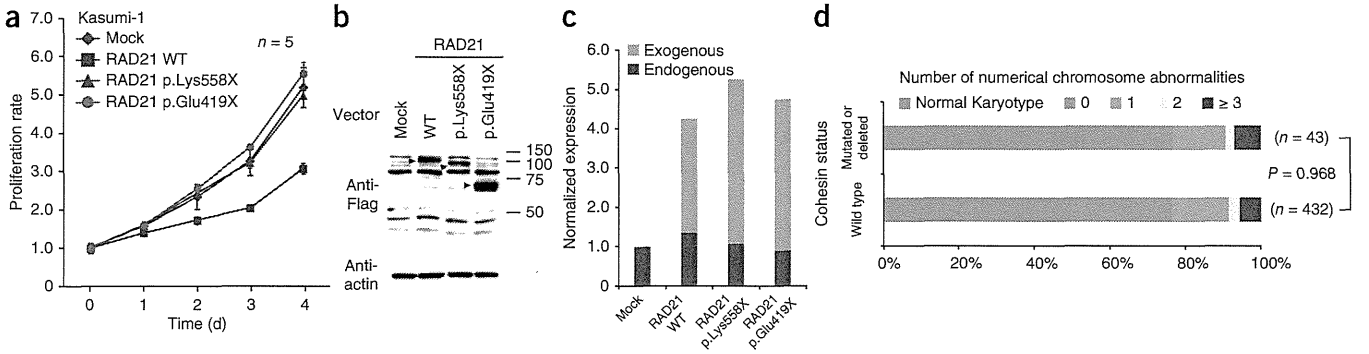

Figure 4 Impact of cohesin mutations on cell proliferation and karyotypes. (**a**) Proliferation of the Kasumi-1 cell line stably transduced with either wild-type RAD21, a truncated allele of *RAD21* (RAD21 p.Lys558X or p.Glu419X) or a mock construct measured by MTT assays ($n = 5$ wells per group). The data are shown as the means ± s.d. of the absorbance at 450 nm relative to the value at day 0. Representative results of three independent experiments are shown. (**b**) Protein blot analysis showing expression of the transduced wild-type and mutant RAD21 alleles. (**c**) Expression of endogenous and exogenous *RAD21* transcripts in Kasumi-1 cells transduced with indicated constructs measured using RNA sequencing by enumerating the corresponding reads. (**d**) The numbers of cases with numerical cytogenetic abnormalities were compared between two groups, those with and those without cohesin mutations or deletions ($P = 0.968$, χ^2). The numbers of numerical chromosome abnormalities are shown at the top. (**e**) Representative metaphases of cell lines with intact (CMS) or abnormal (Kasumi-1 and MOLM-13) cohesin components showing almost normal sister chromatid cohesion. Scale bars, 10 μm.

Mutations in the cohesin complex have recently been reported in a cohort of *de novo* AML and MDS in which four major cohesin components were mutated in 6.0–13.0% of cases[21–25]. Less frequent mutations of cohesin components have been described in other cancers, including *STAG2* mutations in glioblastoma (4/68), melanoma (1/48) and Ewing's sarcoma (1/24)[16]. In primary colon cancer samples, in which impaired cohesion and consequent aneuploidy have been implicated in oncogenesis, mutations in *SMC1A* (4/132), *NIPBL* (4/132), *STAG3* (1/130) and *SMC3* (1/130) have been reported[26]. In contrast, in our cohort of myeloid neoplasms, we found no significant differences in the number of numerical chromosome abnormalities between cohesin-mutated and non-mutated cases, and the 43 cases with cohesin mutations or deletions showed diploid or near-diploid karyotypes, including 23 cases with completely normal karyotypes (**Fig. 4d**). Therefore, in these euploid cases, cohesin-mutated cells were not clonally selected as a result of aneuploidy. Supporting this finding is the observation that expression of *scc1p*, a *RAD21* homolog, at only 13% of its normal level was sufficient for normal cohesion in yeast[27]. Furthermore, Kasumi-1 and MOLM-13 cells showed almost normal cohesion of sister chromatids, even though Kasumi-1 cells have a truncated *RAD21* allele and MOLM-13 cells have substantially reduced expression of multiple cohesin components (**Fig. 4e**).

A growing body of evidence has suggested that cohesin mediates long-range chromosomal *cis* interactions[28] and regulates global gene expression[11,12]. For example, two cohesin subunits, Rad21 and Smc3, have been implicated in the transcriptional regulation of the hematopoietic transcription factor Runx1 in zebrafish[10]. Furthermore, an up to 80% downregulation of *Nipped-B*, a *NIPBL* homolog in *Drosophila*, does not affect chromosomal segregation but does cause impaired regulation of gene expression[20]. We also previously demonstrated that only mild loss (17–28%) of cohesin binding sites within the genome results in deregulated global gene expression[14,18,19]. These observations suggest the possibility that cohesin mutations participate in leukemogenesis through the deregulated expression of genes that are involved in myeloid development and differentiation.

In conclusion, we report frequent mutations in cohesin components that involve a wide variety of myeloid neoplasms. Genetic evidence suggests that aneuploidy may not be the only leukemogenic mechanism, at least *in vivo*, and that deregulated gene expression and/or other mechanisms, such as DNA hypermutability, might also operate in leukemogenesis. Given the integral functions of cohesin for cell viability, genetic defects in cohesin might be potential targets in myeloid neoplasms[14,29].

URLs. dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/; the 1000 Genomes Project, http://www.1000genomes.org/; the UCSC Genome Browser; http://genome.ucsc.edu/cgi-bin/hgGateway/; hg19, http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/; RefSeq genes, http://www.ncbi.nlm.nih.gov/RefSeq/; CNAG/AsCNAR, http://www.genome.umin.jp/; dChip, http://www.dchip.org/; the Integrative Genomics Viewer, http://www.broadinstitute.org/igv/; SIFT, http://sift.jcvi.org/; PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/; Mutation Taster, http://www.mutationtaster.org/.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Whole-exome sequence data have been deposited in the DNA Data Bank of Japan (DDBJ) repository under accession number DRA000433. RNA sequencing data have been deposited in the DDBJ repository under accession number DRA001013. Microarray data have been deposited in the Gene Expression Omnibus under accession number GSE47684.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by Grants-in-Aid from the Ministry of Health, Labor and Welfare of Japan and KAKENHI (23249052, 22134006 and 21790907; S.O.), the Industrial Technology Research Grant Program from the New Energy and Industrial Technology Development Organization (NEDO; S.O.) (08C46598a), NHRI-EX100-10003NI Taiwan (L.-Y.S.), the project for development of innovative research on cancer therapies (p-direct; S.O.) and the Japan Society for the Promotion of Science through the Funding Program for World-Leading Innovative R&D on Science and Technology, initiated by the Council for Science and Technology Policy (CSTP; S.O.). We thank Y. Hayashi (Gunma Children's Medical Centre), R.C. Mulligan (Harvard Medical School), S. Sugano (The University of Tokyo), M. Onodera (National Center for Child Health and Development, Japan) and L. Ström (Karolinska Institute) for providing materials. We thank Y. Yamazaki for cell sorting. We also thank Y. Mori, M. Nakamura, N. Mizota and S. Ichimura for their technical assistance and M. Ueda for encouragement.

AUTHOR CONTRIBUTIONS

A.K., Y.N., K.Y., A.S.-O., Y. Sato and M.S. processed and analyzed genetic materials and performed sequencing and SNP array analysis. Y. Shiraishi, Y.O., R.N., A.S.-O., H.T., T.S., K.C., M.N. and S. Miyano performed bioinformatics analyses of the sequencing data. L.-Y.S. performed pyrosequencing analysis, and A.N. and S.I. performed digital PCR. G.N. and H.A. performed methylation analysis. M.M., M.B. and K.S. performed studies on protein expression of cohesin components. A.K., M.S., T.Y., R.Y., M.O. and H.N. were involved in the functional studies. A.K. and A.S.-O. performed expression microarray experiments and their analyses. L.-Y.S., D.N., T.A., C.H., F.N., W.-K.H., T.H., H.P.K., T.N., H.M., S. Miyawaki, M.S.-Y., K.I., N.O. and S.C. collected specimens and were involved in project planning. A.K., L.-Y.S., M.M., A.S.-O. and S.O. generated figures and tables. S.O. led the entire project, and A.K. and S.O. wrote the manuscript. All authors participated in the discussion and interpretation of the data.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Bejar, R., Levine, R. & Ebert, B.L. Unraveling the molecular pathophysiology of myelodysplastic syndromes. *J. Clin. Oncol.* **29**, 504–515 (2011).
2. Marcucci, G., Haferlach, T. & Dohner, H. Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. *J. Clin. Oncol.* **29**, 475–486 (2011).
3. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
4. Gruber, S., Haering, C.H. & Nasmyth, K. Chromosomal cohesin forms a ring. *Cell* **112**, 765–777 (2003).
5. Nasmyth, K. & Haering, C.H. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* **43**, 525–558 (2009).
6. Ström, L. *et al.* Postreplicative formation of cohesion is required for repair and induced by a single DNA break. *Science* **317**, 242–245 (2007).
7. Watrin, E. & Peters, J.M. The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells. *EMBO J.* **28**, 2625–2635 (2009).
8. Dorsett, D. Cohesin, gene expression and development: lessons from *Drosophila*. *Chromosome Res.* **17**, 185–200 (2009).
9. Dorsett, D. *et al.* Effects of sister chromatid cohesion proteins on cut gene expression during wing development in *Drosophila*. *Development* **132**, 4743–4753 (2005).
10. Horsfield, J.A. *et al.* Cohesin-dependent regulation of Runx genes. *Development* **134**, 2639–2649 (2007).
11. Parelho, V. *et al.* Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
12. Wendt, K.S. *et al.* Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801 (2008).
13. Bose, T. & Gerton, J.L. Cohesinopathies, gene expression, and chromatin organization. *J. Cell Biol.* **189**, 201–210 (2010).
14. Deardorff, M.A. *et al.* HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle. *Nature* **489**, 313–317 (2012).
15. Deardorff, M.A. *et al.* RAD21 mutations cause a human cohesinopathy. *Am. J. Hum. Genet.* **90**, 1014–1027 (2012).

16. Solomon, D.A. *et al.* Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science* **333**, 1039–1043 (2011).
17. Beckouët, F. *et al.* An Smc3 acetylation cycle is essential for establishment of sister chromatid cohesion. *Mol. Cell* **39**, 689–699 (2010).
18. Liu, J. *et al.* Transcriptional dysregulation in NIPBL and cohesin mutant human cells. *PLoS Biol.* **7**, e1000119 (2009).
19. Liu, J. *et al.* Genome-wide DNA methylation analysis in cohesin mutant human cell lines. *Nucleic Acids Res.* **38**, 5657–5671 (2010).
20. Schaaf, C.A. *et al.* Regulation of the *Drosophila* enhancer of split and invected-engrailed gene complexes by sister chromatid cohesion proteins. *PLoS ONE* **4**, e6202 (2009).
21. Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
22. Walter, M.J. *et al.* Clonal architecture of secondary acute myeloid leukemia. *N. Engl. J. Med.* **366**, 1090–1098 (2012).
23. Welch, J.S. *et al.* The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264–278 (2012).
24. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult *de novo* acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
25. Walter, M.J. *et al.* Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia* **27**, 12785–1282 (2013).
26. Barber, T.D. *et al.* Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *Proc. Natl. Acad. Sci. USA* **105**, 3443–3448 (2008).
27. Heidinger-Pauli, J.M., Mert, O., Davenport, C., Guacci, V. & Koshland, D. Systematic reduction of cohesin differentially affects chromosome segregation, condensation, and DNA repair. *Curr. Biol.* **20**, 957–963 (2010).
28. Hadjur, S. *et al.* Cohesins form chromosomal *cis*-interactions at the developmentally regulated IFNG locus. *Nature* **460**, 410–413 (2009).
29. Chan, D.A. & Giaccia, A.J. Harnessing synthetic lethal interactions in anticancer drug discovery. *Nat. Rev. Drug Discov.* **10**, 351–364 (2011).

ONLINE METHODS

Patients and samples. Twenty-nine cases analyzed by whole-exome sequencing were described previously[3]. Anonymized genomic DNA from an additional 581 patients with different myeloid neoplasms were collected from collaborating institutes and used for the analyses described below. All the analyses were performed after written informed consent was obtained. This study was approved by the ethics boards of the University of Tokyo, University Hospital Mannheim, University of Tsukuba, the Munich Leukemia Laboratory, Showa University, Tokyo Metropolitan Ohtsuka Hospital and Chang Gung Memorial Hospital.

Cell lines. The CMS, CMY, UTP-DSAL-1, MOLM-1, MOLM-7, HEL, SS9;22 and TS9;22 cell lines were provided by Y. Hayashi. 293gp and 293gpg cells were provided by R.C. Mulligan. P31FUJ and CMK-86 cells were purchased from the Health Science Research Resources Bank (Osaka, Japan). 293T, KG-1, K562 and F-36P cells were obtained from RIKEN BioResource Center Cell Bank (Tsukuba, Japan), and Kasumi-1, HL-60, MOLM-13 and TF-1 cells were from the American Type Culture Collection. Chromosome spreads were performed for the CMS, Kasumi-1 and MOLM-13 cell lines as previously described[14], except that cells were treated with colcemid (100 μg/ml) and hypotonically swollen in 75 mM KCl for 20 min.

Whole-exome sequencing. The whole-exome sequencing of the 29 paired samples of myelodysplasia was previously described[3], through which we identified a total of 497 candidate single-nucleotide variants and insertions/deletions (indels), of which 268 and 167 were determined by Sanger sequencing as true positives and negatives, respectively, with 62 mutations unconfirmed. In the present study, we updated the list of somatic mutations by rigorously validating the remaining 62 unconfirmed mutations by Sanger sequencing and also by deep sequencing (**Supplementary Table 1**).

