

Supplementary Table 1: Coverage depth of Exome-Seq, RNA-Seq, and Whole Genome Sequencing

| Sample | Exome Sequencing | RNA-Seq | Whole Genome Sequencing |
|--------|------------------|---------------------------------|-------------------------|
| | Coverage Depth | Millions of Unique Mapped Reads | Coverage Depth |
| 1 | 212.296 | - | - |
| 2 | 229.655 | 79.41 | - |
| 3 | 221.612 | - | - |
| 4 | 191.248 | - | - |
| 5 | 237.056 | - | - |
| 6 | 159.702 | 119.08 | - |
| 7 | 232.187 | - | - |
| 8 | 206.878 | - | - |
| 9 | 238.57 | 28.76 | - |
| 10 | 176.488 | - | - |
| 11 | 200.226 | - | - |
| 12 | 225.346 | 188.08 | - |
| 13 | 148.801 | 194.55 | 33.97 |
| 14 | 333.623 | - | - |
| 15 | 207.053 | 86.94 | - |
| 16 | 256.855 | - | - |
| 17 | 212.514 | 200.65 | 32.13 |
| 18 | 228.644 | - | - |
| 19 | 142.219 | 93.12 | - |
| 20 | 251.65 | - | - |
| 21 | 219.13 | - | - |
| 22 | 211.489 | - | - |
| 23 | 313.064 | - | - |
| 24 | 235.35 | - | - |
| 25 | 230.113 | - | - |
| 26 | 277.415 | 86.38 | - |
| 27 | 187.807 | - | - |
| 28 | 256.855 | - | - |
| 29 | 227.849 | - | - |
| 30 | 185.597 | - | - |
| 31 | 220.994 | - | - |
| 32 | 152.91 | 76.53 | - |
| 33 | 257.772 | - | - |
| 34 | 209.109 | - | - |
| 35 | 307.703 | 42.41 | - |
| 36 | 168.945 | - | - |
| 37 | 236.272 | - | - |
| 38 | 190.454 | - | - |
| 39 | 202.544 | - | - |
| 40 | 169.274 | - | - |

Supplementary Table 2. C>T mutations in ECP-naïve and ECP-experienced patients.

| ECP status | Total # SSNVs | Total # C>T SSNVs | % C>T dipyrimidine | % C>T non-dipyrimidine |
|-------------------|--------------------------|---------------------------------|----------------------------------|--------------------------------------|
| ECP-neg (n = 3) | 127 | 83 | 80 | 20 |
| ECP-pos (n = 37) | 2,722 | 1,933 | 76 | 24 |

ECP-neg, ECP-naïve; ECP-pos, ECP-experienced.

Supplementary Table 3a. Concordance of calls between exome sequencing and WGS (n=2)

| | |
|--|-------|
| Number of exome somatic variants with >10 coverage in WGS | 120 |
| Number of exome somatic variants with mutant base in WGS | 119 |
| % concordance of somatic variants between exome sequencing and WGS | 99.4% |

Table 3b. Concordance of calls between exome sequencing and RNA-seq (n=11)

| | |
|--|-----|
| Number of exome somatic variants with >20x coverage in RNA-seq | 192 |
| Number of exome somatic variants with mutant base in RNA-seq | 184 |
| % concordance of somatic variants between exome sequencing and RNA-seq | 96% |

Supplementary Table 4. SCNVs confirmed by whole genome sequencing.

| | |
|---|----|
| SCNVs identified by exome sequencing | 43 |
| SCNVs identified by WGS and Patchwork-R | 41 |
| SCNVs identified by WGS and Breakdancer | 1 |

Supplementary Table 5. Read counts supporting SSNVs in putative driver genes in CTCL.

| Sample | Tumor Purity | Gene | AA change | Position (hg18) | CTCL | | | Monocyte | | |
|--------|--------------|---------------|-----------|-------------------|---------|------------|------|----------|------------|------|
| | | | | | Ref cov | Nonref cov | MAF | Ref cov | Nonref cov | MAF |
| 19 | 93.40% | <i>ARID1A</i> | N1705fs | chr1:26,974,774 | 118 | 53 | 31% | 300 | 1 | 0.3% |
| 15 | 99.70% | <i>ARID1A</i> | Q1454X | chr1:26,973,665 | 49 | 53 | 52% | 113 | 0 | 0.0% |
| 14 | 95.00% | <i>ATM</i> | G2694A | chr11:107,710,976 | 21 | 203 | 91% | 170 | 1 | 0.6% |
| 13 | 99.30% | <i>BRAF</i> | D594N | chr7:140,099,624 | 197 | 72 | 27% | 283 | 1 | 0.4% |
| 5 | 99.80% | <i>CD28</i> | Q77P | chr2: 204,299,778 | 179 | 140 | 44% | 287 | 0 | 0.0% |
| 5 | 99.80% | <i>CD28</i> | K81N | chr2:204,299,791 | 179 | 140 | 44% | 287 | 0 | 0.0% |
| 14 | 95.00% | <i>CD28</i> | F51V | chr2:204,299,699 | 177 | 149 | 46% | 137 | 1 | 0.7% |
| 10 | 99.60% | <i>CD28</i> | F51I | chr2:204,299,699 | 74 | 82 | 53% | 202 | 2 | 1.0% |
| 11 | 99.00% | <i>CDKN2A</i> | W15X | chr9:21,964,783 | 0 | 10 | 100% | 14 | 1 | 6.7% |
| 12 | 99.70% | <i>CTCF</i> | L309V | chr16:66,203,498 | 141 | 87 | 38% | 256 | 0 | 0.0% |
| 32 | 99.00% | <i>DNMT3A</i> | P777L | chr2:25,315,581 | 115 | 86 | 43% | 224 | 6 | 2.6% |
| 32 | 99.00% | <i>DNMT3A</i> | S669F | chr2:25,318,011 | 72 | 75 | 51% | 198 | 5 | 2.5% |
| 11 | 99.00% | <i>DNMT3A</i> | Y584X | chr2:25,320,627 | 0 | 35 | 100% | 52 | 2 | 3.7% |
| 30 | 97.80% | <i>DNMT3A</i> | P233L | chr2:25,324,567 | 55 | 25 | 31% | 93 | 1 | 1.1% |
| 5 | 99.80% | <i>FAS</i> | E261K | chr10:90,763,960 | 111 | 177 | 61% | 392 | 0 | 0.0% |
| 13 | 99.30% | <i>FAS</i> | D265E | chr10:90,763,974 | 3 | 85 | 97% | 259 | 0 | 0.0% |
| 8 | 99.60% | <i>FAS</i> | S212C | chr10:90,761,802 | 10 | 125 | 93% | 273 | 0 | 0.0% |
| 32 | 99.00% | <i>IRF4</i> | F358I | chr6:346,753 | 64 | 35 | 35% | 98 | 0 | 0.0% |
| 11 | 99.00% | <i>IRF4</i> | F359L | chr6:346,755 | 67 | 34 | 34% | 97 | 5 | 4.9% |
| 37 | 99.60% | <i>NFKB2</i> | SS* | chr10:104,150,692 | 53 | 30 | 36% | 41 | 0 | 0.0% |
| 7 | 97.60% | <i>PLCG1</i> | E1163K | chr20:39,235,798 | 125 | 60 | 32% | 170 | 1 | 0.6% |
| 6 | 99.60% | <i>PLCG1</i> | D342N | chr20:39,225,988 | 103 | 104 | 50% | 165 | 1 | 0.6% |
| 14 | 95.00% | <i>PLCG1</i> | R48W | chr20:39,199,837 | 41 | 37 | 47% | 45 | 2 | 4.3% |
| 3 | 96.00% | <i>PLCG1</i> | S345F | chr20:39,225,998 | 118 | 87 | 42% | 191 | 0 | 0.0% |
| 33 | 97.00% | <i>RHOA</i> | N117I | chr3:49,374,991 | 173 | 157 | 48% | 153 | 1 | 0.6% |
| 13 | 99.30% | <i>RHOA</i> | N117I | chr3:49,374,991 | 124 | 128 | 51% | 212 | 4 | 1.9% |
| 4 | 99.20% | <i>RHOA</i> | R70K | chr3:49,380,933 | 146 | 114 | 44% | 150 | 0 | 0.0% |
| 21 | 97.60% | <i>STAT5B</i> | N642H | chr17:37,613,255 | 149 | 119 | 44% | 198 | 5 | 2.5% |
| 32 | 99.00% | <i>TP53</i> | R273P | chr17:7,517,845 | 48 | 131 | 73% | 280 | 4 | 1.4% |

| | | | | | | | | | | |
|----|---------|-------------|-------|------------------|-----|-----|------|-----|----|-------|
| 25 | 99.10% | <i>TP53</i> | A215V | chr17:7,514,712 | 11 | 81 | 88% | 80 | 14 | 14.9% |
| 24 | 98.30% | <i>TP53</i> | R196X | chr17:7,518,988 | 3 | 199 | 99% | 211 | 5 | 2.3% |
| 11 | 99.00% | <i>TP53</i> | S94X | chr17:7,520,131 | 1 | 50 | 98% | 177 | 1 | 0.6% |
| 20 | 100.00% | <i>TP53</i> | T155N | chr17:7,519,191 | 6 | 97 | 94% | 90 | 10 | 10.0% |
| 40 | 100.00% | <i>TP53</i> | I254T | chr17:7,518,245 | 0 | 45 | 100% | 68 | 8 | 10.5% |
| 4 | 99.20% | <i>TP53</i> | S34X | chr17:7,519,158 | 5 | 47 | 90% | 67 | 0 | 0.0% |
| 2 | 97.70% | <i>ZEB1</i> | Q855X | chr10:31,850,829 | 8 | 74 | 90% | 212 | 0 | 0.0% |
| 37 | 99.60% | <i>ZEB1</i> | Q538X | chr10:31,849,929 | 74 | 103 | 58% | 65 | 0 | 0.0% |
| 31 | 97.50% | <i>ZEB1</i> | Y251C | chr10:31,843,604 | 156 | 238 | 60% | 278 | 30 | 9.7% |

Ref Cov; number of reads of reference allele, Nonref Cov, number of reads of non-reference allele; MAF, minor allele frequency.

*Splice-site mutation upstream of exon 17.

Supplementary Table 6. SSNVs in CTCL identified in other cancers

| Gene | Mutations | Other Cancers |
|---------------|------------------|----------------------|
| <i>BRAF</i> | D594N | Melanoma** |
| <i>CD28</i> | F51V | AITL |
| <i>CARD11</i> | M360K | DLBCL |
| <i>CDKN2A</i> | W15X | HNSCC |
| <i>DNMT3A</i> | S669F | Lung |
| <i>FAS</i> | E261K | Lung |
| | D265E | DLBCL |
| <i>STAT5B</i> | N642H | T-LGL** |
| <i>TP53</i> | R273P | Breast |
| | A215V | CNS |
| | R196X | Breast |
| | S94X | Esophagus |
| | T155N | Lung |
| | I254T | Colon |
| | S34X | Lung |

Other cancer, cancer previously described to be harboring identical somatic SNV. AITL, angioimmunoblastic T-cell lymphoma; CNS, central nervous system; DLBCL, diffuse large B-cell lymphoma; HNSCC, head and neck squamous cell carcinoma; T-LGL, T- cell large granular lymphocytic leukemia. **functionally validated to be a gain-of-function mutation.