Mutation analysis of cohesin components. In total, 534 tumor DNA samples from a variety of myeloid neoplasms were analyzed for possible mutations in nine components of the cohesin complex, *STAG1*, *STAG2*, *SMC1A*, *SMC3*, *RAD21*, *PDS5B*, *ESCO1*, *ESCO2* and *NIPBL*, using high-throughput sequencing of pooled exons amplified from pooled genomic DNA samples. In an additional 47 samples, mutations in *STAG2*, *RAD21*, *SMC1A* and *SMC3* were examined by deep sequencing after enrichment for these targets using a SureSelect custom kit (Agilent) designed to capture all of the coding exons from the target genes, performed as previously described with minor modifications in the algorithm for mutation call[30].

For pooled-DNA sequencing, all target exons (n = 232) encompassing 89,323 nucleotides were PCR amplified using a set of primers having common NotI adaptor sequences on their 5′ ends, digested with NotI, ligated using T4 ligase and sonicated to approximately 200-bp fragments using an ultrasonicator (Covaris); these fragments were used for the generation of sequencing libraries according to a modified pair-end protocol from Illumina. The libraries were then sequenced using HiSeq 2000 (Illumina) with a standard 100-bp paired end–reads protocol. On average, 99.5% of the target bases were analyzed at the depth of 12,000 per pool or 1,000 per sample. Data processing and variant calling were performed as previously described[3] with minor modifications. First, each read from a given DNA pool was aligned to the set of target sequences using BLAT[31] with the -fine option. The mapping information in a .psl format was transformed into a .sam format using the my_psl2sam script, which was further converted into the .bam format using SAMtools[32]. Among the successfully mapped reads, reads were removed from further analysis that either mapped to multiple sites, mapped with more than four mismatched bases or had more than ten clipped bases. Next, the Estimation_CRME script was run to eliminate strand-specific errors and exclude PCR-derived errors. Then, a strand-specific mismatch ratio was calculated for each nucleotide variation for both strands using the bases corresponding to 11–50 cycles. By excluding the top five cycles showing the highest mismatch rates, strand-specific mismatch rates were recalculated, and the smaller value between both strands was adopted as the nominal mismatch ratio. In addition, the nucleotide variations that were present across multiple pools were removed based on permutations across different pools using the Permut_Rm_com script because it is probable that such variations result from systemic sequencing errors. Finally, after excluding variations found in the dbSNP database, the database from the 1000 Genomes project or our in-house SNP database, the variants whose mismatch rate exceeded 0.009 were adopted as candidate mutations. Each candidate mutation was validated by Sanger sequencing of the 12 original individual DNAs from the corresponding DNA pools.

The functional impact of each amino acid substitution was evaluated by computer prediction using SIFT[33], PolyPhen-2 (ref. 34) and Mutation Taster[35]. The significance of nonsilent mutations in each cohesin component was evaluated assuming a uniform distribution of the background mutations within the coding regions, which was estimated to be ~0.3 Mb^{-1} on the basis of a previous whole-exome sequencing of myelodysplasia[3].

Determination of variant allele frequencies. Variant allele frequencies were evaluated by deep sequencing of PCR amplicons, pyrosequencing[36,37] and/or digital PCR (Fluidigm CA, US)[38–40] of the variants using nonamplified DNA. For amplicon sequencing, genomic fragments harboring the variants of interest were PCR amplified using NotI-tagged primers. Ninety-two randomly selected SNP loci that do not contain repetitive sequences were amplified using normal genomic DNA as a template, which served as the control. Touch-down PCRs using high-fidelity DNA polymerase KOD-Plus-Neo (TOYOBO, Tokyo) were performed, and an equimolar mixture of all PCR products was prepared for deep sequencing using HiSeq2000 or Miseq (Illumina), as described above, with a 75-bp or 100-bp pair end–read option. To calculate the allele frequency of each variant, all reads were mapped to the target reference sequence using BLAT[31], followed by differential enumeration of the dichotomic variant alleles. For indels, individual reads were first aligned to each of the wild-type and altered sequences and then assigned to the one with better alignment in terms of the number of matched bases.

Array-based copy-number and methylation analyses. Genomic DNA from 453 bone marrow samples with myeloid neoplasms was analyzed using GeneChip SNP genotyping microarrays as previously described using CNAG/AsCNAR software[41,42]. The results of the SNP array karyotyping for 290 of the 453 cases have been previously published[3,41–44]. The promoter methylation of each cohesin component gene was analyzed using the HumanMethylation450 BeadChip (Illumina), as previously described[30,45], in which methylation status was evaluated by calculating the ratio of methylation-specific and demethylation-specific fluorophores (β value) at each CpG site using iScan software (Illumina).

RT-PCR. Complementary DNA synthesis and quantitative RT-PCR analyses were performed as previously described[3]. The primer sequences used are listed in **Supplementary Tables 16** and **17**.

Protein expression of cohesin components in whole-cell extracts and chromatin-enriched fractions. Whole-cell extracts of myeloid cell lines were separated into soluble supernatant and chromatin-containing pellet fractions and analyzed by SDS-PAGE and protein blot analysis for the expression of different cohesin components as previously described[12,14]. Antibodies used for protein blot analysis are described in **Supplementary Table 18**.

Gene expression and cell proliferation assays. A full-length *RAD21* cDNA (BC050381) was provided by S. Sugano. A full-length *STAG2* cDNA was obtained from total cDNA derived from bone marrow cells and cloned into pBluescript. The truncated mutant of *RAD21* was subcloned by PCR. Flag-tagged *RAD21* or *STAG2* cDNAs were constructed into the retrovirus vector pGCDNsamIRESEGFP (provided by M. Onodera)[46] or a tetracycline-inducible lentiviral vector, CS-TRE-Ubc-tTA-IRESPuro. The wild-type *RAD21*, the mutant *RAD21* and/or a mock-induced retroviral vector were generated as previously described[3] and transduced into Kasumi-1, K562 and TF1 cells, which were sorted by GFP marking using a MoFlo FACS cell sorter (Beckman Coulter) or a BD FACSAria cell sorter (BD Biosciences) 48–96 h after retroviral transduction. The wild-type *RAD21*, the wild-type *STAG2* and a mock-induced lentiviral vector were generated as described previously[47], transduced into MOLM-13 cells and selected by 1 μg/ml puromycin. Gene expression was induced by 1 μg/ml doxycycline. For cell growth assays, the cells were inoculated into 96-well culture plates in RPMI 1640 medium supplemented

with 5% FCS (and 5 ng/ml GM-CSF for TF1 cells), and cell growth was monitored in three independent experiments by MTT assay using the Cell Counting Kit-8 (Dojindo Co.).

Expression microarray analysis. RNA was extracted from Kasumi-1 cells that were either mock transduced or transduced with wild-type RAD21 and analyzed in triplicate using the Human Genome U133 Plus 2.0 Array (Affymetrix) according to the manufacturer's protocol. For data analysis, raw array signals were first extracted from .CEL files using dChip Software[48]. After background correction and normalization across the six array data sets, the standardized signal value was obtained for each probe set in each of triplicate array experiments, which were compared between mock-transduced and wild-type *RAD21*–transduced cells. Two independent microarray experiments were performed. To identify transcriptionally altered genes, we used the criteria of fold change greater than ±1.2 and $P < 0.05$ (two-tailed paired t test) in two independent experiments.

RNA sequencing. RNA sequencing of *RAD21*-transduced Kasumi-1 cells and subsequent data analyses were performed as previously described[3] with minor modifications. For quantifications of expression values from the RNA sequencing data, we used a slightly modified version of RKPM (reads per kb of exon per million mapped reads) measures[49]. After removing the sequencing reads that were inappropriately aligned or that had low mapping quality, the number of bases on each exonic region for each RefSeq gene[50] was counted. Then the number of bases was normalized per kb of exon and per 100 million aligned bases. Finally, the expression value of each gene was determined by taking the maximum values among the RefSeq genes corresponding to the gene symbol.

We measured RAD21 expression by differentially enumerating endogenous and exogenous *RAD21* sequence reads, which were discriminated by the absence and presence of the Flag sequence, respectively. After normalization by the number of total reads for each sample, the raw differential read counts were further calibrated against the read counts containing the stop codon in *RAD21*.

Statistical analyses. The significance of the difference in frequency of cohesin component mutations between disease subtypes was tested by one-tailed Fisher's exact test. The coexistence of mutations was tested by two-tailed Fisher's direct method. The significance of the difference in the total number of somatic mutations between cohesin-mutated or -deleted and non-mutated or -deleted samples was tested by Mann-Whitney U test. Differences in the number of numerical abnormalities in cytogenetics between two groups with and without cohesin mutations or deletions was assessed by one-sided χ^2 test.

30. Sato, Y. *et al.* Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat. Genet.* doi:10.1038/ng.2699 (24 June 2013).
31. Kent, W.J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
32. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
33. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
34. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
35. Schwarz, J.M., Rodelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* **7**, 575–576 (2010).
36. Ronaghi, M. Pyrosequencing sheds light on DNA sequencing. *Genome Res.* **11**, 3–11 (2001).
37. Shih, L.Y. *et al.* Emerging kinetics of BCR-ABL1 mutations and their effect on disease outcomes in chronic myeloid leukemia patients with imatinib failure. *Leuk. Res.* **37**, 43–49 (2013).
38. Qin, J., Jones, R.C. & Ramakrishnan, R. Studying copy number variations using a nanofluidic platform. *Nucleic Acids Res.* **36**, e116 (2008).
39. Dube, S., Qin, J. & Ramakrishnan, R. Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. *PLoS ONE* **3**, e2876 (2008).
40. Totoki, Y. *et al.* High-resolution characterization of a hepatocellular carcinoma genome. *Nat. Genet.* **43**, 464–469 (2011).
41. Nannya, Y. *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* **65**, 6071–6079 (2005).
42. Yamamoto, G. *et al.* Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am. J. Hum. Genet.* **81**, 114–126 (2007).
43. Hosoya, N. *et al.* Genomewide screening of DNA copy number changes in chronic myelogenous leukemia with the use of high-resolution array-based comparative genomic hybridization. *Genes Chromosom. Cancer* **45**, 482–494 (2006).
44. Sanada, M. *et al.* Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature* **460**, 904–908 (2009).
45. Nagae, G. *et al.* Tissue-specific demethylation in CpG-poor promoters during cellular differentiation. *Hum. Mol. Genet.* **20**, 2710–2721 (2011).
46. Nabekura, T., Otsu, M., Nagasawa, T., Nakauchi, H. & Onodera, M. Potent vaccine therapy with dendritic cells genetically modified by the gene-silencing–resistant retroviral vector GCDNsap. *Mol. Ther.* **13**, 301–309 (2006).
47. Agarwal, S. *et al.* Isolation, characterization, and genetic complementation of a cellular mutant resistant to retroviral infection. *Proc. Natl. Acad. Sci. USA* **103**, 15933–15938 (2006).
48. Li, C. & Wong, W.H. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA* **98**, 31–36 (2001).
49. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
50. Pruitt, K.D., Tatusova, T., Brown, G.R. & Maglott, D.R. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* **40**, D130–D135 (2012).

PLOS ONE

The CD3 versus CD7 Plot in Multicolor Flow Cytometry Reflects Progression of Disease Stage in Patients Infected with HTLV-I

Seiichiro Kobayashi[1], Yamin Tian[1,2], Nobuhiro Ohno[3], Koichiro Yuji[3], Tomohiro Ishigaki[4], Masamichi Isobe[3], Mayuko Tsuda[3], Naoki Oyaizu[5], Eri Watanabe[4], Nobukazu Watanabe[4], Kenzaburo Tani[2], Arinobu Tojo[1,3], Kaoru Uchimaru[3]*

1 Division of Molecular Therapy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, **2** Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan, **3** Department of Hematology/Oncology, Research Hospital, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, **4** Laboratory of Diagnostic Medicine, Division of Stem Cell Therapy, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, **5** Clinical Laboratory, Research Hospital, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Abstract

Purpose: In a recent study to purify adult T-cell leukemia-lymphoma (ATL) cells from acute-type patients by flow cytometry, three subpopulations were observed in a CD3 versus CD7 plot (H: $CD3^{high}CD7^{high}$; D: $CD3^{dim}CD7^{dim}$; L: $CD3^{dim}CD7^{low}$). The majority of leukemia cells were enriched in the L subpopulation and the same clone was included in the D and L subpopulations, suggesting clonal evolution. In this study, we analyzed patients with indolent-type ATL and human T-cell leukemia virus type I (HTLV-I) asymptomatic carriers (ACs) to see whether the CD3 versus CD7 profile reflected progression in the properties of HTLV-I-infected cells.

Experimental Design: Using peripheral blood mononuclear cells from patient samples, we performed multi-color flow cytometry. Cells that underwent fluorescence-activated cell sorting were subjected to molecular analyses, including inverse long PCR.

Results: In the D(%) versus L(%) plot, patient data could largely be categorized into three groups (Group 1: AC; Group 2: smoldering- and chronic-type ATL; and Group 3: acute-type ATL). Some exceptions, however, were noted (*e.g.*, ACs in Group 2). In the follow-up of some patients, clinical disease progression correlated well with the CD3 versus CD7 profile. In clonality analysis, we clearly detected a major clone in the D and L subpopulations in ATL cases and, intriguingly, in some ACs in Group 2.

Conclusion: We propose that the CD3 versus CD7 plot reflects progression of disease stage in patients infected with HTLV-I. The CD3 versus CD7 profile will be a new indicator, along with high proviral load, for HTLV-I ACs in forecasting disease progression.

Citation: Kobayashi S, Tian Y, Ohno N, Yuji K, Ishigaki T, et al. (2013) The CD3 versus CD7 Plot in Multicolor Flow Cytometry Reflects Progression of Disease Stage in Patients Infected with HTLV-I. PLoS ONE 8(1): e53728. doi:10.1371/journal.pone.0053728

Editor: Jean-Pierre Vartanian, Institut Pasteur, France

Received August 31, 2012; **Accepted** December 4, 2012; **Published** January 22, 2013

Funding: This study was supported by the Ministry of Education, Culture, Sports, Science and Technology, Japan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: uchimaru@ims.u-tokyo.ac.jp

Introduction

Human T-cell leukemia virus type I (HTLV-I) is the agent that causes HTLV-I-associated diseases, such as adult T-cell leukemia-lymphoma (ATL), HTLV-I-associated myelopathy/tropical spastic paraparesis (HAM/TSP), and HTLV-I uveitis (HU) [1–3]. Approximately 10–20 million people are infected with the HTLV-I virus worldwide [4]. The lifetime risk of developing ATL is estimated to be approximately 2.5–5% [5,6]. ATL includes a spectrum of diseases that are referred to as smoldering-, chronic-, lymphoma-, and acute-type [7,8]. The chronic and smoldering types of ATL are considered indolent and are usually managed with watchful waiting until the disease progresses to aggressive (lymphoma- or acute-type) ATL [9]. Because the prognosis of ATL is poor with current treatment strategies, factors to forecast progression to ATL from asymptomatic carriers (ACs) have been researched [10–13] in the hope that they will be useful for preventive therapy under development in the early malignant stage.

Various cellular dysfunctions induced by viral genes (*e.g.*, tax and HBZ), genetic and epigenetic alterations, and the host immune system are considered to cooperatively contribute to leukemogenesis in ATL [14–16]. However, the complex mechanism may hinder determination of a clear mechanism of the pathology and make discovery of risk factors difficult. In a prospective nationwide study in Japan, high proviral load (VL,

Table 1. Clinical profile of patients infected with HTLV-I and normal controls.

Clinical sutype	Number of cases	Male	Female	Age (range)	WBC(/μl) (range)	Lymphocytes(%) (range)	Abnormal lymphocytes(%) (range)
HTLV-1 AC	40	12	28	49.9 (28–70)	5525 (2680–10360)	35.9 (22.4–59.5)	0.9 (0.0–4.4)
Smoldering	7	4	3	55.3 (43–77)	5944 (3680–8710)	32.5 (13.4–47.5)	5.8 (0.7–16.5)
Chronic	7	4	3	52.7 (37–60)	9180 (4070–12790)	45.8 (35.0–61.5)	9.2 (3.4–12.7)
Acute	13	4	9	58.8 (42–74)	15328 (4450–41480)	16.3 (1.7–50.5)	40.3 (3.0–89.6)
Normal controls	10	6	4	47.4 (27–66)	ND	ND	ND

WBC: white blood cells (normal range, 3500–9100/μl).
AC: asymptomatic carrier.
ND: analysis were not performed.
Average of age, WBC, lymphocytes (%) and abnormal lymphocytes (%) are shown.
The proportion of abonormal lymphocytes in peripheral blood WBCs was evaluated by morphological exmanination.
doi:10.1371/journal.pone.0053728.t001

over 4.17 copies/100 peripheral blood mononuclear cells) was found to be a major risk factor for HTLV-I AC developing into ATL [13]. Although VL indicates the proportion of HTLV-I-infected cells, it does not indicate size or degree of malignant progression in each clone; *i.e.*, it does not directly indicate progression of disease stage in HTLV-I infection. Moreover, the majority of ACs with high VL remained intact during the study period, indicating that a more accurate indicator of progression is needed.

In our recent study to purify monoclonal ATL cells from acute-type patients by flow cytometry, three subpopulations were observed in a CD3 versus CD7 plot of $CD4^+$ cells (H: $CD3^{high}CD7^{high}$, D: $CD3^{dim}CD7^{dim}$, L: $CD3^{dim}CD7^{low}$), and the majority of ATL cells were enriched in the L subpopulation [17]. Clonality analyses revealed that the D and L subpopulations contained the same clone, suggesting clonal evolution of HTLV-I-infected cells to ATL cells. From these findings, we speculated that the CD3 versus CD7 profile may reflect disease progression in HTLV-I infection. In this study, the CD3 versus CD7 profile by flow cytometry, combined with molecular (clonality and proviral load) characterizations, were analyzed in patients with various clinical subtypes (HTLV-I AC, and indolent and aggressive ATL). We found that the CD3 versus CD7 profile reflected disease progression of HTLV-I-infected cells to ATL cells. We also discuss the significance of this analysis as a novel risk indicator for HTLV-I ACs in forecasting progression to ATL.

Materials and Methods

Cell lines and patient samples

TL-Oml, an HTLV-I-infected cell line, established Dr. Hinuma's laboratory [18], was provided by Dr. Toshiki Watanabe (The University of Tokyo, Tokyo, Japan) and was cultured in RPMI-1640 medium containing 10% fetal bovine serum. Peripheral blood samples were collected from inpatients and outpatients at our hospital from August 2009 to November 2011. All patients with ATL were categorized according to Shimoyama's criteria [7,8]. Patients with various complications, such as autoimmune disorder and systemic infections, were excluded. Lymphoma-type patients were excluded because ATL cells are not considered to exist in peripheral blood of this clinical subtype. In patients with ATL receiving chemotherapy, blood samples were collected before treatment or during the recovery phase between chemotherapy sessions. Samples collected from 10 healthy volunteers (mean age: 47.4 years; range: 27–66 years) were used as normal controls.

The present study was approved by the research ethics committee of the institute of medical science, the university of Tokyo. Subjects provided written informed consent.

Flow cytometry and cell sorting

Peripheral blood mononuclear cells (PBMCs) were isolated from heparin-treated whole blood by density gradient centrifugation, as described previously [17]. Cells were stained using a combination of phycoerythrin (PE)-CD7, APC-Cy7-CD3, Pacific Blue-CD4, and Pacific Orange-CD14. Pacific Orange-CD14 was purchased from Caltag-Invitrogen (Carlsbad, CA). All other antibodies were obtained from BD BioSciences (San Jose, CA). Propidium iodide (PI; Sigma, St. Louis, MO) was added to the samples to stain dead cells immediately prior to flow cytometry. A BD FACS Aria instrument (BD Immunocytometry Systems, San Jose, CA) was used for all multicolor flow cytometry and cell sorting. Data were analyzed using the FlowJo software (Treestar, San Carlos, CA).

Quantification of HTLV-I proviral load by real-time quantitative polymerase chain reaction (PCR)

The HTLV-I proviral load in FACS-sorted PBMCs was quantified by real-time quantitative polymerase chain reaction (PCR; TaqMan method) using the ABI Prism 7000 sequence detection system (Applied Biosystems, Foster City, CA) as described previously [13,17]. Briefly, 50 ng of genomic DNA was extracted from human PBMCs using a QIAamp DNA blood Micro kit (Qiagen, Hilden, Germany). Triplicate samples of the DNA were amplified. Each PCR mixture, containing an HTLV-I pX region-specific primer pair at 0.1 μM (forward primer 5′-CGGATACCCAGTCTACGTGTT-3′ and reverse primer 5′-

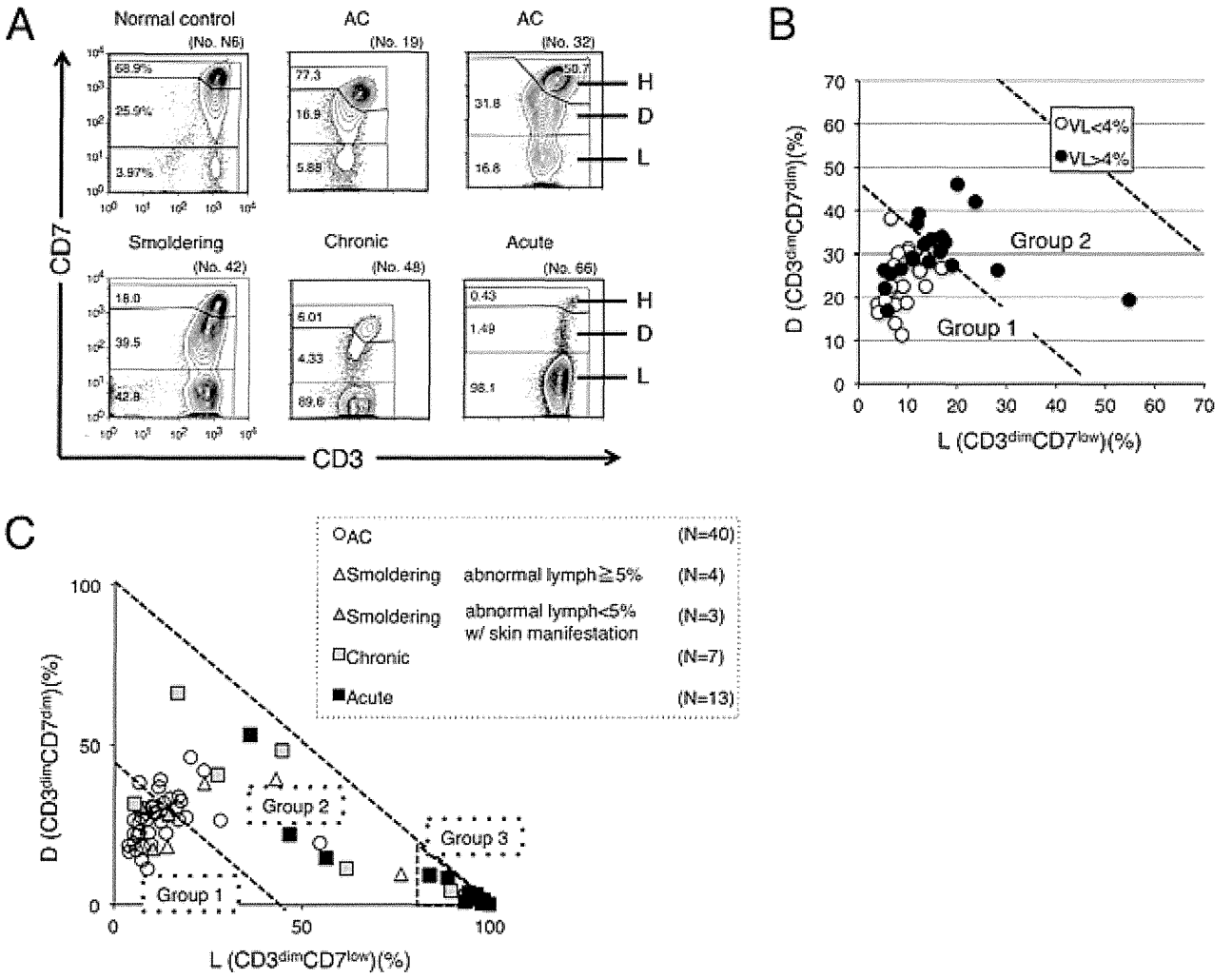

Figure 1. CD3 versus CD7 plots in flow cytometric analysis of patients who are asymptomatic HTLV-I carriers (ACs) and have various clinical subtypes of adult T-cell leukemia-lymphoma (ATL) suggest disease progression in HTLV-I infection. (A) Flow cytometric profile of an AC, various clinical subtypes of ATL (smoldering, chronic, and acute), and a normal control. Representative cases of CD3 versus CD7 plots in $CD4^+$ cells are shown. (B) A two-dimensional plot of AC cases showing the percentage of the D and L subpopulations by flow cytometry. AC cases were divided into two groups according to HTLV-I VL (greater or less than 4%). The border line (45% of D+L subpopulations) between Group 1 and 2 was set based on proviral load (VL). All AC cases with less than 4% VL were included in Group 1. All AC cases included in Group 2 had greater than 4% VL. VL<4%: n = 21; VL>4%: n = 19. All VL data in this figure were provided from the database of the Joint Study on Predisposing Factors of ATL Development (JSPFAD). (C) A two-dimensional plot of all patients showing the percentage of the D and L subpopulations. The smoldering type was divided into two categories: smoldering type with greater than 5% abnormal lymphocytes and smoldering type with less than 5% abnormal lymphocytes with skin manifestation. The two diagonal dotted lines indicate 45% and 100% of D+L subpopulations (*i.e.*, 55% and 0% of the H subpopulation). Data were categorized into three groups.
doi:10.1371/journal.pone.0053728.g001

CAGTAGGGCGTGACGATGTA-3′), FAM-labeled probe at 0.1 μM (5′- CTGTGTACAAGGCGACTGGTGCC-3′), and 1× TaqMan Universal PCR master mix (Applied Biosystems), was subjected to 50 cycles of denaturation (95°C, 15 seconds) and annealing to extension (60°C, 1 minute), following an initial Taq polymerase activation step (95°C, 10 minutes). The RNase P control reagent (Applied Biosystems) was used as an internal control for calculating the input cell number (using VIC reporter dye). DNAs extracted from TL-Om1 and normal human PBMCs were used as positive and negative controls, respectively. The HTLV-I proviral load (%) was calculated as the copy number of the pX region per input cell number. To correct the deviation of data acquired in each experiment, data from TL-Om1 (positive control) were adjusted to 100%, and the sample data were corrected accordingly by a proportional calculation.