Supplementary Table 7. Significant Focal Deletions

| Cytoband | Q-Value | Residual Q-Value | Wide Peak Boundaries | % of CTCLs With Deletions [‡] | | | # Genes in Peak | Candidate Gene | GRAIL P-value |
|----------|------------------------|------------------------|-------------------------------|--|-------|-------|-----------------|----------------|---------------|
| | | | | Total | Focal | Broad | | | |
| 1p36.11 | 6.59x10 ⁻¹⁶ | 6.59x10 ⁻¹⁶ | chr1:26,673,679–27,110,886 | 57.5% | 57.5% | 0.0% | 7 | ARID1A* | - |
| 9p21.3 | 6.59x10 ⁻¹⁶ | 6.59x10 ⁻¹⁶ | chr9:21,849,450–21,996,041 | 40.0% | 37.5% | 2.5% | 1 | CDKN2A* | - |
| 10p11.22 | 4.57x10 ⁻¹⁴ | 2.01x10 ⁻¹³ | chr10:31,205,853–32,136,797 | 60.0% | 32.5% | 27.5% | 1 | ZEB1 | - |
| 10q24.32 | 1.74x10 ⁻⁸ | 4.77x10 ⁻⁷ | chr10:104,131,016–104,230,563 | 67.5% | 42.5% | 25.0% | 7 | NFKB2 | - |
| 2p23.3 | 5.57x10 ⁻⁷ | 5.57x10 ⁻⁷ | chr2:24,949,121–25,818,354 | 37.5% | 30.0% | 7.5% | 4 | DNMT3A* | - |
| 11q22.3 | 1.69x10 ⁻⁵ | 1.76x10 ⁻⁵ | chr11:107,234,547–107,759,139 | 30.0% | 22.5% | 7.5% | 5 | ATM* | - |
| 2q37.3 | 0.00010519 | 0.00010354 | chr2:239,018,134–242,951,149 | 20.0% | 20.0% | 0.0% | 38 | PDCD1 | 0.0035 |
| 19p13.3 | 1.19x10 ⁻⁷ | 0.00039797 | chr19:1–55,701,376 | 42.5% | 42.5% | 0.0% | 1078 | STK11** | - |
| 9q21.32 | 2.94x10 ⁻⁵ | 0.00048049 | chr9:80,386,282–90,795,694 | 32.5% | 17.5% | 15.0% | 33 | DAPK1 | 0.00062 |
| 13q14.2 | 0.00057769 | 0.00056364 | chr13:47,731,103–56,619,937 | 30.0% | 20.0% | 10.0% | 40 | RB1* | - |
| 19p13.3 | 1.35x10 ⁻⁷ | 0.000581 | chr19:2,029,483–3,927,502 | 32.5% | 32.5% | 0.0% | 53 | GADD45B | 0.025 |
| 12p13.2 | 0.00088263 | 0.00085553 | chr12:9,777,007–14,468,100 | 17.5% | 17.5% | 0.0% | 64 | CDKN1B* | - |
| 6q23.3 | 1.76x10 ⁻⁵ | 0.0016017 | chr6:138,243,081–138,525,040 | 25.0% | 25.0% | 0.0% | 2 | TNFAIP3*** | - |
| 10q21.2 | 0.00018739 | 0.00236 | chr10:63,698,286–64,053,033 | 52.5% | 17.5% | 35.0% | 2 | ZNF365 | 0.58 |
| 10q26.3 | 0.00010236 | 0.004758 | chr10:127,512,506–135,374,737 | 55.0% | 35.0% | 20.0% | 46 | MGMT | 0.00016 |
| 7p21.1 | 0.011766 | 0.011766 | chr7:15,566,451–36,519,345 | 27.5% | 22.5% | 5.0% | 107 | NFE2L3* | - |
| 9q31.1 | 0.00078864 | 0.022893 | chr9:96,260,592–105,897,449 | 25.0% | 20.0% | 5.0% | 58 | XPA | 0.39 |
| 6q25.2 | 0.00078369 | 0.04894 | chr6:153,645,511–170,899,992 | 20.0% | 20.0% | 0.0% | 73 | CCR6 | 0.12 |
| 12q21.33 | 0.048579 | 0.04894 | chr12:87,113,009–111,341,205 | 20.0% | 17.5% | 2.5% | 160 | SOCS2 | 0.076 |
| 6q21 | 0.052955 | 0.051525 | chr6:106,661,861–112,488,015 | 17.5% | 15.0% | 2.5% | 42 | TRAF3IP2 | 0.02 |
| 16q22.1 | 0.056991 | 0.055785 | chr16:66,074,546–66,236,753 | 15.0% | 15.0% | 0.0% | 2 | CTCF* | - |
| 10q26.11 | 0.00030478 | 0.095529 | chr10:120,779,987–124,142,656 | 55.0% | 35.0% | 20.0% | 19 | BAG3 | 0.46 |
| 10q23.31 | 0.015023 | 0.11663 | chr10:90,763,175–90,964,816 | 40.0% | 12.5% | 27.5% | 2 | FAS | - |
| 11p13 | 0.11475 | 0.11663 | chr11:31,441,250–35,596,790 | 20.0% | 10.0% | 10.0% | 28 | WT1* | - |

| | | | | | | | | | |
|----------|------------|---------|----------------------------------|-------|-------|-------|-----|---------|-------|
| 6q25.1 | 0.00050411 | 0.1338 | chr6:149,953,508– 150,112,805 | 22.5% | 22.5% | 0.0% | 3 | LATS1 | 0.17 |
| 16q24.2 | 0.13018 | 0.1338 | chr16:86,667,813–88,154,820 | 12.5% | 12.5% | 0.0% | 21 | BANP | 0.021 |
| 13q12.13 | 0.13373 | 0.1338 | chr13:1–40,031,653 | 17.5% | 10.0% | 7.5% | 101 | BRCA2** | - |
| 1p22.1 | 0.15633 | 0.15669 | chr1:71,317,019–96,960,088 | 15.0% | 15.0% | 0.0% | 106 | RPL5* | - |
| 8p23.1 | 0.24317 | 0.23282 | chr8:1–146,274,826 | 40.0% | 20.0% | 20.0% | 644 | EGR3** | - |

Bold, Genes with statistically significant gene localizing data. * Putative tumor suppressors published by the TCGA to be significantly mutated across tumors of multiple lineage types¹. **There are multiple consensus tumor suppressors in the GISTIC interval, the gene with the highest SCNv mutation rate is listed. ****TNFAIP3* was listed because it was recently implicated in leukemic CTCL². The remaining candidate genes reflect the highest-scoring gene in the interval identified by G-RAIL. G-RAIL P-values are corrected for multiple hypothesis testing. †Only the focal SCNvs are listed for CTCL samples with both focal SCNvs and broad SCNvs involving the indicated GISTIC interval.

Supplementary Table 8. Significant Focal Amplifications

| Cytoband | Q-value | Residual Q-value | Wide Peak Boundaries | % of CTCLs With Amplifications [‡] | | | # Genes in Peak | Candidate Gene | GRAIL P-value |
|----------|------------|---------------------|----------------------------------|--|-------|-------|--------------------|-------------------|-------------------------|
| | | | | Total | Focal | Broad | | | |
| 10p15.1 | 0.00029474 | 0.00029474 | chr10:5,882,522–7,245,908 | 30.0% | 17.5% | 12.5% | 9 | PRKCQ | 6.78 x 10 ⁻⁵ |
| 10p12.33 | 0.063691 | 0.063691 | chr10:15,250,463–18,469,816 | 22.5% | 10.0% | 12.5% | 21 | TRDMT1 | 0.18 |
| 7p22.2 | 0.098621 | 0.098621 | chr7:1-4,248,090 | 22.5% | 17.5% | 5.0% | 34 | CARD11 | 0.00031 |
| 9p24.2 | 0.11963 | 0.11963 | chr9:2,612,161–5,512,649 | 12.5% | 12.5% | 0.0% | 20 | JAK2 | 3.27 x10 ⁻⁵ |
| 6p25.3 | 0.21318 | 0.21318 | chr6:1–1,258,328 | 5.0% | 5.0% | 0.0% | 5 | IRF4 | 0.0015 |
| 7q34 | 0.21318 | 0.21318 | chr7:127,929,037– 158,821,424 | 17.5% | 10.0% | 7.5% | 222 | BRAF** | - |
| 17q11.2 | 0.21318 | 0.21318 | chr17:21,750,215–37,983,285 | 62.5% | 10.0% | 52.5% | 283 | STAT5B | - |

Bold, Genes with gene localizing data. * Putative oncogene published by the TCGA to be significantly mutated across tumors of multiple lineage types¹. **There are multiple consensus oncogenes in the GISTIC interval, the gene with the highest SCNv mutation rate is listed. The remaining candidate genes reflect the highest-scoring gene in the interval identified by GRAIL. GRAIL P-values are corrected for multiple hypothesis testing. [‡]Only the focal SCNvs are listed for CTCL samples with both focal SCNvs and broad SCNvs involving the indicated GISTIC interval.

Supplementary Table 9. Broad deletions in CTCL

| Chromosome Arm | # CTCL Samples With Deletions | Q-value |
|-----------------------|--------------------------------------|------------------------|
| 1p | 0 | 0.8992776 |
| 1q | 0 | 0.8992776 |
| 2p | 2 | 0.8992776 |
| 2q | 0 | 0.8992776 |
| 3p | 0 | 0.8992776 |
| 3q | 0 | 0.8992776 |
| 4p | 0 | 0.8992776 |
| 4q | 0 | 0.8992776 |
| 5p | 0 | 0.8992776 |
| 5q | 0 | 0.8992776 |
| 6p | 0 | 0.8992776 |
| 6q | 1 | 0.8992776 |
| 7p | 2 | 0.8992776 |
| 7q | 0 | 0.8992776 |
| 8p** | 9 | 0.000680835 |
| 8q | 0 | 0.8992776 |
| 9p | 3 | 0.7896044 |
| 9q | 4 | 0.4562849 |
| 10p** | 8 | 0.002859392 |
| 10q** | 12 | 2.83×10^{-06} |
| 11p | 4 | 0.4562849 |
| 11q | 3 | 0.7896044 |
| 12p | 0 | 0.8992776 |
| 12q | 1 | 0.8992776 |
| 13q | 3 | 0.7896044 |
| 14q | 0 | 0.8992776 |
| 15q | 0 | 0.8992776 |
| 16p | 0 | 0.8992776 |
| 16q | 0 | 0.8992776 |
| 17p** | 34 | 0 |
| 17q | 0 | 0.8992776 |
| 18p | 0 | 0.8992776 |
| 18q | 0 | 0.8992776 |
| 19p | 0 | 0.8992776 |
| 19q | 0 | 0.8992776 |
| 20p | 1 | 0.8992776 |
| 20q | 0 | 0.8992776 |
| 21q | 0 | 0.8992776 |
| 22q | 0 | 0.8992776 |

* False discovery rate between 0.1 and 0.25.

** False discovery rate below 0.1

Supplementary Table 10. Broad amplifications in CTCL

| Chromosome Arm | # CTCL Samples With Amplifications | Q-value |
|---------------------------|---|------------------------|
| 1p | 0 | 0.9096584 |
| 1q | 0 | 0.9096584 |
| 2p | 0 | 0.9096584 |
| 2q | 0 | 0.9096584 |
| 3p | 0 | 0.9096584 |
| 3q | 0 | 0.9096584 |
| 4p | 4 | 0.3537442 |
| 4q** | 6 | 0.07615465 |
| 5p | 2 | 0.9096584 |
| 5q | 1 | 0.9096584 |
| 6p | 0 | 0.9096584 |
| 6q | 0 | 0.9096584 |
| 7p | 2 | 0.9096584 |
| 7q* | 5 | 0.1336047 |
| 8p* | 5 | 0.1336047 |
| 8q** | 17 | 3.88×10^{-11} |
| 9p | 1 | 0.9096584 |
| 9q | 1 | 0.9096584 |
| 10p* | 5 | 0.1336047 |
| 10q | 1 | 0.9096584 |
| 11p | 0 | 0.9096584 |
| 11q | 0 | 0.9096584 |
| 12p | 0 | 0.9096584 |
| 12q | 0 | 0.9096584 |
| 13q | 1 | 0.9096584 |
| 14q | 0 | 0.9096584 |
| 15q | 0 | 0.9096584 |
| 16p | 1 | 0.9096584 |
| 16q | 0 | 0.9096584 |
| 17p | 1 | 0.9096584 |
| 17q** | 23 | 0 |
| 18p* | 5 | 0.1336047 |
| 18q* | 5 | 0.1336047 |
| 19p | 0 | 0.9096584 |
| 19q | 0 | 0.9096584 |
| 20p | 1 | 0.9096584 |
| 20q | 1 | 0.9096584 |
| 21q | 3 | 0.7925349 |
| 22q | 0 | 0.9096584 |

* False discovery rate between 0.1 and 0.25.

** False discovery rate below 0.1

Supplementary Table 11. Clonality of recurrent mutations in driver genes in CTCL

| Gene | SCNV | | SSNV | | Total | |
|---------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | <75% | 76-100% | <75% | 76-100% | <75% | 76-100% |
| <i>TP53</i> | 1 | 35 | 0 | 7 | 1 | 42 |
| <i>ZEB1</i> | 2 | 22 | 0 | 3 | 2 | 25 |
| <i>ARID1A</i> | 4 | 19 | 1 | 1 | 5 | 20 |
| <i>DNMT3A</i> | 3 | 12 | 1 | 3 | 4 | 15 |
| <i>CDKN2A</i> | 0 | 16 | 0 | 1 | 0 | 17 |
| <i>FAS</i> | 2 | 14 | 0 | 3 | 2 | 17 |
| <i>NFKB2</i> | 0 | 4 | 1 | 0 | 1 | 4 |
| <i>CD28</i> | 0 | 0 | 0 | 4 | 0 | 4 |
| <i>PLCG1</i> | 0 | 0 | 1 | 3 | 1 | 3 |
| <i>RHOA</i> | 0 | 0 | 0 | 3 | 0 | 3 |
| <i>STAT5B</i> | 5 | 20 | 0 | 1 | 5 | 21 |
| <i>BRAF</i> | 0 | 0 | 1 | 0 | 1 | 0 |

Each SCNV and SSNV was assessed for the % of tumor cells with the mutation. Samples were binned with <75% clonality and 75-100% clonality. Listed are numbers of samples with each SCNV and SSNV in each bin.