Inverse long PCR

For clonality analysis, inverse long PCR was performed [17]. First, 1 μg of genomic DNA extracted from the FACS-sorted cells was digested with *Eco*RI and *Pst*I at 37°C overnight. Purification of DNA fragments was performed using a QIAEX2 gel extraction kit (Qiagen). The purified DNA was self-ligated with T4 DNA ligase (Takara Bio, Otsu, Japan) at 16°C overnight. The circular DNA obtained from the *Eco*RI digestion fragment was then digested

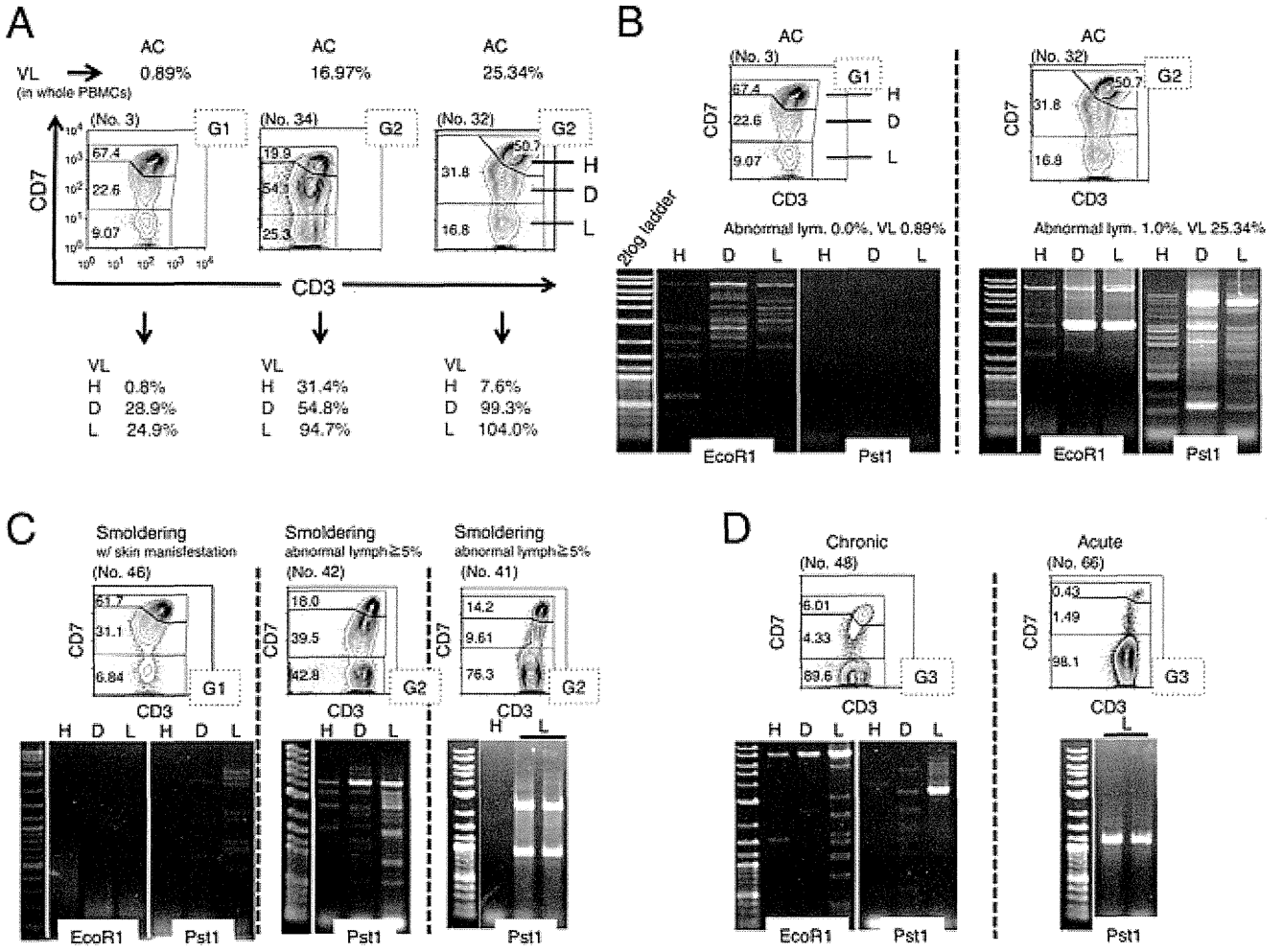

Figure 2. HTLV-I proviral load (VL) and clonality in each subpopulation, based on the CD3 versus CD7 plot. (A) The three subpopulations (H, D, L) based on the CD3 versus CD7 plot were subjected to fluorescence-activated cell sorting (FACS) and VL analysis. Three representative cases are shown. G1 or G2 in the dotted box indicates Group 1 or Group 2, categorized by the percentage of the D and L subpopulations, respectively. (B)–(D) Analysis of clonality in the three subpopulations based on the CD3 versus CD7 plot. Genomic DNA was extracted from FACS-sorted cells of each subpopulation and subjected to inverse long polymerase chain reaction (PCR). Representative data of two cases of AC (B), three cases of smoldering type, including one with skin manifestations (C), and cases of a chronic type and an acute type (D) are shown. PCR was performed in duplicate (black bars) in cases when a sufficient amount of DNA was obtained.
doi:10.1371/journal.pone.0053728.g002

with *Mlu*I, which cuts the pX region of the HTLV-I genome and prevents amplification of the viral genome. Inverse long PCR was performed using Takara LA *Taq* polymerase (Takara Bio). For the *Eco*RI-treated template, the forward primer was 5′-TGCCTGACCCTGCTTGCTCAACTCTACGTCTTTG-3′ and the reverse primer was 5′-AGTCTGGGCCCT-GACCTTTTCAGACTTCTGTTTC-3′. For the *Pst*I-treated group, the forward primer was 5′-CAGCCCATTCTATAG-CACTCTCCAGGAGAG-3′ and the reverse primer was 5′-CAGTCTCCAAACACGTAGACTGGGTATCCG-3. Each 50-μL reaction mixture contained 0.4 mM of each dNTP, 25 mM $MgCl_2$, 10× LA PCR buffer II containing 20 mM Tris-HCl and 100 mM KCl, 0.5 mM of each primer, 2.5 U LA *Taq* polymerase, and 50 ng of the processed genomic DNA. The reaction mixture was subjected to 35 cycles of denaturation (94°C, 30 seconds) and annealing to extension (68°C, 8 minutes). Following PCR, the products were subjected to electrophoresis on 0.8% agarose gels. In samples from which a sufficient amount of DNA was extracted, PCRs were performed in duplicate.

Results

CD3 versus CD7 profile in flow cytometry in various clinical subtypes of patients infected with HTLV-I

The clinical profiles of the 77 cases analyzed in this study are shown in Table 1. According to the gating procedure, as shown in Figure S1 [17], we constructed a CD3 versus CD7 plot of $CD4^+$ cells in PBMCs of various clinical subtypes from patients infected with HTLV-I and normal controls. The three subpopulations ($CD3^{high}CD7^{high}$, $CD3^{dim}CD7^{dim}$, and $CD3^{dim}CD7^{low}$) observed are referred to as the H, D, and L subpopulations, respectively. Representative results for each clinical subtype of HTLV-I infection are shown in Figure 1A. Regarding the data for an acute-type patient (no. 66), the dominant population was the L subpopulation, in which we previously demonstrated that monoclonal ATL cells are enriched [17]. Regarding the AC (no. 19), the CD3 versus CD7 profile was close to that of the normal control, although in some AC cases, such as no. 32, the profile differed from that of the normal control, because in contrast to case no. 19,

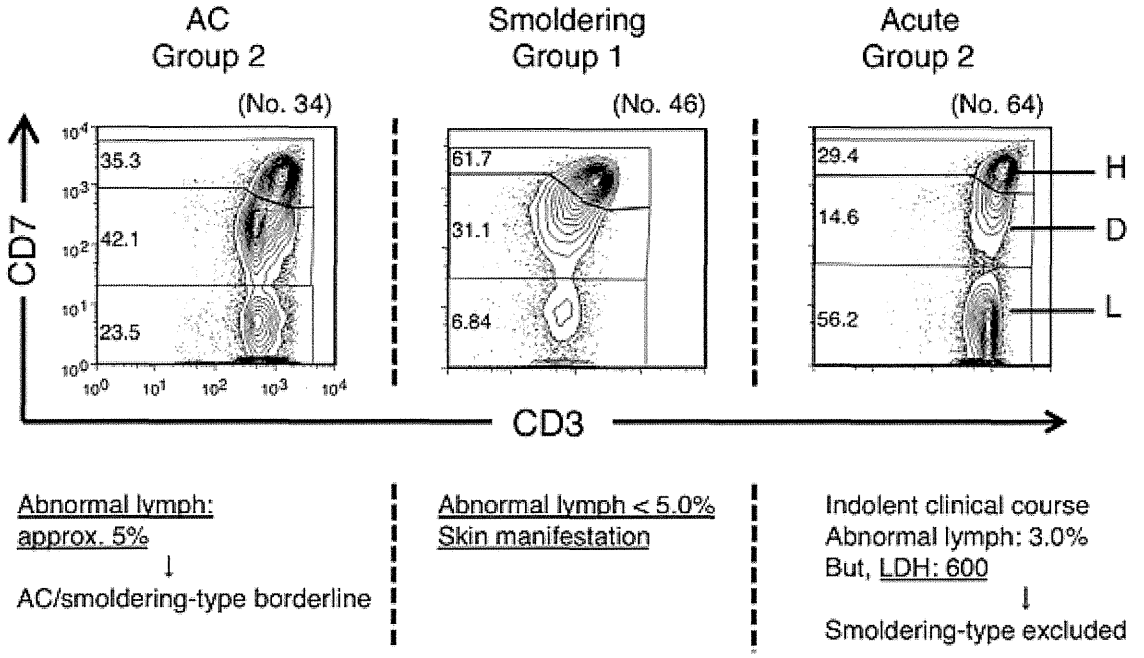

Figure 3. Study of exceptional cases categorized by proportion of the $CD3^{dim}CD7^{dim}$ (D) and $CD3^{dim}CD7^{low}$ (L) subpopulations. Left: An HTLV-I AC patient who was categorized in Group 2 in the D(%) versus L(%) plot. Middle: A patient with smoldering-type ATL who was categorized in Group 1. Right: A patient with acute-type ATL who was categorized in Group 2.
doi:10.1371/journal.pone.0053728.g003

these cases had increased D and L subpopulations. Regarding the data for indolent-type disease (smoldering and chronic), increases in the D and L subpopulations were intermediate between ACs and patients with acute-type disease. These representative flow cytometric data suggest that continuity in the CD3 versus CD7 profile seemed to exist among the various clinical subtypes of patients infected with HTLV-I.

The proportions of D and L subpopulations in all AC cases analyzed are shown in Figure 1B. Because the high HTLV-I proviral load (VL) in whole PBMCs, a VL of >4%, was reported to be a major risk indicator for progression to ATL [13], a border line was set based on VL. Group 1, the area under the diagonal line (D+L = 45%), included all AC cases with VLs of <4%. ACs with VLs of >4% were distributed between Groups 1 and 2. The proportions of D and L subpopulations in normal controls are shown in Figure S2. In this plot, all data for normal controls were distributed in Group 1. Data for all clinical subtypes are shown in Figure 1C. Most data for acute-type patients were located in the area beyond 80% of the L subpopulation and we designated this area as Group 3. Group 2, which is located between Group 1 and Group 3, included the majority of indolent-type (smoldering and chronic) cases. From these results, the three groups in the D(%) versus L(%) plot seemed to represent disease stage in each case.

Proviral load and clonality in each subpopulation in the CD3 versus CD7 plot

To further characterize each subpopulation (H, D, and L) in the CD3 versus CD7 plot, cells in each subpopulation were FACS-sorted and subjected to analysis of VL to determine the percentage of HTLV-I-infected cells in each subpopulation. Results for representative cases are shown in Figure 2A. The VL in whole PBMCs of an AC (no. 3) was low (0.89%). As expected, the VL in H, the major subpopulation, was low (0.8%). However, VLs in the D and L subpopulations were considerably higher (28.9% and 24.9%, respectively), indicating that HTLV-I-infected cells are relatively concentrated in these subpopulations. In the cases with high VLs in whole PBMCs (no. 32 with 25.34%; no. 34 with 16.97%), the VLs were also higher in the D and L subpopulations, and almost all cells in the L subpopulation were HTLV-I-infected.

In HTLV-I infection, progression to ATL requires several pathological steps, including clonal expansion [15]. To further characterize the three subpopulations based on the CD3 versus CD7 plot, we analyzed clonality in each subpopulation in patients with various clinical subtypes using the inverse long PCR method. Figure 2B shows two cases of AC. In the left case (no. 3), included in Group 1 in the D(%) and L(%) plot, multiple bands suggestive of multiple small clones were detected in the three subpopulations. However, no major band suggestive of a dominant clone was observed. In the right case (no. 32), included in Group 2, inverse long PCR of the FACS-sorted subpopulations suggested that the D and L subpopulations contained a major clone (Figure 2B). The D subpopulation had bands of the same size as those of the L subpopulation, indicating that the two distinct subpopulations contained a common major clone. Eleven cases of AC were included in Group 2. All three cases analyzed by Southern blotting (whole blood samples) were positive for clonal bands (Figure S3). In Figure 2C, data for three smoldering cases are shown. In case no. 46 (left), whose only manifestation was a skin eruption with few abnormal lymphocytes (less than 5% of white blood cells) in the peripheral blood, only faint minor bands suggestive of small clones were observed. In contrast, in the other two cases (nos. 42 and 41), intense bands suggestive of major clones were observed in both the D and L subpopulations. In no. 41 (right), weak bands were not visible, which suggested the selection of dominant clones. In Figure 2D, data for a chronic-type case and an acute-type case are shown. In both cases, intense bands in the L subpopulation suggest the existence of a major clone. The series of clonality analyses indicated that a major clone became more evident and the clinical

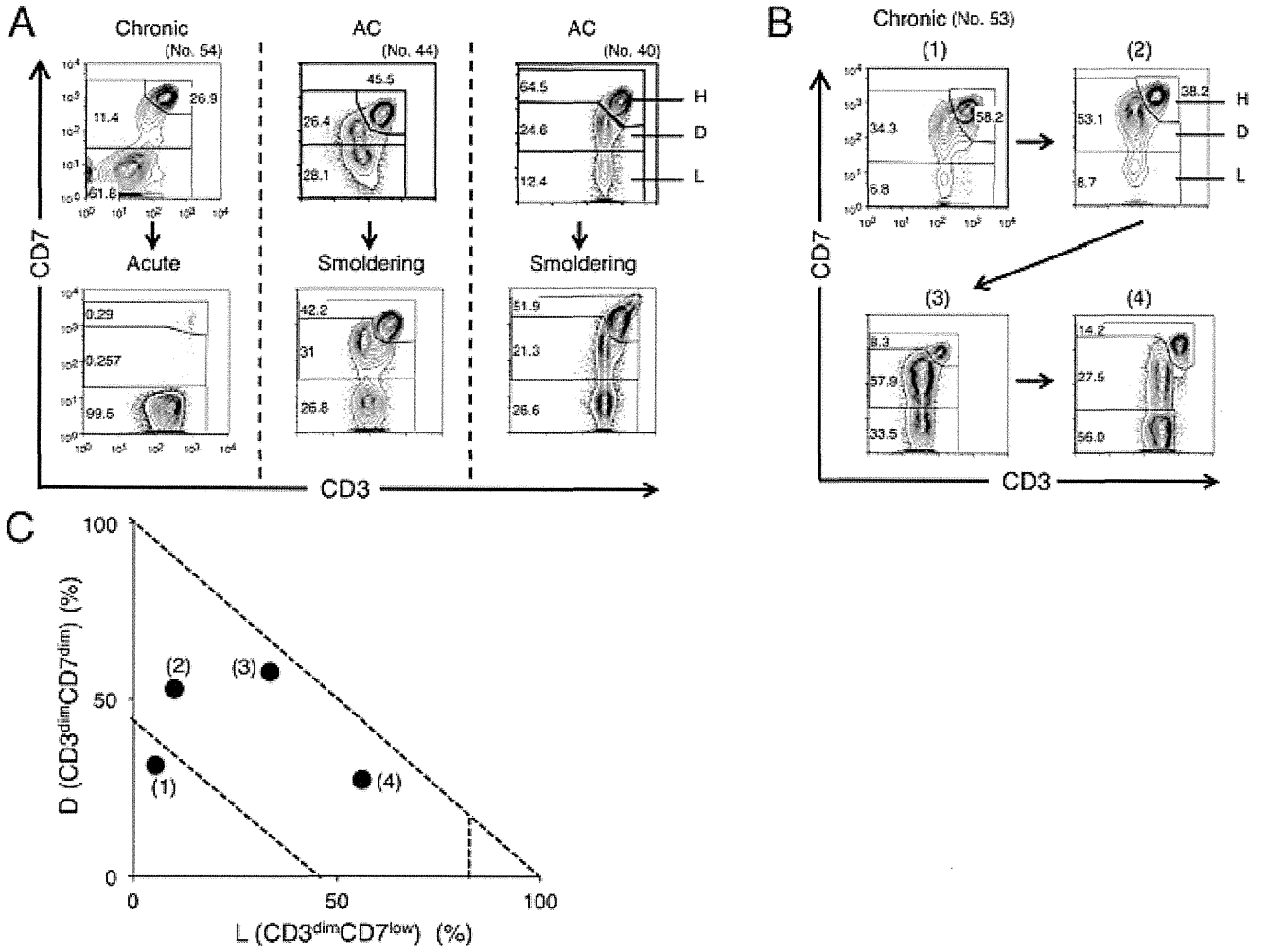

Figure 4. Alteration in the CD3 versus CD7 profile by flow cytometry in accordance with disease progression. (A) Change in the CD3 versus CD7 profile in representative cases. In all three cases shown, change in clinical data (*e.g.*, abnormal lymphocyte, LDH) resulted in progression of the clinical subtype. (B) Change in the CD3 versus CD7 profile in a time course in the case of chronic-type ATL. Clinical data are shown in Table S1. (C) Flow cytometric data in (B) are summarized in the D(%) versus L(%) plot.
doi:10.1371/journal.pone.0053728.g004

stage became more advanced as the D and L subpopulations increased.

Clinical evaluation of exceptional cases categorized by proportions of the $CD3^{dim}CD7^{dim}$ (D) and $CD3^{dim}CD7^{low}$ (L) subpopulations

As noted above, the D(%) versus L(%) plot generally represented disease stage in HTLV-I infection. However, we observed one case of chronic-type disease and three cases of smoldering-type disease in Group 1 and three cases of acute-type disease in Group 2. Furthermore, some ACs with VLs of >4% were observed in Group 2. Representative data from these apparently exceptional cases are shown in Figure 3. On the left, a case of AC (no. 34) observed in Group 2 is shown. 4.7% of lymphocytes in this blood sample were abnormal and clonality analysis by Southern blotting showed oligoclonal bands suggestive of clones of substantial size (Figure S3). These clinical data suggest that the disease stage would be around the AC/smoldering borderline. In the middle, a case of a smoldering type (no. 46) observed in Group 1 is shown. In this case, the percentage of abnormal lymphocytes in the peripheral blood was only 1%, but she had a histologically proven ATL lesion in the skin and was diagnosed with smoldering-type ATL. The other two smoldering cases categorized in Group 1 were the same as this case. These results indicate that ATL cells in these three smoldering cases infiltrated the skin, but not the peripheral blood. On the right, a case of acute-type disease categorized as Group 2 (no. 64) is shown. The clinical course of this patient was relatively indolent compared with typical acute-type disease. He had skin infiltration of ATL cells, but no lymph node swelling. However, LDH exceeded 1.5 times the upper limit of the normal range, which excludes a diagnosis of smoldering-type disease. Other acute-type cases categorized in Group 2 were diagnosed as such according to Shimoyama's criteria, but also had the same indolent clinical course as case no. 64. These cases should have been regarded as indolent ATL.

Changes in the CD3 versus CD7 profile in flow cytometry with disease progression

In several cases, we could obtain time-sequential samples (Figure 4). The patient (no. 54) shown on the left in Figure 4A progressed from chronic-type to acute-type disease. In flow cytometric analysis, decreases in the H and D subpopulations

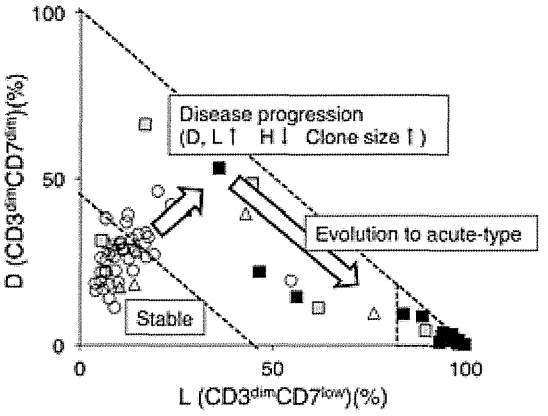

Figure 5. Summary of the study: the CD3 versus CD7 profile reflects progression of disease stage in patients infected with HTLV-I. In the percentage of D ($CD3^{dim}CD7^{dim}$) versus L ($CD3^{dim}CD7^{low}$) plot, Group 1 includes the majority of AC cases. As disease stage progresses, the CD3 versus CD7 profile then changes. With downregulation of CD3 and CD7, the D and L subpopulations increase gradually (Group 2). During this step, clones in the D and L subpopulations increase in size. Further accumulation of genetic alterations will result in rapid expansion of ATL clones—*i.e.*, evolution to acute-type ATL. In this step, the CD3 versus CD7 profile will progress from Group 2 to 3.
doi:10.1371/journal.pone.0053728.g005

and an increase in the L subpopulation were observed, indicating that disease progression correlated well with the change in the CD3 versus CD7 profile. The patients in the middle (no. 44) and on the right (no. 40) were included in Group 2 at the AC stage and later progressed to smoldering-type ATL. Although variation in the change of the flow cytometric profile was seen between these patients, the results suggest that ACs in Group 2 are at high risk of developing ATL.

The patient in Figure 4B (no. 53) was initially diagnosed with AC and later progressed to chronic-type ATL. Although the initial clinical course was stable, an increase in abnormal lymphocyte numbers was later observed, and low-dose VP-16 therapy (50 mg/day) was initiated because of hypoxemia due to lung infiltration of ATL cells. Table S1 and Figure 4C show summaries of the clinical data and the flow cytometric analyses, respectively. The flow cytometric data correlated well with disease progression.

Discussion

Findings in our previous analysis of acute-type ATL samples prompted our analysis of various clinical subtypes of patients infected with HTLV-I to examine whether the CD3 versus CD7 profile reflects the progression of oncogenesis in HTLV-I-infected cells [17]. Representative flow cytometric data shown in Figure 1A suggested that the CD3 versus CD7 profile changed during disease progression. As the disease stage progressed, the D and L subpopulations increased with concomitant decreases in the $CD3^{high}CD7^{high}$(H) subpopulation. Figure 1C, a summary of the flow cytometric data of all cases analyzed, reveals that the two-dimensional plot of the proportions of the D versus L subpopulations could divide all cases into three groups. Group 1, the area under the diagonal line, equivalent to 55% of the H subpopulation in which all normal controls were included (Figure S2), contained the majority of HTLV-I ACs. Group 3 was the area beyond 80% of the L subpopulation, and the majority of acute-type cases were included in this group. Group 2, located between Groups 1 and 3 (*i.e.*, less than 55% of the H subpopulation and 80% of the L subpopulation), included indolent-type (smoldering and chronic) cases and some AC cases. These results suggest that the CD3 versus CD7 expression profile reflects disease stage. Initially, both the D and L subpopulations gradually and simultaneously increased. However, at the clinically advanced stage, the increase in the L subpopulation was prominent. The change is considered to reflect the biological difference between the D and L subpopulations, which needs to be clarified.

In HTLV-I infection, the small clones of infected cells are considered to coexist from the AC stage [19,20]. A selected clone from the multiple small clones then grows and progresses to the malignant state, and the emergence of a dominant clone indicates disease progression in ATL [19,20]. As shown in Figure 2B–D, major bands suggesting dominant clones were evident in patients with progressed clinical subtypes or those in the advanced group in the CD3 versus CD7 profile, and major bands existed exclusively in the D and L subpopulations. These data also support the idea that increases in the D and L subpopulations correlate with the progression of disease stage. AC cases in Group 2 had high HTLV-I proviral loads (>4%; Figure 1B) and clear major bands were observed by inverse long PCR in these cases (Figure 2B, right). Sasaki *et al.* reported that two cases of HTVL-I AC with oligoclonal bands on Southern blots and high VLs (20%) had progressed to ATL by 4 and 3.5 years later [21]. The two cases may correspond to HTLV-I AC in Group 2 proposed in our study. In fact, two cases of ACs in our series that were included in Group 2 progressed to smoldering ATL (Figure 4A). AC cases in Group 2 could be regarded as advanced carriers. Our flow cytometric analysis could apparently discriminate high-risk AC cases from stable ones. Follow-up analysis of these cases is warranted to determine whether AC cases included in Group 2 progress to ATL. Flow cytometric data for these AC cases included in Group 2 (Figure 1A and 1C) were similar to those for indolent ATL cases in Group 2. These ACs in Group 2 can be considered essentially the same as smoldering ATL cases. Some of the ACs categorized according to Shimoyama's criteria should in fact be separated and regarded as a subtype together with at least some of the smoldering ATL cases.

Iwanaga *et al.* reported that high HTLV-I proviral load (>4%) in whole PBMCs was a risk factor for progression to ATL [13]. In Figure 1B, the ACs with VLs>4% were distributed between Groups 1 and 2. These findings suggest that not all ACs with high VLs are currently in an advanced stage, although they may have the potential to develop ATL in the future.

In general, the categorization by flow cytometric profile correlated well with the current classification of clinical subtypes, with some exceptional cases of acute-type and smoldering-type disease (Figure 3). The only manifestation of three smoldering cases categorized in Group 1 was skin lesions; they fell into Group 1 because they showed minimal abnormalities in peripheral blood [22]. Three acute-type ATL cases categorized in Group 2 had indolent clinical courses. A diagnosis of acute-type disease is made when the indolent-type and lymphoma-type are excluded, according to Shimoyama's criteria. The CD3 versus CD7 plot may discriminate the cases that will follow an indolent clinical course from the aggressive acute-type ATL.

The VL in each subpopulation indicated that HTLV-I-infected cells were relatively concentrated in the D and L subpopulations (representative data are shown in Figure 2A). These data are consistent with downregulation of CD3 and CD7 being relevant to HTLV-I infection, although cells without HTLV-I infection may also contribute to this change to some extent, as a substantial subpopulation of T cells has been reported to be CD7-deficient under physiological [23,24] and certain pathological conditions,

including autoimmune disorders and viral infection [25–29]. To more precisely analyze phenotypic changes in HTLV-I-infected cells, markers that indicate HTLV-I infection should be incorporated in future studies.

A summary of this study is shown in Figure 5. In the CD3 versus CD7 profile, most AC cases were included in Group 1, in which the D and L subpopulations were relatively small. Consistent with disease progression to smoldering- and chronic-type ATL, a decrease in the H subpopulation and increases in the D and L subpopulations occur (Group 2). In this step, increases in the sizes of clones in the D and L subpopulations are observed. Further expansion of the leukemic clone results in progression to acute-type ATL in which the L subpopulation has expanded (Group 3). According to a study by Yamaguchi *et al.*, the natural course of ATL is to progress from the HTLV-I carrier state through the intermediate state, smoldering ATL, and chronic ATL, and finally to the acute ATL, indicating a process of multistage leukemogenesis [19]. We consider this study to successfully link the progressive clinical status and phenotypic changes in HTLV-I-infected cells. However, the way in which this profile reflects multistep oncogenesis in HTVL-I infection at the molecular level remains unclear. Further molecular analyses of the three subpopulations will help in understanding the mechanism(s).