Supplementary Table 12. SCNVs and SSNVs of candidate driver genes residing on significant focal deletions.

| Cytoband | GISTIC Residual Q Value | # Genes in Peak | Candidate Gene | GRAIL P value | % of CTCLs With Deletions | | | | # of SSNVs |
|----------|-------------------------|-----------------|--------------------------|---------------|---------------------------|-----------|-------|-------|----------------|
| | | | | | Total | Biallelic | Focal | Broad | |
| 1p36.11 | 6.59x10 ⁻¹⁶ | 7 | ARID1A* | - | 57.5% | 0.0% | 57.5% | 0.0% | NS (1), FS(1) |
| 9p21.3 | 6.59x10 ⁻¹⁶ | 1 | CDKN2A* | - | 40.0% | 30.0% | 7.5% | 2.5% | NS (1) |
| 10p11.22 | 2.01x10 ⁻¹³ | 1 | ZEB1 | - | 60.0% | 12.5% | 20.0% | 27.5% | NS (2), MS (1) |
| 10q24.32 | 4.77x10 ⁻⁷ | 7 | NFKB2^t | - | 10.0% | 5.0% | 5.0% | 0.0% | SS (1) |
| 2p23.3 | 5.57x10 ⁻⁷ | 4 | DNMT3A* | - | 37.5% | 2.5% | 27.5% | 7.5% | NS (1), MS (3) |
| 11q22.3 | 1.76x10 ⁻⁵ | 5 | ATM* | - | 30.0% | 2.5% | 20.0% | 7.5% | MS (1) |
| 2q37.3 | 0.00010354 | 38 | PDCD1 | 0.0035 | 20.0% | 5.0% | 15.0% | 0.0% | - |
| 19p13.3 | 0.00039797 | 1078 | STK11** | - | 30.0% | 0.0% | 30.0% | 0.0% | - |
| 9q21.32 | 0.00048049 | 33 | DAPK1 | 0.00062 | 25.0% | 0.0% | 12.5% | 12.5% | - |
| 13q14.2 | 0.00056364 | 40 | RB1* | - | 25.0% | 0.0% | 15.0% | 10.0% | MS (1) |
| 19p13.3 | 0.000581 | 53 | GADD45B | 0.025 | 27.5% | 0.0% | 27.5% | 0.0% | - |
| 12p13.2 | 0.00085553 | 64 | CDKN1B* | - | 15.0% | 0.0% | 15.0% | 0.0% | - |
| 6q23.3 | 0.0016017 | 2 | TNFAIP3*** | - | 25.0% | 2.5% | 22.5% | 0.0% | - |
| 10q21.2 | 0.00236 | 2 | ZNF365 | 0.58 | 50.0% | 0.0% | 15.0% | 35.0% | - |
| 10q26.3 | 0.004758 | 46 | MGMT | 0.00016 | 47.5% | 0.0% | 30.0% | 17.5% | - |
| 7p21.1 | 0.011766 | 107 | NFE2L3* | - | 15.0% | 0.0% | 10.0% | 5.0% | - |
| 9q31.1 | 0.022893 | 58 | XPA | 0.39 | 20.0% | 0.0% | 15.0% | 5.0% | - |
| 6q25.2 | 0.04894 | 73 | CCR6 | 0.12 | 15.0% | 0.0% | 15.0% | 0.0% | - |
| 12q21.33 | 0.04894 | 160 | SOCS2 | 0.076 | 15.0% | 0.0% | 12.5% | 2.5% | - |
| 6q21 | 0.051525 | 42 | TRAF3IP2 | 0.02 | 12.5% | 0.0% | 10.0% | 2.5% | - |
| 16q22.1 | 0.055785 | 2 | CTCF* | - | 12.5% | 0.0% | 12.5% | 0.0% | MS (1) |
| 10q26.11 | 0.095529 | 19 | BAG3 | 0.46 | 52.5% | 0.0% | 32.5% | 20.0% | - |
| 10q23.31 | 0.11663 | 2 | FAS | - | 40.0% | 2.5% | 10.0% | 27.5% | MS (3) |
| 11p13 | 0.11663 | 28 | WT1* | - | 17.5% | 0.0% | 7.5% | 10.0% | - |
| 6q25.1 | 0.1338 | 3 | LATS1 | 0.17 | 22.5% | 0.0% | 22.5% | 0.0% | - |

| | | | | | | | | | |
|----------|---------|-----|---------|-------|-------|------|-------|-------|--------|
| 16q24.2 | 0.1338 | 21 | BANP | 0.021 | 2.5% | 0.0% | 2.5% | 0.0% | - |
| 13q12.13 | 0.1338 | 101 | BRCA2** | - | 15.0% | 0.0% | 5.0% | 10.0% | MS (1) |
| 1p22.1 | 0.15669 | 106 | RPL5* | - | 10.0% | 0.0% | 10.0% | 0.0% | - |
| 8p23.1 | 0.23282 | 644 | EGR3** | - | 30.0% | 0.0% | 7.5% | 22.5% | - |

Bold, Genes with statistically significant gene localizing data. * Putative tumor suppressors published by the TCGA to be significantly mutated across tumors of multiple lineage types¹. **There are multiple consensus tumor suppressors in the GISTIC interval, the gene with the highest SCNv mutation rate is listed. ****TNFAIP3* was listed because it was recently implicated in leukemic CTCL². The remaining candidate genes were identified by GRail. GRail P-values are corrected for multiple hypothesis testing. [†] For NFKB2, only the truncating mutations are shown. NS, represents nonsense mutations; FS, frameshift mutations; MS, missense mutations; SS, splice site mutations.

Supplementary Table 13. SCNVs and SSNVs for candidate driver genes on significant focal amplifications

| Cytoband | GISTIC Residual Q-value | # Genes in Peak | Candidate Gene | GRAIL P-value | % of CTCLs With Amplifications | | | # of SSNVs |
|----------|-------------------------|-----------------|----------------|-----------------------|--------------------------------|-------|-------|---------------------|
| | | | | | Total | Focal | Broad | |
| 10p15.1 | 0.00029474 | 9 | PRKCQ | 6.78×10^{-5} | 30.0% | 17.5% | 12.5% | - |
| 10p12.33 | 0.063691 | 21 | TRDMT1 | 0.18 | 22.5% | 10.0% | 12.5% | - |
| 7p22.2 | 0.098621 | 34 | CARD11 | 0.00031 | 22.5% | 17.5% | 5.0% | MS (3) [#] |
| 9p24.2 | 0.11963 | 20 | JAK2 | 3.27×10^{-5} | 12.5% | 12.5% | 0.0% | - |
| 6p25.3 | 0.21318 | 5 | IRF4 | 0.0015 | 5.0% | 5.0% | 0.0% | MS (2) |
| 7q34 | 0.21318 | 222 | BRAF ** | - | 15.0% | 2.5% | 12.5% | MS (1) [#] |
| 17q11.2 | 0.21318 | 283 | STAT5B | - | 62.5% | 5.0% | 57.5% | MS (1) [#] |

Bold, Genes with gene localizing data.* Putative oncogene published by the TCGA to be significantly mutated across tumors of multiple lineage types¹. **There are multiple consensus oncogenes in the GISTIC interval, the gene with the highest SCNV mutation rate is listed. The remaining candidate genes were identified by GRAIL. GRAIL P-values are corrected for multiple hypothesis testing. MS, missense mutations. [#]Missense mutations were previously reported in other cancers.

Supplementary Table 14a. SCNVs and SSNVs of putative driver genes residing on narrow, significant focal deletions intervals (<10 Genes in GISTIC interval).

| Candidate Gene | # Genes Residing on GISTIC Peak | GISTIC Residual Q-value | # SSNVs | % of CTCLs with Deletions | | % of CTCLs with SCNV or SSNV | GRAIL P-value | |
|----------------|---------------------------------|-------------------------|---------|---------------------------|-------------|------------------------------|---------------|---|
| | | | | Biallelic | Monoallelic | | | |
| | | | | Focal | Broad | | | |
| ATM* | 5 | 1.8 x10 ⁻⁵ | 1 | 2.5 | 20 | 7.5 | 30 | - |
| TNFAIP3** | 2 | 0.0016 | 0 | 2.5 | 22.5 | 0 | 25 | - |
| CTCF* | 2 | 0.056 | 1 | 0 | 12.5 | 0 | 15 | - |

Table 14b. SCNVs and SSNVs of putative driver genes residing on narrow, significant focal amplification intervals (<10 Genes in GISTIC interval).

| Candidate Gene | # Genes Residing on GISTIC Peak | GISTIC Residual Q-value | # SSNVs | % of CTCLs with Amplifications | | % of CTCLs with SCNV or SSNV | GRAIL P-value |
|----------------|---------------------------------|-------------------------|---------|--------------------------------|-------|------------------------------|------------------------|
| | | | | Focal | Broad | | |
| PRKCQ | 9 | 0.00029 | 0 | 17.5 | 12.5 | 30 | 6.8 x 10 ⁻⁵ |
| IRF4 | 5 | 0.21318 | 2 | 5 | 0 | 10 | 0.0015 |

*Genes with statistically significant mutations across multiple lineage types as published by the TCGA. ***TNFAIP3* was listed because it was recently implicated in leukemic CTCL². The remaining candidate genes were identified by GRAIL.

Supplementary Table 15a. Transcript levels of driver genes for which there are at least 3 CTCLs with the normal diploid gene copy number and at least 3 CTCLs with deletions.

| Gene | Number of Diploid CTCLs | Mean FPKM of Normal | Standard Error | Number of CTCLs with SCNVs | Mean FPKM of Mutant | Standard Error | P-value |
|---------|-------------------------|---------------------|----------------|----------------------------|---------------------|----------------|---------|
| ARID1A | 5* | 32.37 | 4.103 | 5 | 20.5 | 2.733 | 0.016 |
| CTCF | 8 | 30.55 | 1.393 | 3 | 23.43 | 2.911 | 0.042 |
| TNFAIP3 | 8 | 268.6 | 73.39 | 3 | 59.9 | 23.93 | 0.012 |
| ZEB1 | 6* | 12.76 | 1.476 | 3** | 5.095 | 1.27 | 0.024 |

Supplementary Table 15b. Transcript levels of driver genes for which there are at least 3 CTCLs with the normal diploid gene copy number and at least 3 CTCLs with amplifications.

| Gene | Number of Diploid CTCLs | Mean FPKM of Normal | Standard Error | Number of CTCLs with SCNVs | Mean FPKM of Mutant | Standard Error | P-value |
|--------|-------------------------|---------------------|----------------|----------------------------|---------------------|----------------|---------|
| STAT5B | 3 | 91.37 | 9.05 | 8 | 144 | 8.685 | 0.012 |

Diploid CTCLs, CTCLs with two copies of indicated gene. FPKM, fragments per kilobase per million reads. P-values were calculated by one-sided Mann-Whitney t-test. * 1 sample was excluded from diploid CTCLs because it harbored a nonsense mutation. ** 1 sample was excluded from CTCLs with SCNVs because it harbored a biallelic deletion.

† 1 sample was excluded from diploid CTCLs because it harbored an intragenic deletion.