Supporting Information

Figure S1 Representative flow cytometric analysis of an HTLV-I asymptomatic carrier (patient no. 32). The CD3 versus CD7 plot of $CD4^+$ cells was constructed according to the gating procedure shown in this figure. In the plot, we designated three subpopulations: H ($CD3^{high}CD7^{high}$), D ($CD3^{dim}CD7^{dim}$), and L ($CD3^{dim}CD7^{low}$).
(PPTX)

Figure S2 A two-dimensional plot of 10 normal controls showing the percentage of the D and L subpopulations.
(PPTX)

Figure S3 Southern blot analysis of clonal integration of the HTLV-I provirus. Representative data (AC, No. 34) are shown. In *Eco*RI or *Pst*I digestion, a band indicated by a red arrow represents the monoclonal integration of the provirus. The band pattern indicates that two major clones coexist. This analysis was performed by a commercial laboratory (SRL, Tokyo, Japan).
(PPTX)

Table S1 Clinical data in a case of chronic-type ATL (No. 53). Proportion of abonormal lymphocytes in the preipheral blood WBC were evaluated by morphological exmanination. LDH: Lactate dehydrogenase (normal range, 120–240 U/L) sIL-2R: soluble interleukin-2 receptor (normal range, 122–496 U/ml).
(XLSX)

Acknowledgments

We thank Dr. Toshiki Watanabe, Dr. Kazumi Nakano, and Dr. Tadanori Yamochi (The University of Tokyo) for providing the TL-Om1 cell line and the plasmid containing the HTLV-I genome, which was used as a standard for the quantification of proviral load. We also thank Mr. Yuji Zaike (Clinical Laboratory, Research Hospital, Institute of Medical Science, The University of Tokyo) for his excellent technical advice. We are grateful to the hospital staff who have made a commitment to providing high-quality care to all of our patients.

Author Contributions

Conceived and designed the experiments: KT AT KU. Performed the experiments: SK YT. Analyzed the data: EW NW TI NO. Contributed reagents/materials/analysis tools: MI MT KU NO. Wrote the paper: SK KU.

References

1. Yoshida M, Miyoshi I, Hinuma Y (1982) Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. Proc Natl Acad Sci U S A 79: 2031–2035.
2. Osame M, Usuku K, Izumo S, Ijichi N, Amitani H, et al. (1986) HTLV-I associated myelopathy, a new clinical entity. Lancet 1: 1031–1032.
3. Mochizuki M, Watanabe T, Yamaguchi K, Takatsuki K, Yoshimura K, et al. (1992) HTLV-I uveitis: a distinct clinical entity caused by HTLV-I. Japanese journal of cancer research : Gann 83: 236–239.
4. Proietti FA, Carneiro-Proietti AB, Catalan-Soares BC, Murphy EL (2005) Global epidemiology of HTLV-I infection and associated diseases. Oncogene 24: 6058–6068.
5. Yamaguchi K, Watanabe T (2002) Human T lymphotropic virus type-I and adult T-cell leukemia in Japan. International journal of hematology 76 Suppl 2: 240–245.
6. Murphy EL, Hanchard B, Figueroa JP, Gibbs WN, Lofters WS, et al. (1989) Modelling the risk of adult T-cell leukemia/lymphoma in persons infected with human T-lymphotropic virus type I. International journal of cancer Journal international du cancer 43: 250–253.
7. Shimoyama M (1991) Diagnostic criteria and classification of clinical subtypes of adult T-cell leukaemia-lymphoma. A report from the Lymphoma Study Group (1984–87). Br J Haematol 79: 428–437.
8. Tsukasaki K, Hermine O, Bazarbachi A, Ratner L, Ramos JC, et al. (2009) Definition, prognostic factors, treatment, and response criteria of adult T-cell leukemia-lymphoma: a proposal from an international consensus meeting. J Clin Oncol 27: 453–459.
9. Takasaki Y, Iwanaga M, Imaizumi Y, Tawara M, Joh T, et al. (2010) Long-term study of indolent adult T-cell leukemia-lymphoma. Blood 115: 4337–4343.
10. Hisada M, Okayama A, Shioiri S, Spiegelman DL, Stuver SO, et al. (1998) Risk factors for adult T-cell leukemia among carriers of human T-lymphotropic virus type I. Blood 92: 3557–3561.
11. Imaizumi Y, Iwanaga M, Tsukasaki K, Hata T, Tomonaga M, et al. (2005) Natural course of HTLV-1 carriers with monoclonal proliferation of T lymphocytes ("pre-ATL") in a 20-year follow-up study. Blood 105: 903–904.
12. Kamihira S, Atogami S, Sohda H, Momita S, Yamada Y, et al. (1994) Significance of soluble interleukin-2 receptor levels for evaluation of the progression of adult T-cell leukemia. Cancer 73: 2753–2758.
13. Iwanaga M, Watanabe T, Utsunomiya A, Okayama A, Uchimaru K, et al. (2010) Human T-cell leukemia virus type I (HTLV-1) proviral load and disease progression in asymptomatic HTLV-1 carriers: a nationwide prospective study in Japan. Blood 116: 1211–1219.
14. Okamoto T, Ohno Y, Tsugane S, Watanabe S, Shimoyama M, et al. (1989) Multi-step carcinogenesis model for adult T-cell leukemia. Japanese journal of cancer research : Gann 80: 191–195.
15. Matsuoka M, Jeang KT (2007) Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. Nat Rev Cancer 7: 270–280.
16. Yoshida M (2010) Molecular approach to human leukemia: isolation and characterization of the first human retrovirus HTLV-1 and its impact on tumorigenesis in adult T-cell leukemia. Proceedings of the Japan Academy Series B, Physical and biological sciences 86: 117–130.
17. Tian Y, Kobayashi S, Ohno N, Isobe M, Tsuda M, et al. (2011) Leukemic T cells are specifically enriched in a unique CD3(dim) CD7(low) subpopulation of CD4(+) T cells in acute-type adult T-cell leukemia. Cancer science 102: 569–577.
18. Sugamura K, Fujii M, Kannagi M, Sakitani M, Takeuchi M, et al. (1984) Cell surface phenotypes and expression of viral antigens of various human cell lines carrying human T-cell leukemia virus. International journal of cancer Journal international du cancer 34: 221–228.
19. Yamaguchi K, Kiyokawa T, Nakada K, Yul LS, Asou N, et al. (1988) Polyclonal integration of HTLV-I proviral DNA in lymphocytes from HTLV-I seropositive individuals: an intermediate state between the healthy carrier state and smouldering ATL. British journal of haematology 68: 169–174.
20. Mortreux F, Gabet AS, Wattel E (2003) Molecular and cellular aspects of HTLV-1 associated leukemogenesis in vivo. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK 17: 26–38.
21. Sasaki D, Doi Y, Hasegawa H, Yanagihara K, Tsukasaki K, et al. (2010) High human T cell leukemia virus type-1(HTLV-1) provirus load in patients with HTLV-1 carriers complicated with HTLV-1-unrelated disorders. Virology journal 7: 81.
22. Setoyama M, Katahira Y, Kanzaki T (1999) Clinicopathologic analysis of 124 cases of adult T-cell leukemia/lymphoma with cutaneous manifestations: the smouldering type with skin manifestations has a poorer prognosis than previously thought. The Journal of dermatology 26: 785–790.
23. Reinhold U, Abken H (1997) CD4+ CD7- T cells: a separate subpopulation of memory T cells? J Clin Immunol 17: 265–271.

24. Reinhold U, Abken H, Kukel S, Moll M, Muller R, et al. (1993) CD7- T cells represent a subset of normal human blood lymphocytes. J Immunol 150: 2081–2089.
25. Aandahl EM, Quigley MF, Moretto WJ, Moll M, Gonzalez VD, et al. (2004) Expansion of CD7(low) and CD7(negative) CD8 T-cell effector subsets in HIV-1 infection: correlation with antigenic load and reversion by antiretroviral treatment. Blood 104: 3672–3678.
26. Autran B, Legac E, Blanc C, Debre P (1995) A Th0/Th2-like function of CD4+CD7- T helper cells from normal donors and HIV-infected patients. J Immunol 154: 1408–1417.
27. Legac E, Autran B, Merle-Beral H, Katlama C, Debre P (1992) CD4+CD7-CD57+ T cells: a new T-lymphocyte subset expanded during human immunodeficiency virus infection. Blood 79: 1746–1753.
28. Schmidt D, Goronzy JJ, Weyand CM (1996) CD4+ CD7- CD28- T cells are expanded in rheumatoid arthritis and are characterized by autoreactivity. J Clin Invest 97: 2027–2037.
29. Willard-Gallo KE, Van de Keere F, Kettmann R (1990) A specific defect in CD3 gamma-chain gene transcription results in loss of T-cell receptor/CD3 expression late after human immunodeficiency virus infection of a CD4+ T-cell line. Proc Natl Acad Sci U S A 87: 6713–6717.

Cell Stem Cell
Short Article

Generation of Rejuvenated Antigen-Specific T Cells by Reprogramming to Pluripotency and Redifferentiation

Toshinobu Nishimura,[1] Shin Kaneko,[1,9,*] Ai Kawana-Tachikawa,[2] Yoko Tajima,[1] Haruo Goto,[1] Dayong Zhu,[2] Kaori Nakayama-Hosoya,[2] Shoichi Iriguchi,[1] Yasushi Uemura,[6] Takafumi Shimizu,[1] Naoya Takayama,[3,10] Daisuke Yamada,[7] Ken Nishimura,[8] Manami Ohtaka,[8] Nobukazu Watanabe,[4] Satoshi Takahashi,[5] Aikichi Iwamoto,[2] Haruhiko Koseki,[7] Mahito Nakanishi,[8] Koji Eto,[3,10] and Hiromitsu Nakauchi[1,*]

[1]Division of Stem Cell Therapy, Center for Stem Cell Biology and Regenerative Medicine
[2]Division of Infectious Diseases, Advanced Clinical Research Center
[3]Stem Cell Bank, Center for Stem Cell Biology and Regenerative Medicine
[4]Laboratory of Diagnostic Medicine, Center for Stem Cell Biology and Regenerative Medicine
[5]Division of Molecular Therapy, Advanced Clinical Research Center
The Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
[6]Division of Immunology, Aichi Cancer Center Research Institute, 1-1 Kanakoden, Chikusa-ku Nagoya, Aichi 464-8681, Japan
[7]Laboratory for Lymphocyte Development, RIKEN Center for Allergy and Immunology, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
[8]Research Center for Stem Cell Engineering, National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Central 4, Tsukuba, Ibaraki 305-8562, Japan
[9]Present address: Department of Fundamental Cell Technology, Center for iPS Cell Research and Application (CiRA), Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan
[10]Present address: Department of Clinical Application, CiRA, Kyoto University, 53 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan
*Correspondence: kaneko.shin@cira.kyoto-u.ac.jp (S.K.), nakauchi@ims.u-tokyo.ac.jp (H.N.)
http://dx.doi.org/10.1016/j.stem.2012.11.002

SUMMARY

Adoptive immunotherapy with functional T cells is potentially an effective therapeutic strategy for combating many types of cancer and viral infection. However, exhaustion of antigen-specific T cells represents a major challenge to this type of approach. In an effort to overcome this problem, we reprogrammed clonally expanded antigen-specific $CD8^+$ T cells from an HIV-1-infected patient to pluripotency. The T cell-derived induced pluripotent stem cells were then redifferentiated into $CD8^+$ T cells that had a high proliferative capacity and elongated telomeres. These "rejuvenated" cells possessed antigen-specific killing activity and exhibited T cell receptor gene-rearrangement patterns identical to those of the original T cell clone from the patient. We also found that this method can be effective for generating specific T cells for other pathology-associated antigens. Thus, this type of approach may have broad applications in the field of adoptive immunotherapy.

INTRODUCTION

T cells play a central role in acquired immunity and the configuration of systemic immunity against pathogens. In particular, cytotoxic T lymphocytes (CTLs) are major components of this systemic response to microorganisms, viral infections, and neoplasms (Greenberg, 1991; Zhang and Bevan, 2011). T cells initiate their proliferative and effector functions upon human leukocyte antigen (HLA)-restricted recognition of specific antigen peptides via T cell receptors (TCRs). This is greatly beneficial in enabling the selective recognition and eradication of target cells, and also in long-term immunological surveillance by long-lived memory T cells (Butler et al., 2011; Jameson and Masopust, 2009; MacLeod et al., 2010). However, viruses in chronic infection or cancers often hamper or escape the T cell immunity by decreasing the expression of molecules required for T cell recognition or by inhibiting antigen presentation (Virgin et al., 2009). In addition, continuous exposure to chronically expressed viral antigens or cancer/self-antigens can drive T cells into an "exhausted" state. This is characterized by loss of effector functions and the potential for long-term survival and proliferation, ultimately leading to the depletion of antigen-responding T cell pools (Klebanoff et al., 2006; Wherry, 2011).

The infusion of ex vivo-expanded autologous antigen-specific T cells is being developed clinically for T cell immunotherapy. However, up to now, highly expanded T cells have not proven to be particularly effective (June, 2007). This is in part explained by losses of function that occur during the ex vivo manipulation of patient autologous T cells. In another instance, genetic modification of antigen receptors is an ambitious but only partially successful way to add desired antigen specificity to nonspecific T cells (Morgan et al., 2006; Porter et al., 2011). The therapeutic effect also strongly depends on the extent of functional loss that occurs during the ex vivo manipulation of T cells and on the stability of exogenous antigen receptor expression specific to target molecules in the presence of the endogenous TCR genes (Bendle et al., 2010; Brenner and Okur, 2009).

For the purpose of overcoming these obstacles, the therapeutic potential of induced pluripotent stem cells (iPSCs) is being

explored. Embryonic stem cells (ESCs) or iPSCs have the capacity for self-renewal while maintaining pluripotency (Takahashi et al., 2007) and could potentially form the basis for the unlimited induction of antigen-specific juvenile T cells. However, there are challenges to this approach as well, given that methods for the differentiation and immunological education of ESCs and iPSCs, or indeed that of intermediate hematopoietic stem and/or progenitor cells, into fully matured functional human T cells are not well established (Timmermans et al., 2009). Reprogramming the nuclei of lymphocytes was historically performed for studying whether terminally differentiated or fully matured somatic cells could revert to a pluripotent state. The first demonstration of lymphocyte reprogramming employed somatic cell nuclear transfer in murine B and T cells, proving that terminally differentiated somatic cells were reprogrammable (Hochedlinger and Jaenisch, 2002). Reprogramming murine B cells into pluripotent stem cells by iPSC technology also provided definite proof for fate reversibility in fully matured somatic cells (Hanna et al., 2008). From another point of view, nuclear reprogramming of lymphocytes is seen as having applications for regenerative medicine different than those for scientific research. The irreversible rearrangement of genes encoding immunoglobulins and TCRs was recognized solely as a genetic marker in somatic cell nuclear transfer and iPSC research. However, the preserved rearrangements in genomic DNA can also provide a blueprint of "educated" weapons for attacking cancers and pathogens in adoptive immunotherapy. Although several groups have reported the generation of T cell-derived iPSCs (T-iPSCs), their clinical applications have yet to be thoroughly explored (Brown et al., 2010; Loh et al., 2010; Seki et al., 2010; Staerk et al., 2010).