Supplementary Table 16. Identification by whole genome sequencing of breakpoints contributing to SCNVs in CTCL.

| Sample | Breakpoint #1 | | | Breakpoint #2 | | | Structural Variant | Driver genes affected | Total Reads Supporting | Micro-homology | Non-templated Sequence |
|--------|---------------|-------------|---------------------------|---------------|-------------|-------------------------|--------------------|-----------------------|------------------------|----------------|------------------------|
| | Chr | Position | Location | Chr | Position | Location | | | | | |
| 13 | 1 | 26,713,590 | Intergenic | 1 | 27,941,189 | Intron of <i>FAM76A</i> | Deletion | <i>ARID1A</i> | 18 | 1 | |
| 13 | 1 | 148,622,707 | Intron of <i>RPRD2</i> | 1 | 149,660,970 | Intron of <i>POGZ</i> | Deletion | | 16 | 1 | |
| 13 | 1 | 224,379,591 | Intergenic | 1 | 233,378,232 | Intron of <i>RBM34</i> | Deletion | | 23 | 1 | |
| 13 | 1 | 228,512,286 | Intergenic | 1 | 233,382,968 | Intron of <i>RBM34</i> | Inversion | | 16 | 1 | |
| 13 | 2 | 11,943,988 | Intergenic | 2 | 86,138,204 | Intron of <i>POLR1A</i> | Inversion | | 33 | 0 | |
| 13 | 2 | 12,448,066 | Intron of <i>AK001558</i> | 2 | 85,728,835 | Intron of <i>USP39</i> | Inversion | | 28 | 3 | |
| 13 | 2 | 25,342,324 | Intron of <i>DNMT3A</i> | 11 | 101,668,436 | Intergenic | Translocation | <i>DNMT3A, ATM</i> | 22 | 1 | |
| 13 | 2 | 26,099,685 | Intergenic | 11 | 117,868,561 | Intron of <i>MLL</i> | Translocation | <i>DNMT3A, ATM</i> | 16 | 1 | |
| 13 | 3 | 171,780,269 | Intron of <i>SLC7A14</i> | 10 | 73,141,626 | Intron of <i>CDH23</i> | Translocation | | 18 | 1 | |
| 13 | 3 | 176,284,784 | Intron of <i>NAALADL2</i> | 10 | 66,110,178 | Intergenic | Translocation | <i>ZEB1</i> | 15 | 9 | |
| 13 | 5 | 139,813,420 | Intron of <i>ANKHD1</i> | 11 | 117,857,758 | Exon of <i>MLL</i> | Translocation | | 13 | 2 | |
| 13 | 5 | 139,985,062 | Intergenic | 11 | 117,860,771 | Intron of <i>MLL</i> | Translocation | | 28 | 0 | |
| 13 | 9 | 21,818,243 | Intron of <i>MTAP</i> | 9 | 31,671,084 | Intergenic | Inversion | <i>CDKN2A</i> | 30 | | 4 |
| 13 | 9 | 31,207,570 | Intergenic | 9 | 21,818,204 | Intron of <i>MTAP</i> | Deletion | | 29 | | 1 |
| 13 | 10 | 25,778,773 | Intron of <i>GPR158</i> | 3 | 175,371,776 | Intron of <i>NLGN1</i> | Translocation | <i>ZEB1</i> | 16 | | 0 |
| 13 | 10 | 72,379,661 | Intergenic | 10 | 107,233,329 | Intergenic | Deletion | | 8 | | 2 |
| 13 | 10 | 73,204,904 | Intron of <i>CDH23</i> | 3 | 175,193,528 | Intron of <i>NLGN1</i> | Translocation | | 23 | 0 | |
| 13 | 10 | 73,309,035 | Intergenic | 10 | 57,650,001 | Intergenic | Inversion | | 21 | 0 | |
| 13 | 10 | 89,295,708 | Intron of <i>MINPP1</i> | 10 | 90,907,696 | Intergenic | Deletion | <i>FAS</i> | 25 | | 1 |
| 13 | 10 | 98,565,242 | Intergenic | 10 | 25,538,394 | Intron of <i>GPR158</i> | Deletion | | 17 | | 1 |
| 13 | 10 | 108,330,062 | Intron of <i>SORCS1</i> | 10 | 25,550,092 | Exon of <i>GPR158</i> | Inversion | | 19 | | 2 |
| 13 | 10 | 110,861,244 | Intergenic | 10 | 57,822,392 | Intergenic | Inversion | | 16 | 3 | |
| 13 | 10 | 119,033,460 | Exon of <i>PDZD8</i> | 10 | 42,411,948 | Intron of <i>ZNF338</i> | Deletion | | 33 | 1 | |
| 13 | 10 | 119,267,885 | Intron of <i>EMX2OS</i> | 10 | 56,596,808 | Intron of <i>PCDH15</i> | Inversion | | 27 | 2 | |
| 17 | 2 | 24,367,403 | intron of <i>ITSN2</i> | 2 | 25,537,807 | Intron of <i>DTNB</i> | Deletion | <i>DNMT3A</i> | 19 | 5 | |

| | | | | | | | | | | |
|----|----|-------------|---------------------------|----|-------------|---------------------------|---------------|------------------|----|----|
| 17 | 2 | 197,673,437 | Intron of <i>ANKRD44</i> | 2 | 199,342,112 | Intergenic | Inversion | | 19 | 0 |
| 17 | 2 | 198,522,274 | intron of <i>PLCL1</i> | 2 | 199,414,606 | Intergenic | Deletion | | 25 | 2 |
| 17 | 10 | 24,938,888 | Intergenic | 10 | 129,541,310 | Intron of <i>ARHGAP21</i> | Deletion | <i>ZEB1, FAS</i> | 11 | 2 |
| 17 | 10 | 30,991,190 | Intergenic | 15 | 72,750,546 | Intron of <i>EDC3</i> | Translocation | <i>ZEB1</i> | 16 | 1 |
| 17 | 10 | 32,144,341 | Intron of <i>ARHGAP12</i> | 15 | 73,749,197 | Intergenic | Translocation | <i>ZEB1</i> | 10 | 1 |
| 17 | 10 | 35,838,582 | Intron of <i>CCNY</i> | 11 | 30,352,778 | Intergenic | Translocation | | 20 | 15 |
| 17 | 10 | 57,379,788 | Intergenic | 10 | 59,091,465 | Intergenic | Deletion | | 18 | 9 |
| 17 | 10 | 57,380,995 | Intergenic | 10 | 129,415,540 | Intergenic | Deletion | | 29 | 0 |
| 17 | 10 | 60,008,131 | Intron of <i>BICC1</i> | 10 | 128,178,240 | intron of <i>C10orf90</i> | Inversion | | 23 | 1 |
| 17 | 10 | 65,707,339 | Intergenic | 10 | 60,337,921 | Intergenic | Translocation | | 27 | 0 |
| 17 | 10 | 103,775,255 | Intron of <i>C10orf76</i> | 10 | 104,566,246 | Intron of <i>C10orf26</i> | Deletion | <i>NFKB2</i> | 23 | 1 |
| 17 | 10 | 104,150,525 | Exon of <i>NFKB2</i> | 5 | 92,742,617 | Intergenic | Translocation | <i>NFKB2</i> | 21 | 3 |
| 17 | 10 | 104,165,993 | Intron of <i>PSD</i> | 5 | 87,323,821 | Intergenic | Translocation | <i>NFKB2</i> | 11 | 0 |
| 17 | 10 | 121,408,863 | Intron of <i>BAG3</i> | 19 | 43,536,117 | Intron of <i>C19orf15</i> | Translocation | | 8 | 1 |
| 17 | 10 | 121,501,394 | Intron of <i>INPP5F</i> | 19 | 44,184,624 | Intergenic | Translocation | | 25 | 2 |
| 17 | 10 | 128,122,034 | Intron of <i>C10orf90</i> | 10 | 57,882,148 | Intergenic | Inversion | | 18 | 4 |
| 17 | 10 | 128,777,367 | Intron of <i>DOCK1</i> | 12 | 18,904,702 | Intergenic | Translocation | | 20 | 1 |
| 17 | 10 | 130,061,878 | Intergenic | 10 | 65,535,231 | Intergenic | Inversion | | 21 | 1 |
| 17 | 11 | 26,374,469 | Intron of <i>ANO3</i> | 11 | 30,033,694 | Intergenic | Deletion | | 11 | 1 |
| 17 | 11 | 63,402,680 | Intron of <i>MARK2</i> | 10 | 34,676,505 | Intron of <i>PARD3</i> | Translocation | | 15 | 4 |
| 17 | 12 | 17,974,225 | Intergenic | 12 | 25,562,201 | Intron of <i>LMNTD1</i> | Inversion | | 31 | 3 |
| 17 | 12 | 60,257,721 | Intergenic | 12 | 121,277,149 | Exon of <i>DIABLO</i> | Inversion | | 14 | 2 |
| 17 | 12 | 121,329,259 | Intron of <i>CLIP1</i> | 12 | 57,264,550 | Intergenic | Translocation | | 17 | 0 |
| 17 | 16 | 66,130,715 | Intron of <i>FAM65A</i> | 19 | 44,085,041 | Intron of <i>NFKB1B</i> | Translocation | <i>CTCF</i> | 24 | 1 |
| 17 | 16 | 66,227,930 | Intron of <i>CTCF</i> | 1 | 23,836,991 | Intron of <i>MDS2</i> | Translocation | <i>CTCF</i> | 18 | 1 |
| 17 | 17 | 7,211,140 | Intergenic | 17 | 7,849,358 | Intron of <i>GUCY2D</i> | Deletion | <i>TP53</i> | 17 | 8 |

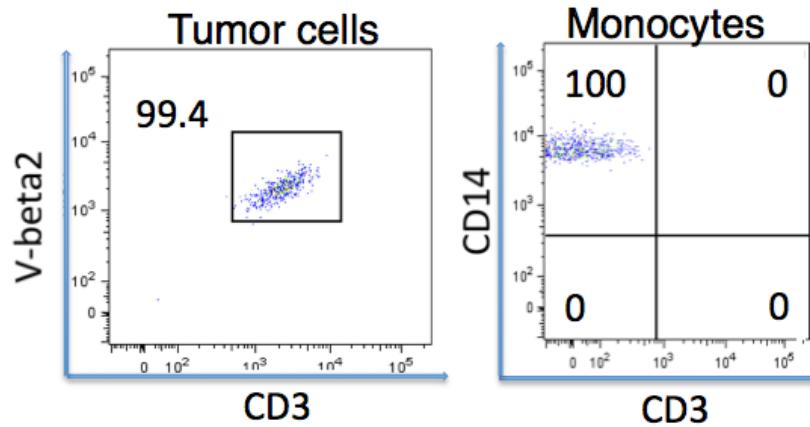
Supplementary Table 17. IL-2 primer sequences

| Primer name | Target gene | Primer sequence |
|-------------|-------------|------------------------------|
| IL-2-FWD | IL-2 | 5' - CAAACCTCTGGAGGAAGTGC-3' |
| IL-2-REV | IL-2 | 5' - GGTTGCTGTCTCATCAGCAT-3' |

References for supplementary tables:

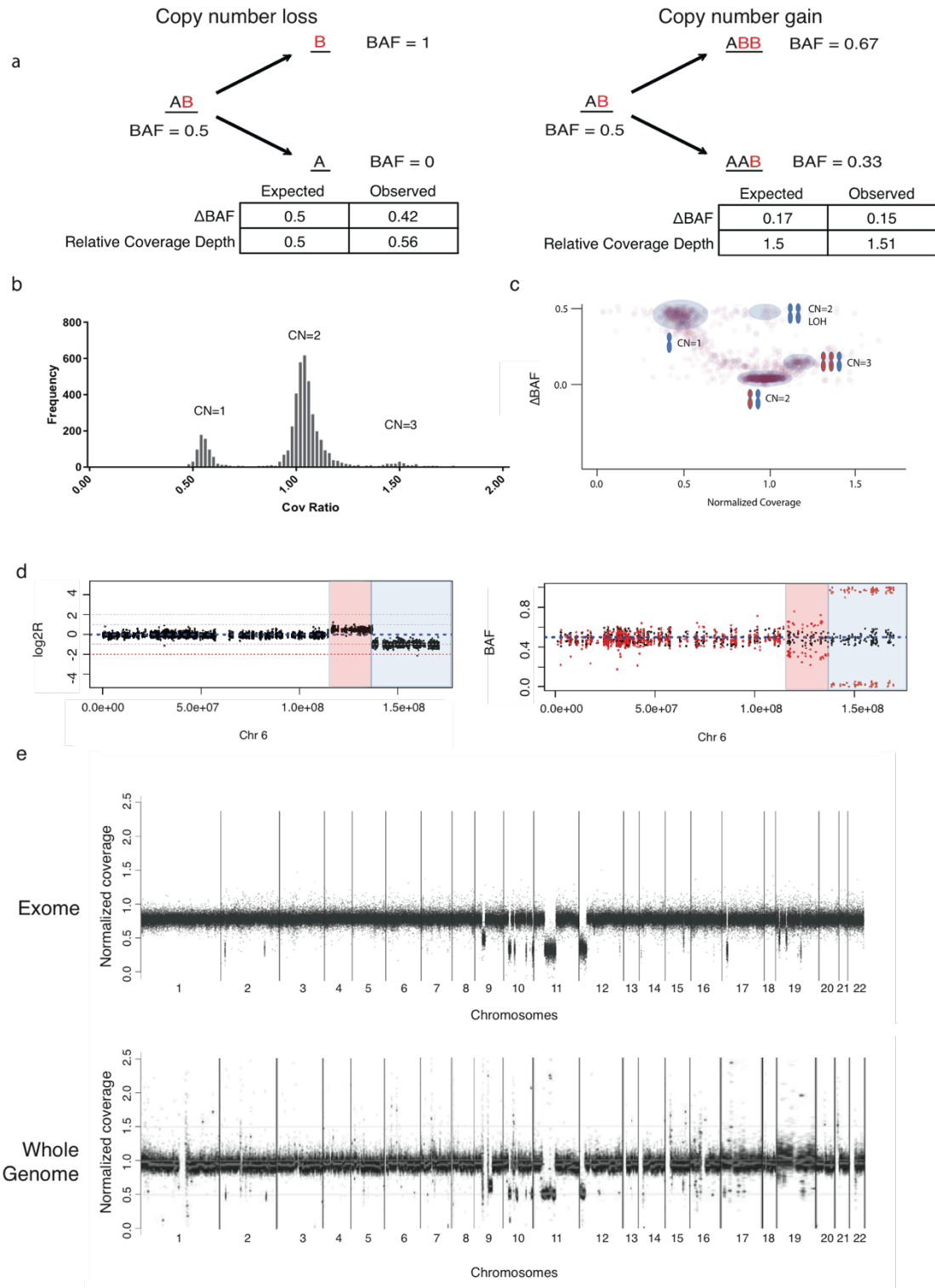
1. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333-9 (2013).
2. Braun, F.C. *et al.* Tumor suppressor TNFAIP3 (A20) is frequently deleted in Sezary syndrome. *Leukemia* **25**, 1494-501 (2011).

Supplementary Figure 1. Cell surface staining of tumor cells and normal cells after flow cytometry activated cell sorting.



Flow cytometry activated cell sorting was performed to isolate CTCL cells and matched monocytes. The post-sort flow cytometry to assess uniformity of cell sorting is shown above. CTCL clones were identified by a T cell marker (CD3) and an antibody targeting the TCR V- β utilized by the leukemic CTCL clone (in this case, TCR-V β 2). Monocytes were identified by the presence of CD14, a monocyte-specific marker, and the absence of CD3 staining.

Supplementary Figure 2. Somatic copy number mutations identified by relative coverage depth and changes in B-allele frequency (BAF) at informative SNPs

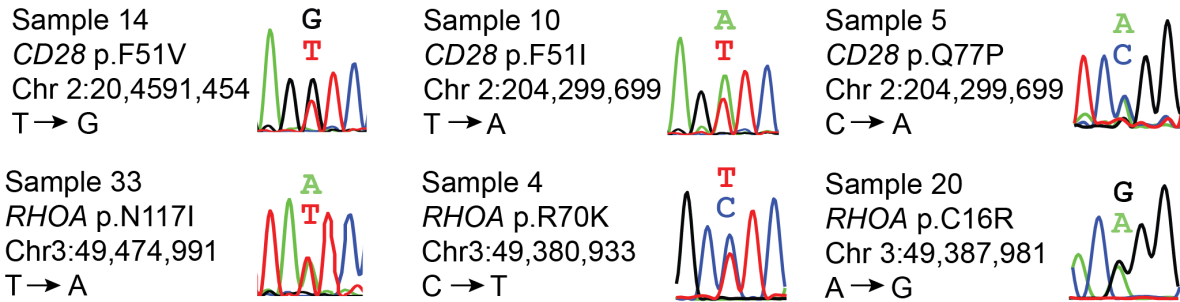


(a) Expected and observed changes in normalized coverage depth and B-allele frequency in copy number deletions and copy number gains identified by exome sequencing. **(b)** Histogram of normalized coverage depth of 50 kb segments in a sample of CTCL. The genome was binned into contiguous 50 kb segments. The normalized coverage depth from exome sequencing was calculated from the tumor: normal read depth of the segment compared to the average tumor:normal read depth across the genome. **(c)** Plots of B-allele frequency (BAF) vs. normalized read coverage generated by exome sequencing of all 40 samples. Copy number loss is marked by loss-of-heterozygosity ($\Delta\text{BAF}=0.5$) and halving of normalized coverage. Copy neutral loss-of-heterozygosity is marked by loss-of-heterozygosity and no change in coverage depth. Copy number amplification is marked by characteristic changes in B-allele-frequency ($\Delta\text{BAF}=0.16$) and increases in coverage depth. Each point represents one segment from one sample of a defined copy number. **(d)** A representative example of a plot of relative coverage depth and corresponding plot of BAF. Each point represents the \log_2 relative ratio (\log_2R) of coverage between tumor and normal sample at a targeted region from the Nimblegen array. The blue dotted line indicates an LRR of 0. The red dotted line indicates an LRR of -2 . For the BAF plot, heterozygous SNVs in the normal (B-allele frequency of 0.4–0.6) are plotted in black. The BAFs of the tumors at the corresponding SNVs are plotted in red. An area designated as a copy number gain from 2→3 is shaded red. An area designated as a copy number loss from 2→1 is shaded blue. **(e)** Plots of normalized coverage generated by exome sequencing and whole genome sequencing. WGS plots were generated with Patchwork-R.

Supplementary Figure 3. Sanger sequencing of select point mutations.

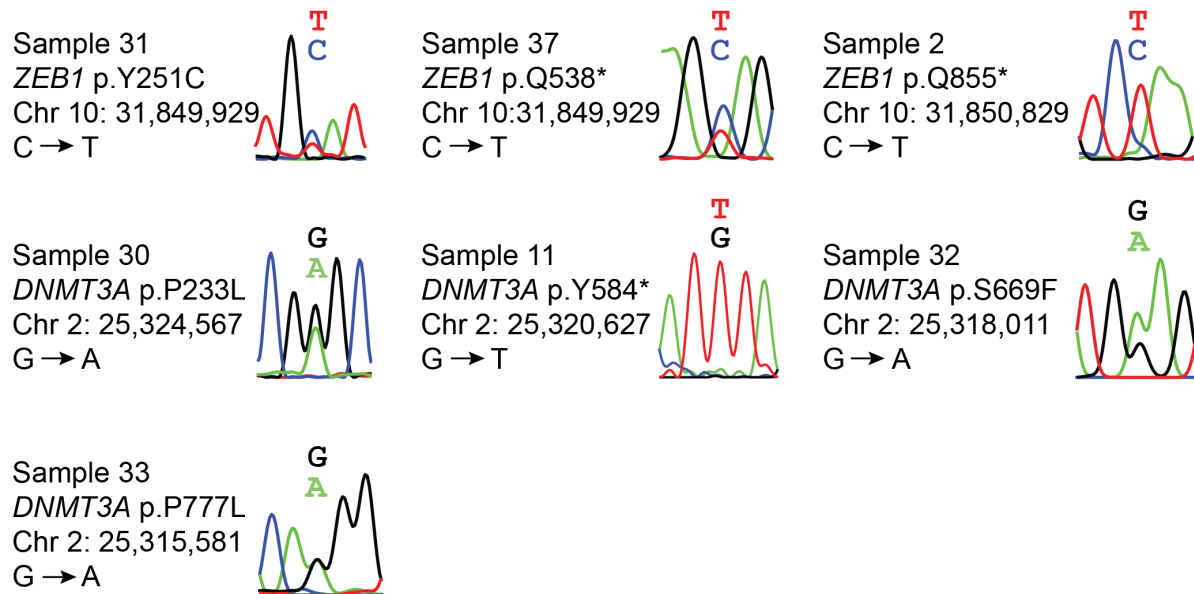
(a) Representative chromatograms of Sanger DNA sequence of CTCL DNA for point mutations in *CD28* and *RHOA*. (b) Representative chromatograms of Sanger DNA sequence of CTCL DNA for point mutations in *DNMT3A* and *ZEB1*.

a



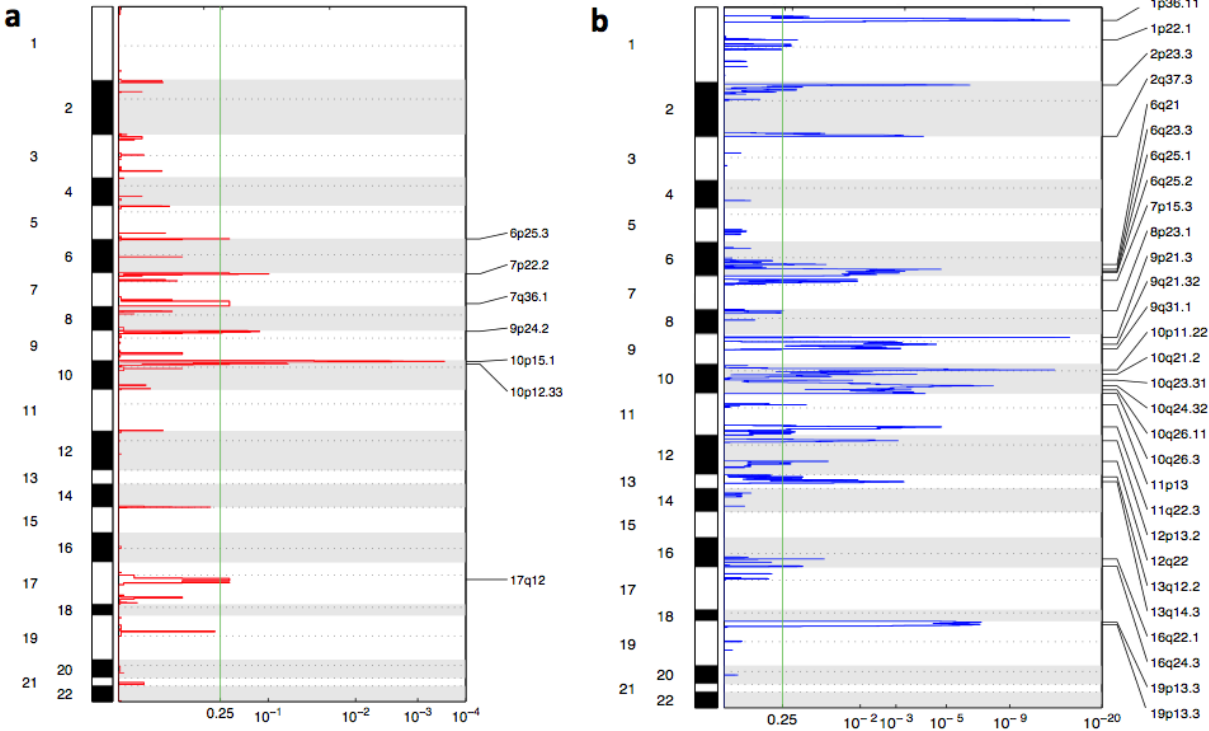
| Gene | <i>CD28</i> | | | <i>RHOA</i> | | |
|--|-------------|------|------|-------------|------|-------|
| | F51V | F51I | Q77P | C16R | R70K | N117I |
| # samples with SSNV identified by exome sequencing | 2 | 1 | 2 | 1 | 1 | 2 |
| # samples with SSNV confirmed by Sanger sequencing | 2 | 1 | 2 | 1 | 1 | 2 |

b



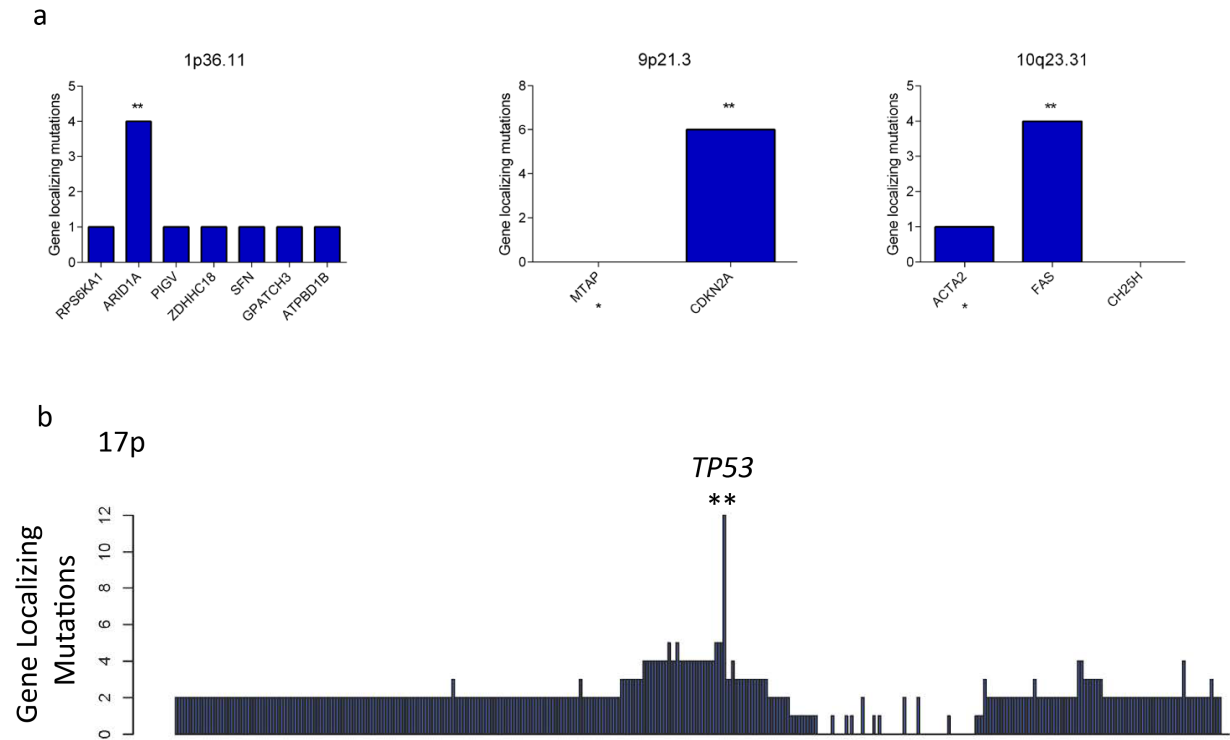
| Gene | <i>DNMT3A</i> | <i>ZEB1</i> |
|--|---------------|-------------|
| # samples with SSNV identified by exome sequencing | 4 | 3 |
| # samples with SSNV confirmed by Sanger sequencing | 4 | 3 |