In the present study, we chose a T cell clone specific to an HIV type 1 (HIV-1) epitope of known structure to act as a generic representation of iPSC-mediated T cell regeneration. We successfully induced iPSCs from antigen-specific T cells and redifferentiated them into functional T cells. This may act as proof of concept for the application of "rejuvenated" T cells in treating various diseases. Crucial to this concept was that T-iPSCs retained the assembled "endogenous" TCR genes even after being subjected to nuclear reprogramming. Furthermore, redifferentiated T cells showed the same pattern of TCR gene arrangement as that in the original T cells. These features may therefore serve as the foundation for the reproduction of unlimited numbers of T cells that express desired TCRs conferring to antigen specificity.

RESULTS

Reprogramming an Antigen-Specific Cytotoxic T Cell Clone into Pluripotency

To establish T cell-derived iPSCs, we magnetically separated the $CD3^+$ T cell population from peripheral blood mononuclear cells (PBMCs) of healthy volunteers. The isolated $CD3^+$ T cells were stimulated with human CD3 and CD28 antibody-coated microbeads (α-CD3/28 beads) in the presence of interleukin-2 (IL-2). We then transduced the activated $CD3^+$ T cells with separate retroviral vectors that individually code for *OCT3/4*, *SOX2*, *KLF4*, and c-*MYC*. Human ESC-like colonies were obtained within 25 days of culture (Figure S1A available online).

We also isolated PBMCs from an HLA-A24-positive patient with a chronic HIV-1 infection. $CD8^+$ CTL clones specific for an antigenic peptide (amino acids [aa] 138–145) from the HIV-1 Nef protein (Nef-138-8(WT); RYPLTFGW) (Altfeld et al., 2006) were established. One of the clones, H25-#4, was stimulated using α-CD3/28 beads in the presence of IL-2 and then transduced simultaneously with six retroviral vectors encoding *OCT3/4*, *SOX2*, *KLF4*, c-*MYC*, *NANOG*, and *LIN28A*. However, we could not reprogram H25-#4 into pluripotency, possibly due to the cells being in a low infectious and exhausted state, or due to insufficient expression of the reprogramming factors. In response, we attempted to increase transduction efficiency and transgene expression by using two Sendai virus (SeV) vectors. One of them encodes tetracistronic factors (*OCT3/4*, *SOX2*, *KLF4*, and *c-MYC*) (Nishimura et al., 2011) with the miR-302 target sequence (SeVp[KOSM302L]; K.N., M.O., and M.N., data not shown), and another encodes SV40 large T antigen (SeV18[T]) (Fusaki et al., 2009). After transduction of phytohemagglutinin (PHA)-activated H25-#4 cells with the SeV vectors in the presence of IL-7 and IL-15, sufficient numbers of human ESC-like colonies appeared within 40 days of culture (Figure 1A). Use of this SeV system and optimization of transduction conditions greatly improved the reprogramming efficiency. It enabled us to reprogram several $CD8^+$ or $CD4^+$ T cell clones specific to pp65 antigen in cytomegalovirus (CMV), glutamic acid decarboxylase (GAD) antigen in type 1 diabetes, and α-GalCer (Table 1).

The resultant $CD3^+$ T cell- and H25-#4-derived ESC-like colonies (TkT3V1-7 and H254SeVT-3, respectively) exhibited alkaline phosphatase (AP) activity and expressed the pluripotent cell markers SSEA-4, Tra-1-60, and Tra-1-81 (Figures S1B–S1E and 1B–1E). H254SeVT-3 expressed HLA-A24 (Figure 1F). Both TkT3V1-7 and H254SeVT-3 also expressed human ESC-related genes (Figures S1F and 1G). The expression of exogenous reprogramming factors from the integrated provirus (TkT3V1-7) was halted (Figure S1F), and nonintegrated SeV genomic RNA was successfully removed from the cytosol by RNAi or by self-degradation caused by temperature-sensitive mutations (H254SeVT-3) (Figure 1H). Comparison of gene-expression profiles revealed that the gene-expression patterns in the ESC-like cells were similar to those in human ESCs, but differed significantly from those in peripheral blood (PB) T cells (Figure S1G). Scant methylation of the *OCT3/4* and *NANOG* promoter regions was confirmed using bisulfite PCR, thus indicating successful reprogramming (Freberg et al., 2007) (Figures S1H and 1I). In addition, when injected into nonobese diabetic severe combined immunodeficient (NOD-Scid) mice, those cells formed teratomas containing characteristic tissues derived from all three germ layers, which is indicative of pluripotency (Brivanlou et al., 2003) (Figures S3 and 2A). Therefore, those colonies were confirmed as typical human iPSCs.

T-iPSCs Carry Preassembled TCR Genes from the Original T Cell

Almost all TCRs are composed of heterodimerically associated α and β chains. *TCRA* or *TCRB* gene (encoding α chain or β chain, respectively) rearrangements are involved in normal αβ T cell development in the thymus. These rearrangements enabled us to determine retrospectively whether the iPSCs were derived from an αβ T cell. The BIOMED-2 consortium designed multiplex-PCR primers for analyzing *TCRB* gene assemblies (van Dongen et al., 2003), and we designed the primers for detecting *TCRA*

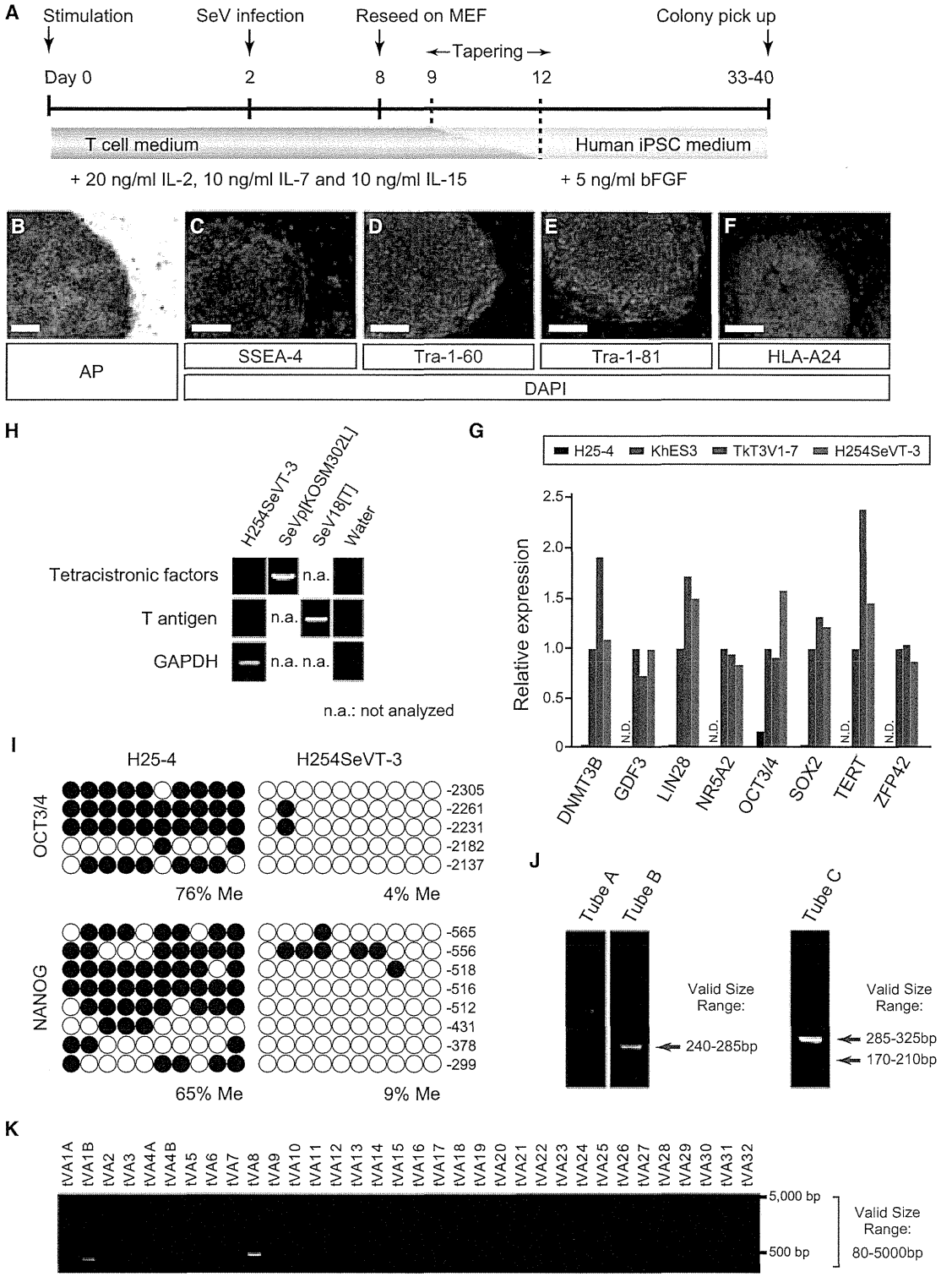

Figure 1. Generation of Human iPSCs from a CTL Clone

(A) Schematic illustration showing the generation of T-iPSCs from H25-#4 T cells using SeV vectors encoding polycistronic *OCT3/4*, *SOX2*, *KLF4*, and c-*MYC*, or SV40 large T antigen. The "tapering" indicates the gradual replacement of culture medium with human iPSC medium.

(B–F) AP activity (B) and expression of pluripotency markers (SSEA-4, C; Tra-1-60, D; and Tra-1-81, E) and HLA-A24 (F) in H254SeVT-3 cells. Nuclei were counterstained with DAPI. The scale bar represents 200 μm.

(G) Quantitative PCR for pluripotency genes in H25-#4, KhES3, TkT3V1-7, and H254SeVT-3 cells. Individual PCR reactions were normalized against 18S ribosomal RNA (rRNA).

(legend continued on next page)

Table 1. Generation of Human T-iPSCs from Various Patient-Derived T Cell Specimens

Antigen	T Cell Source	Initial Cell Number	No. of ESC-like Colonies	No. of Colonies Picked up for Establishing T-iPSC Clones	Date (MM/YYYY)
HIV-1 Nef	monoclonal T cell clone	4×10^5	7	7	05/2011
CMV pp65	polyclonal tetramer-sorted cells	~5,000	15	15	07/2011
GAD	monoclonal T cell clone	1×10^6	>100	not picked up	08/2012
		5×10^5	>100	19	08/2012
α-GalCer	FACS-sorted Vα24$^+$ cells	1×10^6	>100	not picked up	08/2012
		5×10^5	>100	7	08/2012

Sample cells were transduced with *OCT3/4*, *SOX2*, *KLF4*, *c-MYC*, and SV40 large T-antigen by using two Sendai virus (SeV) vectors (SeVp [KOSM302L] and SeV18[T]). After around 40 days, the number of embryonic stem cell (ESC)-like colonies were counted on the basis of morphology and alkaline phosphatase (AP) activity. All established T cell-derived induced pluripotent stem cell (T-iPSC) lines were free from residual SeV vectors (one example in the case of the HIV-1 Nef-specific T-iPSC clone is shown in Figure 1H). CMV, cytomegalovirus; GAD, glutamic acid decarboxylase; FACS, fluorescence-activated cell sorting.

gene assemblies (Figure S2). *TCRB* and *TCRA* gene assemblies were identified as single bands representing each allele in TkT3V1-7 and H254SeVT-3 (Figures S1H, S1I, 1J, and 1K).

We next confirmed the presence of an antigen-recognition site on the TCR that consisted of three complementarity-determining regions (CDR1, CDR2, and CDR3). CDR3 is the most diversifiable among the three because it spans the V(D)J-junction region, where several random nucleotides (N or P nucleotides) are inserted (Alt and Baltimore, 1982; Lafaille et al., 1989). We determined the CDR3 sequences of the assembled *TCRA* and *TCRB* genes in TkT3V1-7 and H254SeVT-3 and identified a set of productive *TCRA* and *TCRB* gene rearrangements (i.e., in-frame junction with no stop codon) (Table S1 and Table 2). Furthermore, the sequences of CDR3 from H254SeVT-3 and H25-#4 were completely identical at both *TCRA* and *TCRB* gene loci. These results indicated that the iPSCs established were derived from a single T cell and that the antigen specificity encoded in the genomic DNA of the T cell was conserved during reprogramming.

Redifferentiation of T-iPSCs into CD8 Single-Positive T Cells Expressing the Desired TCR

Following the application of specific in vitro differentiation protocols, iPSCs can give rise to mesoderm-derived cell types, especially hematopoietic stem and/or progenitor cells (Takayama et al., 2008; Vodyanik et al., 2005) (Figure 2B). This was applied to assess the capacity of T-iPSCs for hematopoietic differentiation by coculturing on C3H10T1/2 feeder cells in the presence of VEGF, SCF, and FLT-3L for the generation of CD34$^+$ hematopoietic stem and/or progenitor cells. On day 14 of culture, the cells were transferred onto Delta-like 1-expressing OP9 (OP9-DL1) feeder cells (Timmermans et al., 2009) and were cocultured in the presence of FLT-3L and IL-7 (Ikawa et al., 2010) (Figure 2B). After 21–28 days of culture, the hematopoietic cells differentiated into CD45$^+$, CD38$^+$, CD7$^+$, CD45RA$^+$, CD3$^+$, and TCRαβ$^+$ T lineage cells (Figure S4). As was the case with TCRαβ transgenic mice (Borgulya et al., 1992) and chimeric mice derived from ESCs produced through nuclear transplantation of T cells (Serwold et al., 2007), aberrant expression of TCRαβ was observed at the CD4/CD8 double-negative (DN) stage. Although some of these T lineage cells differentiated into the CD4/CD8 double-positive (DP) stage and the more mature CD4 or CD8 single-positive (SP) stages (Figure 2C), we could not characterize the small number of SP cells in more detail.