Supplementary Figure 4. GISTIC 2.0 significance plots identifying recurrent focal SCNVs with a false discovery rate < 0.25



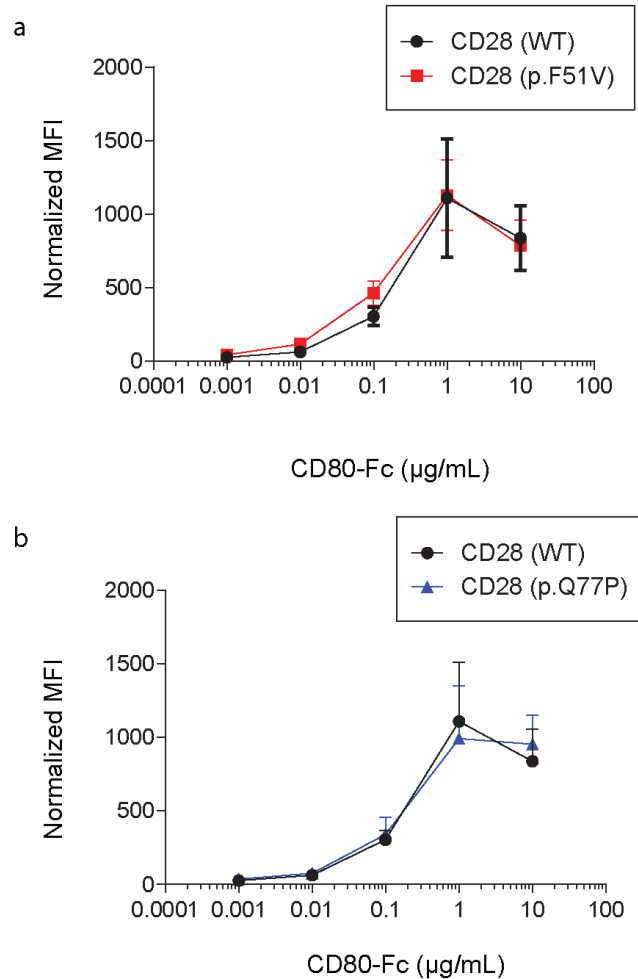
(a) GISTIC 2.0 plot identifying statistically significant recurrent amplifications. **(b)** GISTIC 2.0 plot identifying statistically significant recurrent deletions.

Supplementary Figure 5. Gene localizing mutations in significant SCNVs in CTCL.



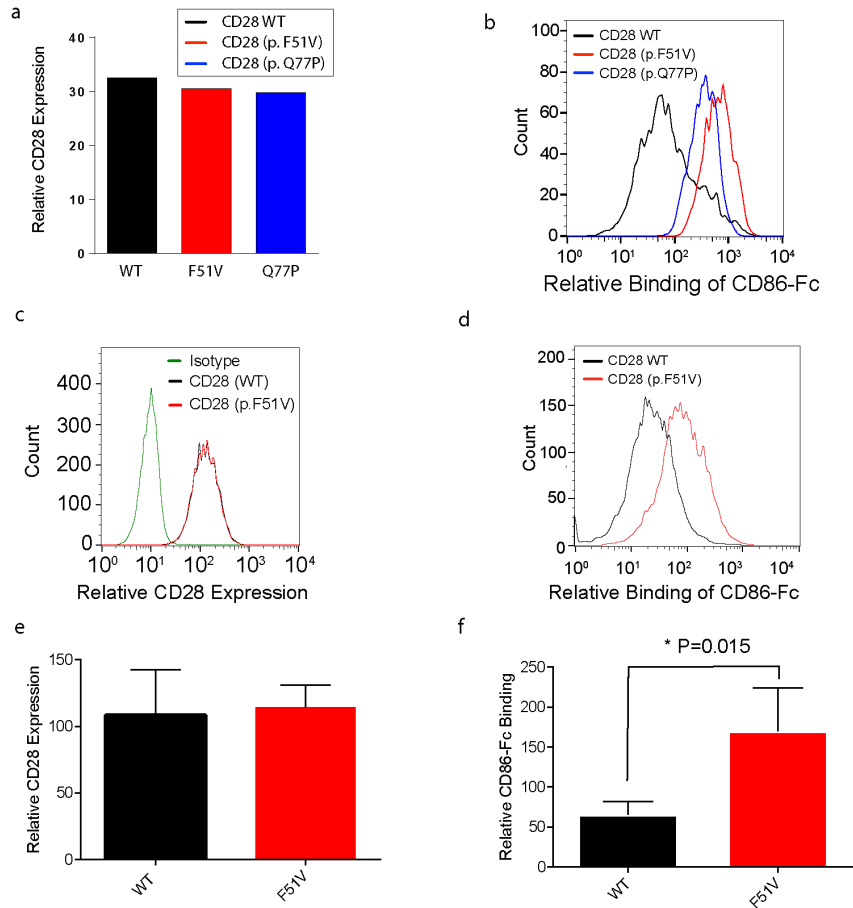
Gene localizing mutations for genes in **(a)** significant focal intervals or **(b)** significant arm-level events. Gene localizing mutations are defined as mutations that would affect at least one but not all genes in the interval. They include non-silent SSNVs and focal SCNVs that overlap with at least one but fewer than all genes in the interval. * indicates gene immediately outside of GISTIC interval. ** indicates likelihood of gene being the target gene in that interval >99.9%.

Supplementary Figure 6. Relative binding of wild-type CD28 or its mutants to CD80-Fc.



Aliquots of 293T cells expressing wild-type CD28 or its mutants (p.F51V (a) or p.Q77P (b)) were independently incubated with anti-CD28 antibody (eBioscience) or the indicated amounts of CD80-Fc (R&D Systems). CD80-binding was subsequently detected by Alexa Fluor 647-labeled secondary antibody (BD Pharmingen). Mean fluorescence intensity was determined from data histograms using FloJo. CD80-Fc binding was then normalized according to the relative expression of the CD28 isoform on the 293T cells. The graph plots normalized mean fluorescence intensity (MFI) as a function of CD80-Fc concentration. The results show the mean of three independent experiments with standard error.

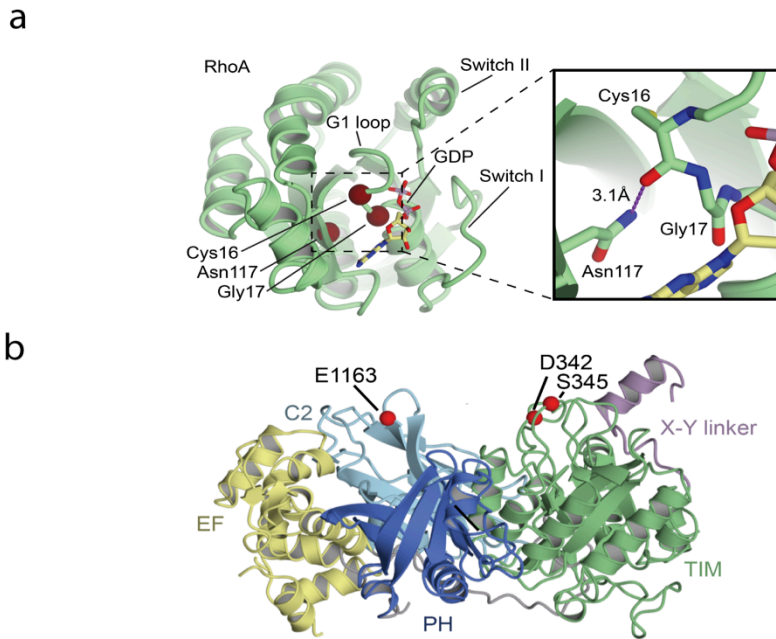
Supplementary Figure 7. Relative CD86-Fc binding in cells expressing equivalent levels of CD28 isoforms.



(a) 293T cells expressing wild-type (WT) CD28 or its mutants were simultaneously stained with both an anti-CD28 antibody (eBioscience) and CD86-Fc (R&D Systems). 293T cells were gated for equivalent levels of CD28 and the mean fluorescence intensity for each gated population is shown here. **(b)** Relative CD28 expression of gated 293T cells stably transduced with human WT CD28 or CD28 mutants (p.F51V or p.Q77P) (CD28 expression for this experiment is shown in **(a)**). CD28-expressing 293T cells were incubated with CD86-Fc (10 μ g/ml), which was subsequently detected by Alexa Fluor 647-labeled secondary antibody (BD Pharmingen). **(c)** Relative expression of CD28 isoforms on Jurkat cells. In a representative experiment, Jurkat cells were stably transduced with WT CD28 or CD28 (p.Phe51Val) and stained with FITC-conjugated anti-CD28 antibody (eBioscience) or isotype control (Green). **(d)** FACS plots

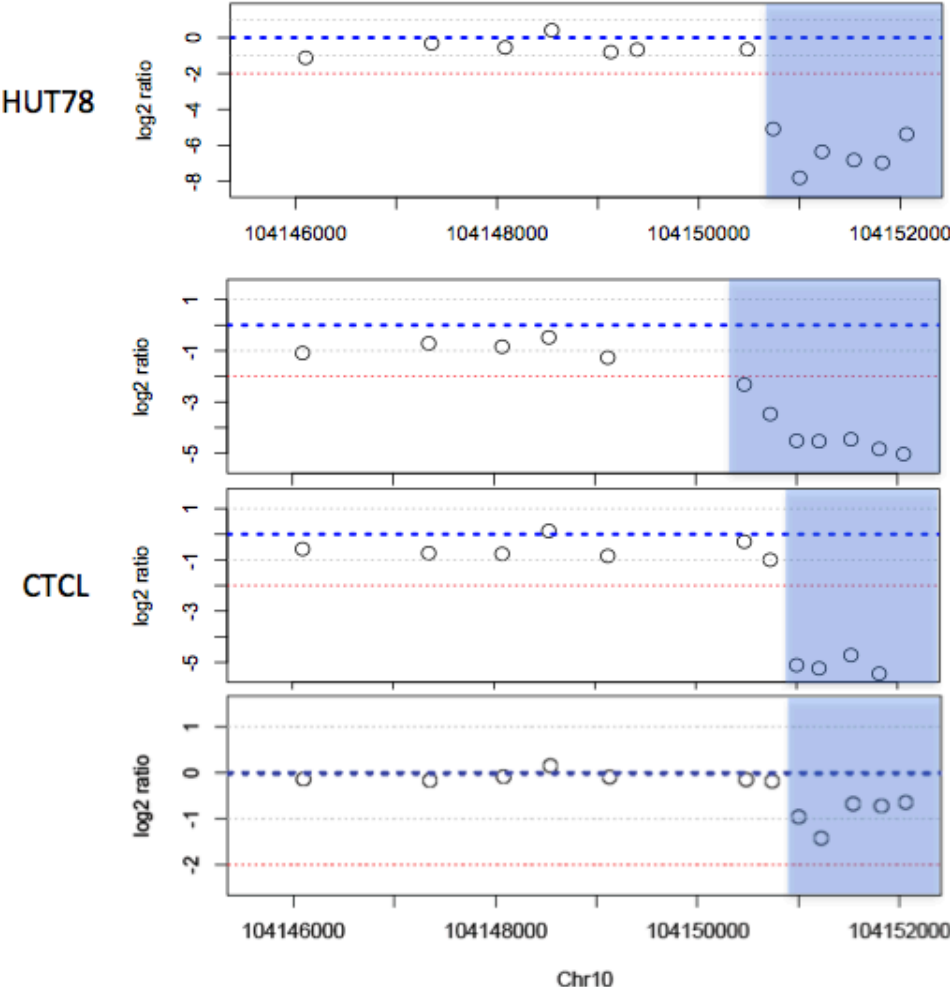
demonstrating relative binding of CD86-Fc (10 $\mu\text{g/ml}$) to Jurkat cells expressing equivalent levels of WT CD28 or CD28 (p.Phe51Val). Relative CD28 expression for this experiment is shown in **(d)**. Binding of CD86-Fc was detected by Alexa Fluor 647-labeled secondary antibody. **(e)** Relative CD28 expression on Jurkat cells expressing human wild-type CD28 or its mutant (p.Phe51Val). The results show the mean of four biological replicates with standard error. **(f)** Relative binding of CD86-Fc (10 $\mu\text{g/ml}$) to Jurkat cells stably transduced with human wild-type CD28 or its mutant (p.Phe51Val). The relative CD28 expression for these experiments are shown in **(e)**. The results show the mean of four biological replicates with standard error. * indicates $P < 0.05$.

Supplementary Figure 8. Structural features of CTCL mutations in RHOA and PLCG1.