During thymocyte development, the CD4/CD8 DN and DP stages correspond respectively to the *TCRB*-encoded β chain and *TCRA*-encoded α chain assembly stages (von Boehmer, 2004). In the *TCRB* locus, the negative-feedback regulation of gene assembly and the capacity to deter further rearrangement are very strict (Khor and Sleckman, 2002). In the *TCRA* locus, by contrast, negative-feedback regulation is relatively loose, and further gene assembly of the preassembled gene, a phenomenon known as "receptor revision," tends to occur (Huang and Kanagawa, 2001; Krangel, 2009). In experiments using TCRα transgenic mice, the reactivation of *Rag1* and *Rag2*, genes related to recombination machinery, occurred in CD4/CD8 DP-stage thymocytes, and gene assembly of endogenous *Tcra* was observed (Padovan et al., 1993; Petrie et al., 1993). Such further gene assembly would be exceedingly undesirable for our purposes, because it would probably convert the tropism of the TCR and render the redifferentiated T cells incapable of attacking the previously targeted antigen. To determine whether such receptor revision could occur in redifferentiating T lineage cells, we collected CD1a$^-$ DN- and CD1a$^+$ DP-stage cells from among the CD45$^+$, CD3$^+$, TCRαβ$^+$, and CD5$^+$ T lineage cells and then analyzed the gene rearrangement of TCR messenger RNAs (mRNAs) (Figures S5A–S5C). Nucleotide sequences of *TCRB* mRNAs in the T lineage cells were identical to those in

(H) Detection of the remnants of SeV genomic RNAs by RT-PCR. Each column represents the template cDNA synthesized from H254SeVT-3 cells, SeVp [KOSM302L] virus solution, and SeV18[T] virus solution. cDNAs from virus solution were the positive controls.
(I) Bisulfite sequencing analyses of the *OCT3/4* and *NANOG* promoter regions in H25-#4 and H254SeVT-3 cells. White and black circles represent unmethylated and methylated (Me) CpG dinucleotides, respectively.
(J) Multiplex PCR analysis to detect *TCRB* gene rearrangements in the H254SeVT-3 genome. Tubes A and B contain Vβ-(D)Jβ assemblies; Tube C contains D-Jβ assemblies.
(K) Multiplex PCR analysis for detection of *TCRA* gene rearrangements (V-Jα assemblies).
See Figures S1, S2, and S3 for additional data.

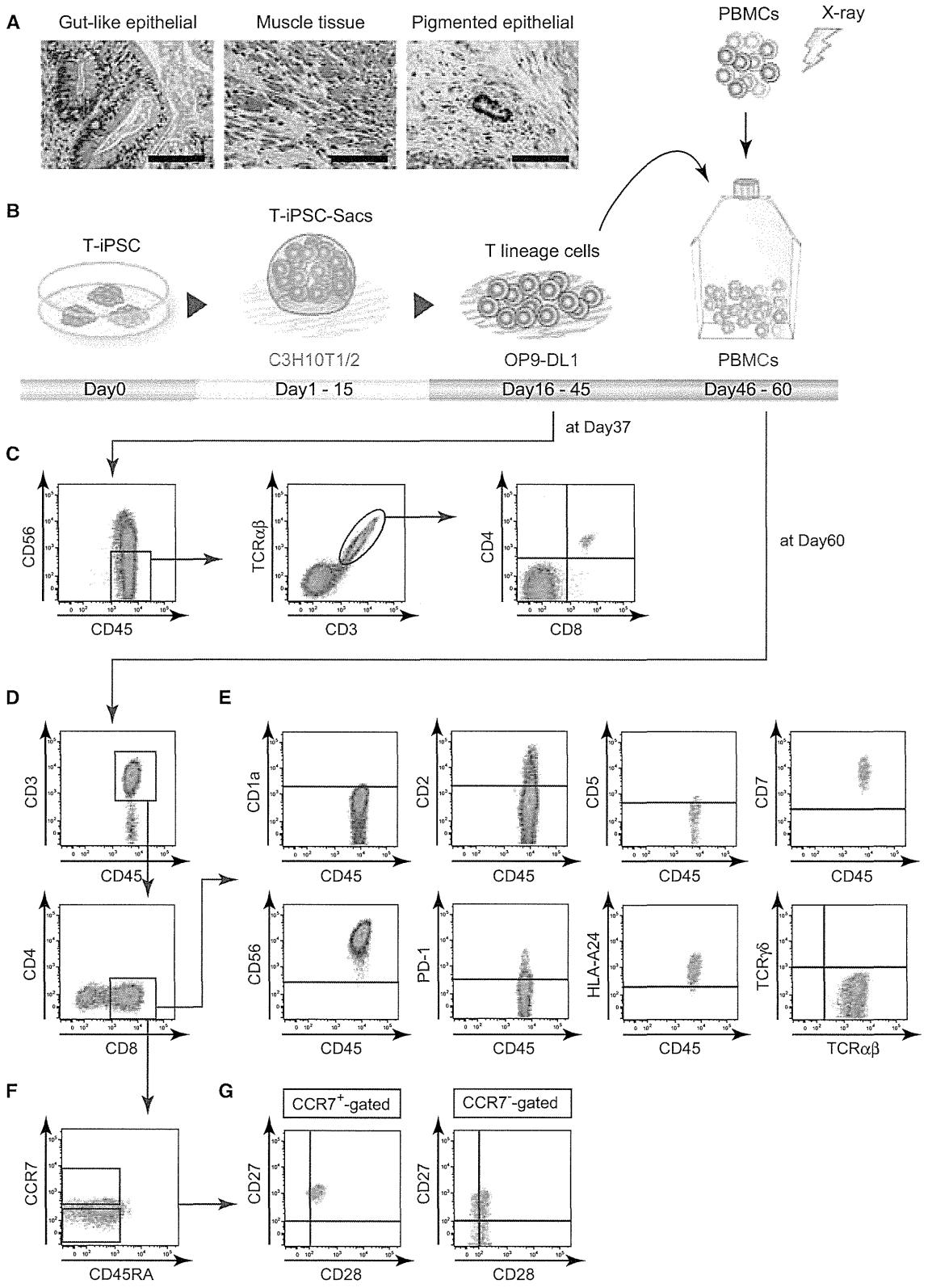

Figure 2. Redifferentiation of T-iPSCs into T Cells

(A) Representative hematoxylin- and eosin-stained sections of a teratoma formed in a NOD/ShiJic-*scid* mouse testis. H254SeVT-3 differentiated into cell lineages derived from endoderm (goblet cells in gut-like epithelial), mesoderm (smooth myocytes in muscle tissue), and ectoderm (retina cells in pigmented epithelial). The scale bar represents 100 μm.

(legend continued on next page)

the T-iPSCs at both the DN and DP stages. By contrast, some *TCRA* mRNAs at the DN and DP stages were identical to those in the T lineage cells, but others differed, and differing sequences were observed more frequently at the DP stage than the DN stage (Table S2). *RAG1* and *RAG2* expression were observed at both the DN and the DP stages, though stronger expression was observed at the DP stage (Figure S5D).

To create mature CD8 SP cells from T-iPSC-derived T lineage cells without receptor revision, we focused on TCR signaling. Turka et al. (1991) reported that TCR signaling via peptide-major histocompatibility complex (MHC) complexes during positive selection ends expression of *RAG* genes and prevents further assembly of TCR genes. They also showed that mimicking TCR signaling using CD3 antibodies had the same effect. Therefore, we tried to stimulate the TCRs of redifferentiating T lineage cells before the completion of the DN-to-DP transition (Figure 2B). For this experiment, we cultured T lineage-committed cells on OP9-DL1, stimulated them with α-CD3/28 beads or PHA (we defined this as the first stimulation) and then cocultured them with irradiated HLA-A24$^-$ PMBCs in the presence of IL-7 and IL-15, which are required for the generation of memory phenotype CD8$^+$ T cells (Kaneko et al., 2009; Prlic et al., 2002; Tan et al., 2002). After 14 days, CD8 SP cells appeared (Figure 2D). These were deemed to be derivatives of H254SeV-3 based on their expression of HLA-A24 (Figure 2E). These CD8 SP cells did not express the immature thymocyte marker CD1a, but they were positive for CD56, which is expressed on CD8$^+$ T cells cultured in vitro (Lu and Negrin, 1994). In addition, these cells expressed CD7 and some CD2, but not CD5. On the one hand, they did not express PD-1, a marker of exhausted T cells (Figure 2E). On the other hand, some of them expressed the memory T cell markers CCR7, CD27, and CD28 simultaneously, thus representing a central memory T cell phenotype (Figures 2F and 2G) (Romero et al., 2007).

To test whether the redifferentiated CD8 SP cells would recognize the same epitope on the same HLA, the entire population of redifferentiated T cells was mixed with the A24/Nef-138-8(WT) tetramer and subjected to flow-cytometric analysis (Kawana-Tachikawa et al., 2002). Most of the CD8 SP cells were stained positively by the A24/Nef-138-8(WT) tetramer, but not by the control tetramer, which represents HIV-1 envelope-derived peptides (RYLRDQQLL; Figure 3A and data not shown). We then collected the A24/Nef-138-8(WT) tetramer-reactive CD8$^+$ cells and expanded them once again using α-CD3/28 beads or PHA stimulation (defined as the second stimulation; Figure 3A). Finally, after several independent redifferentiation experiments, we obtained A24/Nef-138-8(WT) tetramer-reactive CD8 SP cells (reT-1, reT-2.1, reT-2.2, and reT-3). As expected, sequence analysis of *TCRA* and *TCRB* mRNAs in the redifferentiated CD8 SP cells revealed that the TCR gene rearrangement pattern was identical to that in the H25-#4 original T cell clone (Figure 3B and Table 1).

To determine whether the redifferentiated CD8 SP cells were of the T cell lineage, we used quantitative PCR to compare gene-expression profiles among redifferentiated CD8 SP cells, PB CD4$^+$ and CD8$^+$ T cells, and the H25-#4 original T cell clone. As shown in Figure 3C, the expression patterns of CD3, CD4, and CD8 were similar among PB CD8$^+$ T cells, redifferentiated CD8 SP cells, and the H25-#4 original T cell clone. However, the pattern differed from those in PB CD4$^+$ T cells (Figure 3C). Cytotoxic "signature" genes such as granzyme B (*GZMB*), perforin (*PRF1*), interferon-γ (IFN-γ; *IFNG*), and FAS ligand (*FASLG*) were expressed in PB CD8$^+$ T cells. These genes were also expressed relatively strongly in redifferentiated CD8 SP cells and in the H25-#4 original T cell clone; that is, in already-primed T cells (Figure 3D). The expression patterns of several factors involved in transcription or signal transduction and of cell-surface molecules were similar among PB CD8$^+$ T cells, redifferentiated CD8 SP cells, and the H25-#4 original T cell clone (Figure 3E). To exclude the possibility that the redifferentiated CD8 SP cells had acquired natural killer (NK)-like properties during their coculture with OP9-DL1 or PBMCs, we used a complementary DNA (cDNA) microarray to analyze global gene-expression profiles in redifferentiated CD8 cells, the H25-#4 original T cell clone, and PB NK cells. Correlation and cluster analyses of the gene-expression profile of the redifferentiated CD8 SP cells showed it to be similar to that of the H25-#4 original T cell clone but different from that of NK cells (Figures 3F and 3G). These data strongly suggest that T-iPSCs are able to redifferentiate into CD8$^+$ T cells that exhibit the same antigen specificity as that of the original T cell.

Generation of Highly Proliferative T Cells through T-iPSCs

Fewer than 10^5 T lineage cells were obtained from ~3 × 10^5 T-iPSCs after coculture with C3H10T1/2 and OP9-DL1 cells. However, they could be expanded to >10^8 cells with the first stimulation (data not shown). After separating A24/Nef-138-8(WT) tetramer-reactive CD8$^+$ cells, we assessed the expansion rate induced by the second stimulation and also assessed the establishment of reT-1, reT-2.2, and reT-3. We found that these cells expanded from 100-fold to 1,000-fold within 2 weeks in the presence of IL-7 and IL-15, whereas the H25-#4 original T cell clone expanded only about 20-fold (Figure 4A). Even after 100-to-1,000-fold expansions, some cells still expressed central memory T cell markers such as CCR7, CD27, and CD28 (Figure S6). Perhaps with passage through the iPSC state, wherein telomerase activity is quite high (Marion et al., 2009; Takahashi et al., 2007), re-elongation of shortened telomeres in the H25-#4 original T cell clone gives the redifferentiated T cells high replicative potential (Monteiro et al., 1996; Weng et al., 1998). In fact, the redifferentiated T cells carried longer telomeres than the original T cell clone (Figure 4B), an overall process that we call

(B) Schematic illustration of redifferentiation from T-iPSCs into T cells.
(C) Flow-cytometric analysis of the phenotypes of differentiating T lineage cells at 37 days after starting redifferentiation.
(D and E) Flow-cytometric analysis of the phenotypes of T cells at 60 days after starting redifferentiation. Fluorescence-activated cell sorting (FACS) analyses revealed CD8 single-positive maturation (D) and expression of several T cell markers (E).
(F and G) Memory phenotypes of redifferentiated CD8$^+$ T cells. There existed memory-phenotyped cells such as all positive for CCR7 (F), CD27, and CD28 (G). Data are representative of at least three independent experiments. See Figures S3, S4, and S5 and Table S2 for additional data.