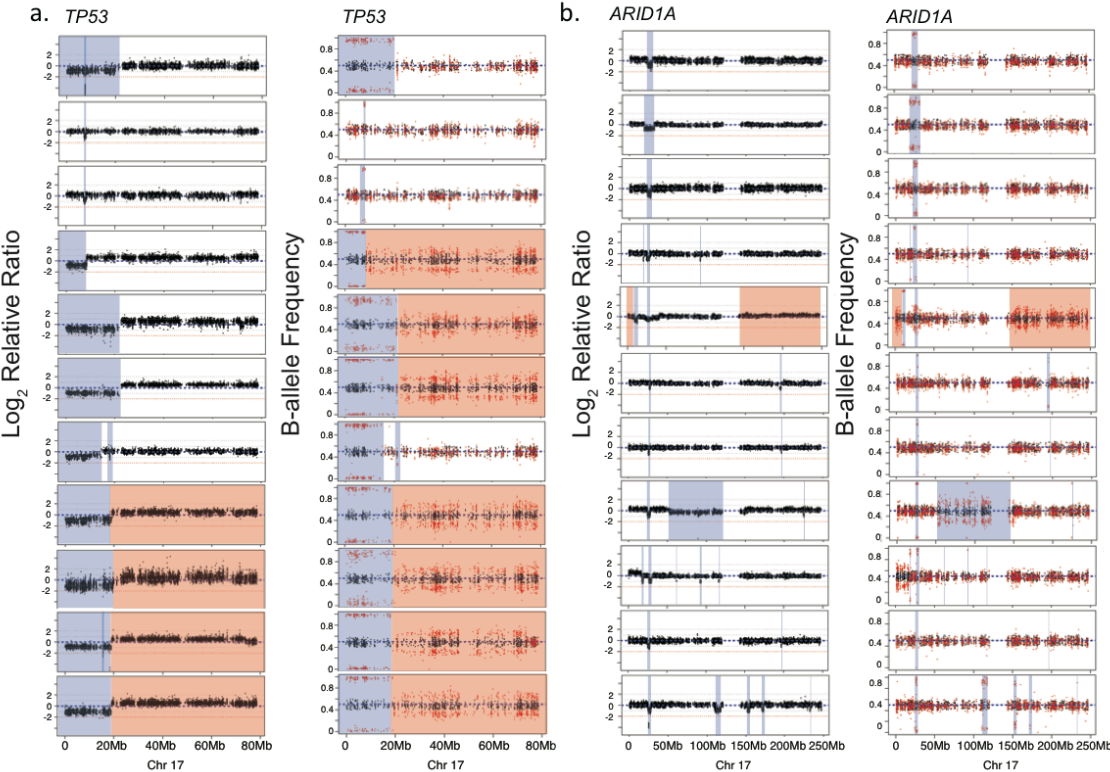
(a) Clustering of residues p.Cys16, p.Gly17 and p.Asn117 in the GTP binding pocket in RhoA. p.Gly17 and p.Asn117 are sites of recurrent mutations in T cell lymphoma that abolish RhoA's ability to bind GTP^{1,2} (**Fig. 2a**). Crystal structure of RhoA in complex with GDP is shown (PDB ID: 1FTN)³. The tight clustering in the GTP binding pocket of residues p.Cys16, p.Gly17, and p.Asn117 are indicated. These amino acids are shown as red spheres. Inset shows the proximity of the three amino acids to the GDP molecule. The G1, Switch I and Switch II loops are indicated. **(b)** Clustering of residues of p.Asp342, p.Ser345 and p.Glu1163 on the surface of PLCG1 facing the plasma membrane. p.Ser345 has been previously shown to be a gain-of-function activating mutation in PLCG1⁴. p.Ser345 and p.Asp342 both cluster in the PLCx domain. These residues were mapped onto the crystal structure of PLCB2 (PDB ID 2ZKM)⁵, a closely related member of the PLC isozyme family. Domains of PLCB2 are indicated. C2, C2 calcium binding domain; EF, EF-hand-domain pair; PH, Pleckstrin homology domain; TIM, catalytic TIM barrel; X-Y linker, linker region between Phospholipase C X and Phospholipase C Y domains. Figure generated using CCP4MG⁶.

Supplementary Figure 9. Copy number plots of *NFKB2* in 3 samples with C-terminal truncating mutations.



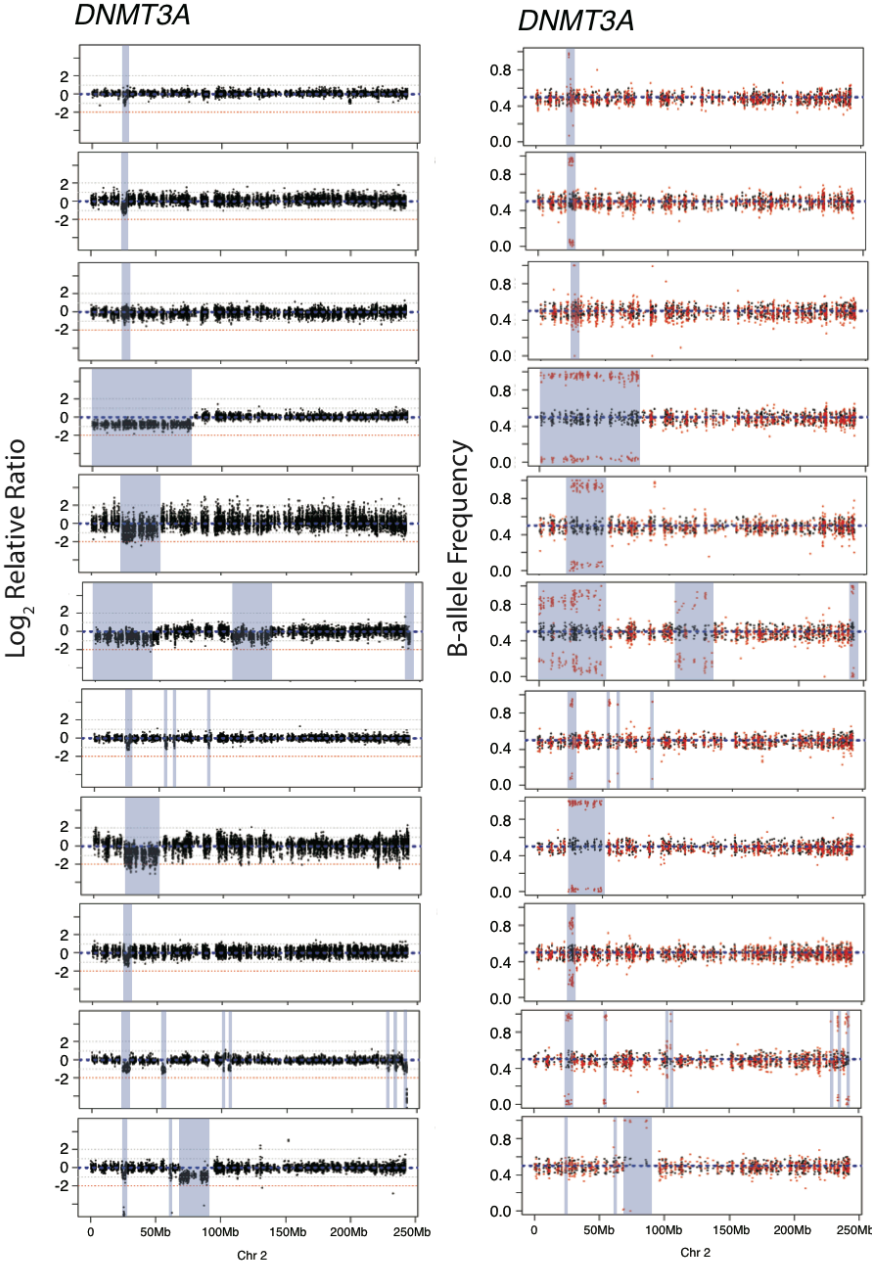
Log₂ relative ratio (log₂ ratio) plots of *NFKB2* in 3 samples and Hut-78, a CTCL cell line with a known truncation in *NFKB2*. Each point represents a targeted region from the Nimblegen capture array. Intragenic deletions are indicated in blue.

Supplementary Figure 10. *TP53* and *ARID1A* deletions in CTCL.



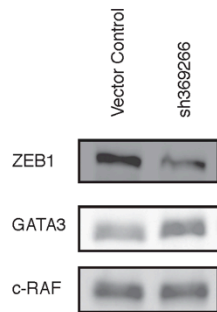
Representative log₂ relative ratio and B-allele frequency plots of 11 samples with deletions in (a) *TP53* (chromosome17) and (b) *ARID1A* (chromosome 1). Deletions are shaded blue, and amplifications shaded in red.

Supplementary Figure 11. *DNMT3A* deletions in CTCL.



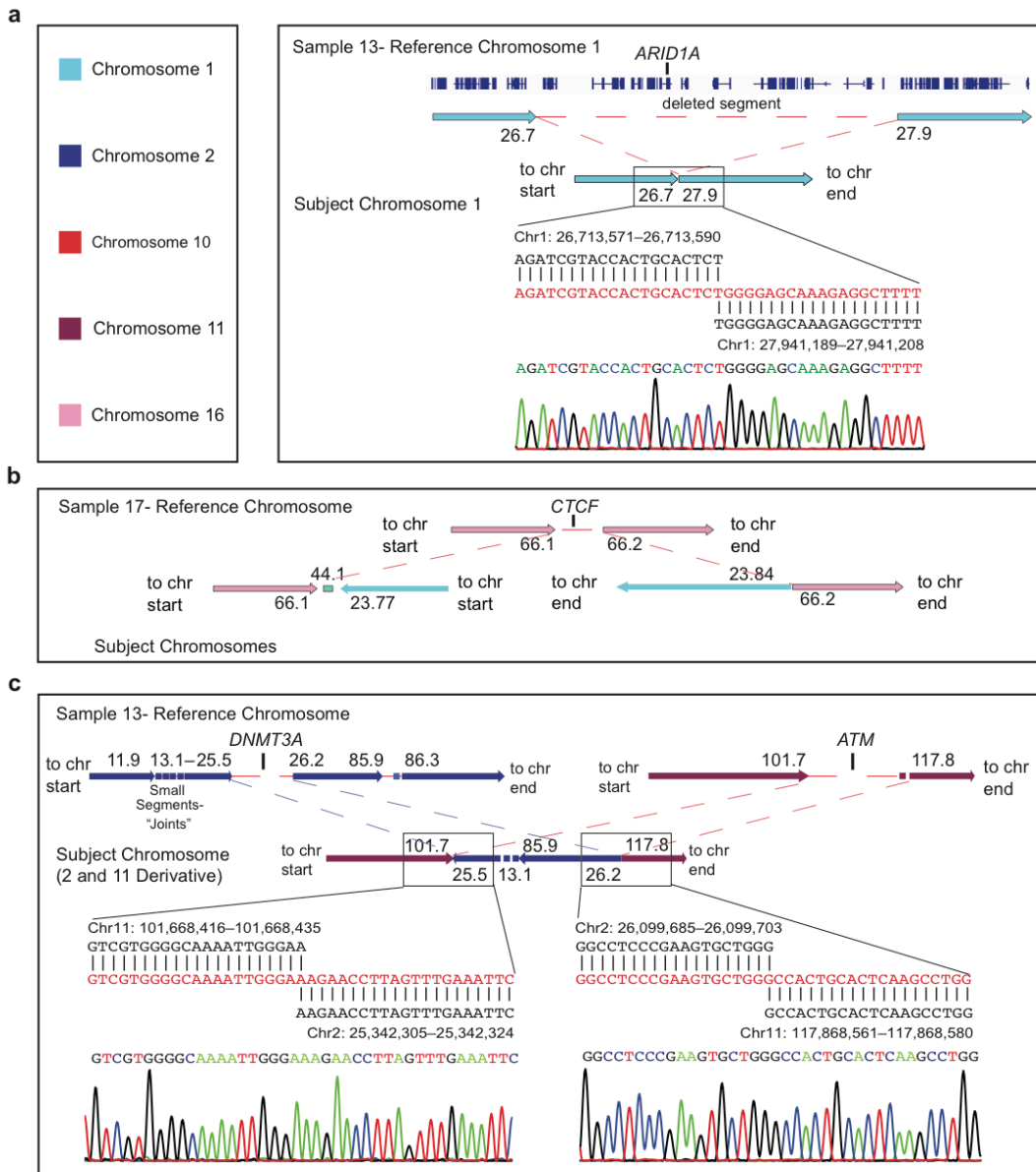
Representative log₂ relative ratio and B-allele frequency plots of 11 samples with deletions in *DNMT3A* (chromosome 2).

Supplementary Figure 12. Loss of ZEB1 by shRNA increases GATA3 expression



Western blots demonstrating the PMA/Ionomycin-induced expression of GATA3 in Jurkat cells transduced with pLKO vector control or pLKO targeting *ZEB1* (sh369266). C-RAF was used as a loading control.

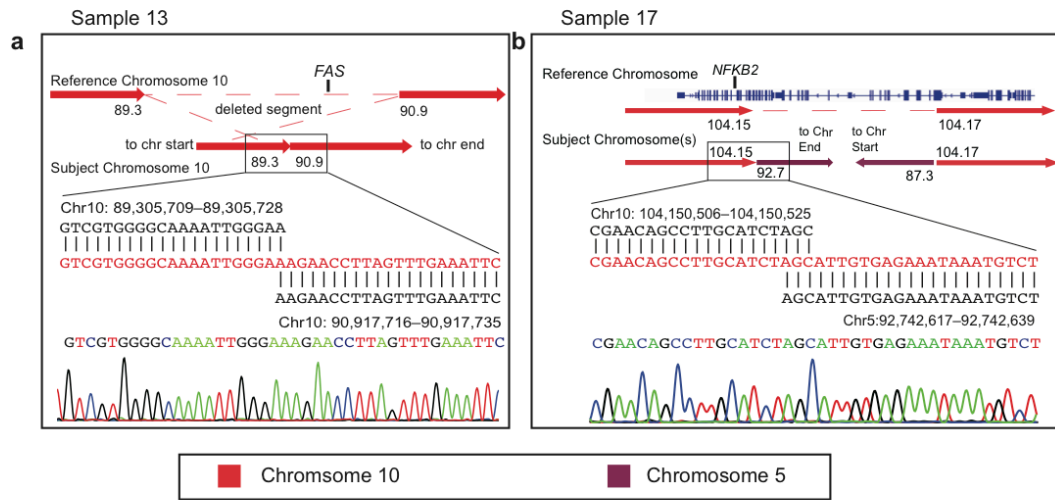
Supplementary Figure 13. Resolution by whole genome sequencing of the breakpoints contributing to structural variants in CTCL.



Structural variants leading to copy number loss of **(a) *ARID1A***, **(b) *CTCF***, and **(c) *DNMT3A*** and ***ATM*** elucidated by whole genome sequencing. Reference chromosomes reflect the wild-type chromosome. The dotted lines in the reference chromosome reference regions of genomic loss in CTCL. The subject chromosomes reflect the rearranged chromosome in the CTCL sample. Each arrow represents a contiguous chromosomal segment with directionality reflected by the arrowhead. The arrowhead

points in the direction of increasing nucleotide position number (hg18). Numbers listed are co-ordinates in hg18 (Mb). The numbers bordering each arrow reflects the nucleotide positions at closest breakpoint. Three breakpoints were confirmed by Sanger DNA sequencing.

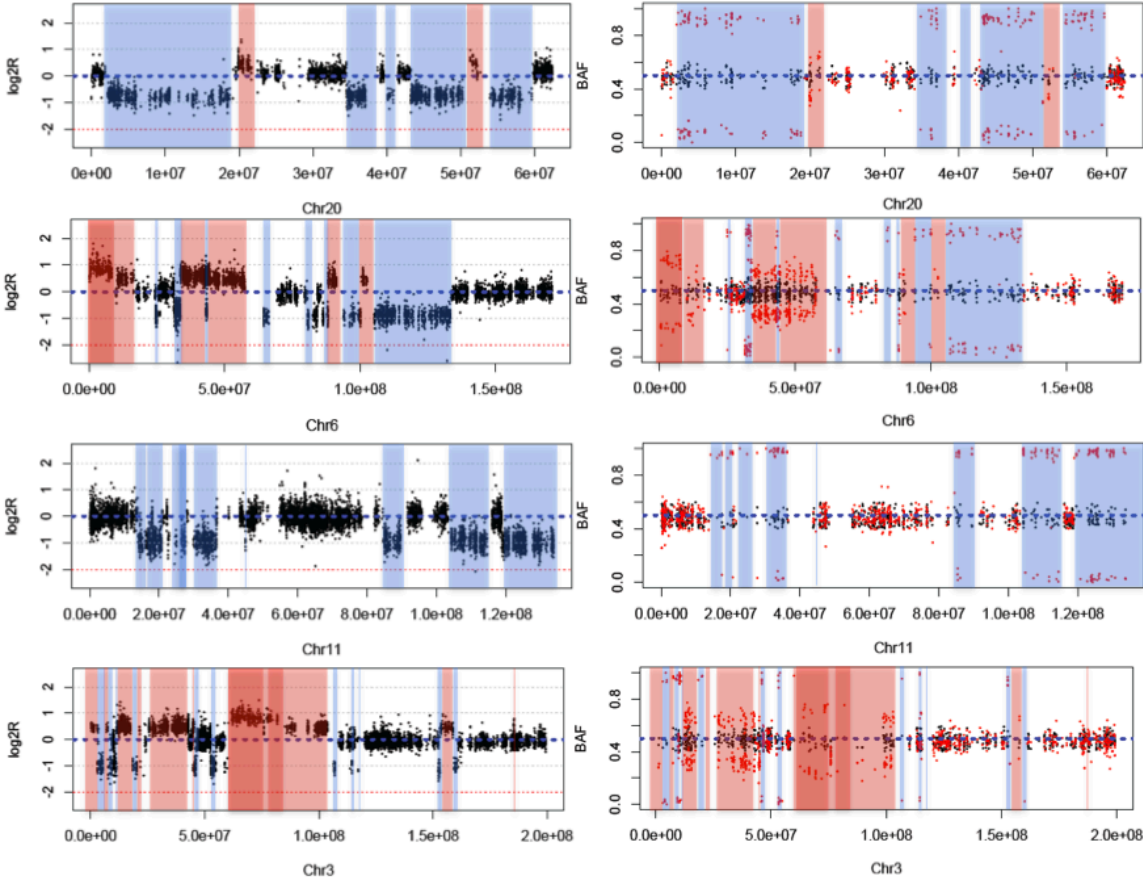
Supplementary Figure 14. Resolution by whole genome sequencing of the breakpoints contributing to significant structural variants on chromosome 10



Structural variants leading to **(a)** copy number loss of *FAS*, **(b)** C-terminal truncation of *NFKB2*. Reference chromosomes reflect the wild-type chromosome. The dotted lines in the reference chromosome represent regions of genomic loss in CTCL. The subject chromosomes reflect the rearranged chromosome in the CTCL sample.

Each arrow represents a contiguous chromosomal segment with directionality reflected by the arrowhead. The arrowhead points in the direction of increasing nucleotide position number (hg18). Numbers listed are co-ordinates in hg18 (Mb). The numbers bordering each arrow reflects the nucleotide positions at closest breakpoint. Two breakpoints were confirmed by Sanger DNA sequencing.

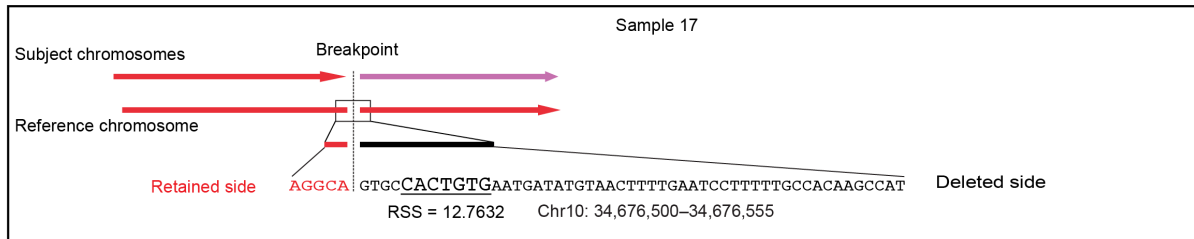
Supplementary Figure 15. Chromothripsis-like complex genomic rearrangements in CTCL.



Log₂ relative ratio (log₂R) and B-allele frequency (BAF) plots of 4 select chromosomes with >10 SCNVs.

Supplementary Figure 16. Enrichment of high scoring RAG heptamers at CTCL breakpoints.

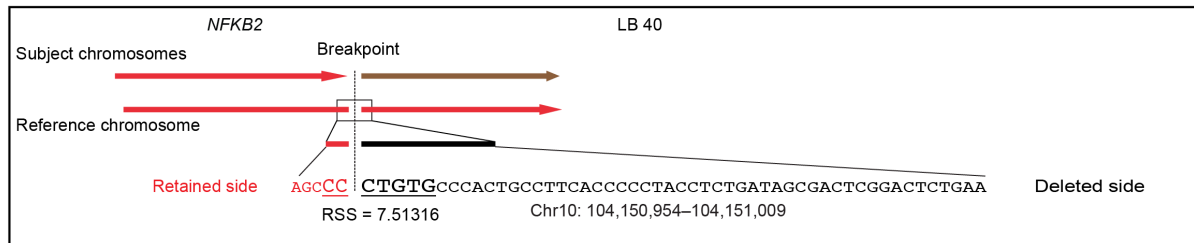
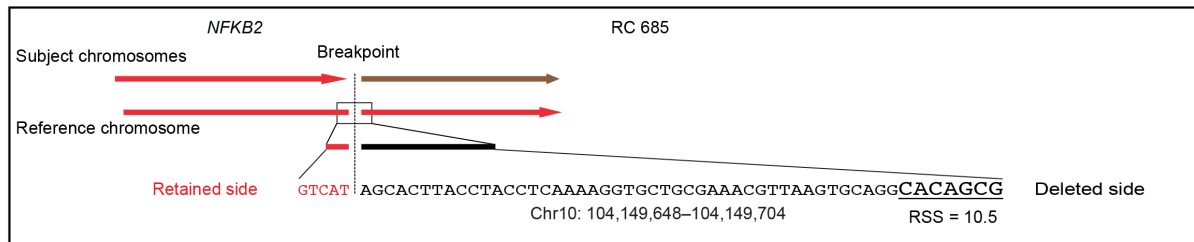
a



b

| Cancer Type | Total Breakpoints | Total breakpoint windows with high scoring heptamers | Prevalence of high scoring heptamers in breakpoint windows | P-value |
|-------------|-------------------|--|--|---------|
| CTCL | 102 | 13 | 12.7% | - |
| Gastric | 24,046 | 1,612 | 6.7% | 0.019 |
| Melanoma | 3,421 | 259 | 7.6% | 0.048 |
| Pancreatic | 754 | 52 | 6.9% | 0.035 |
| Prostate | 4,301 | 272 | 6.3% | 0.014 |

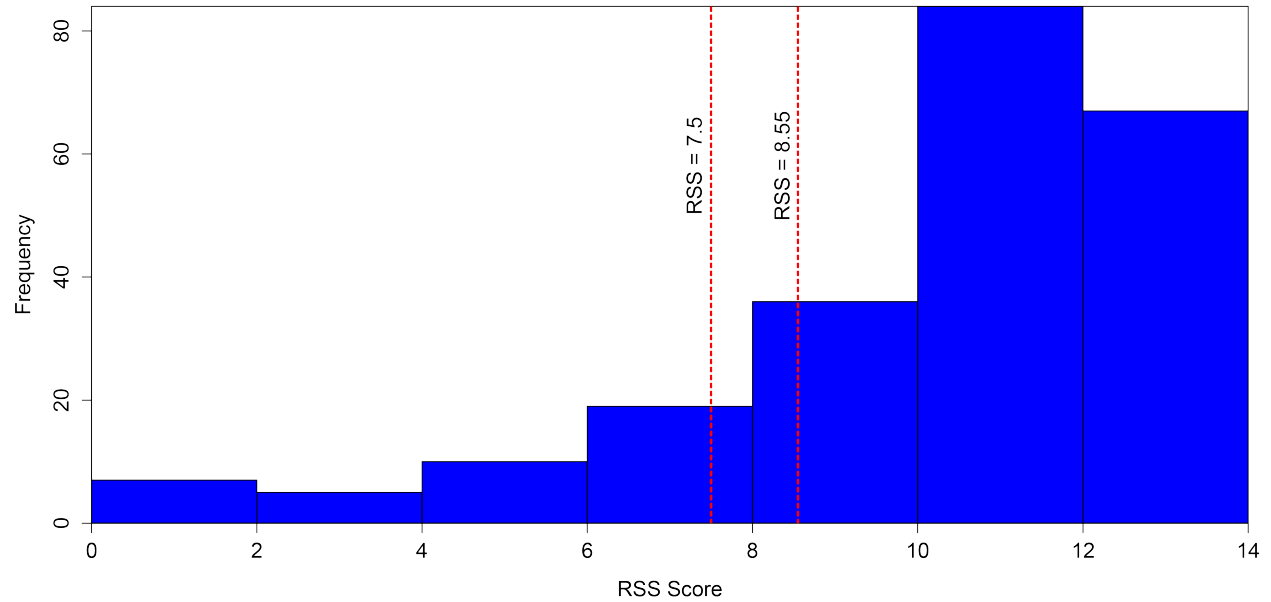
c



(a) Schematic of the 55 bp breakpoint window which was used to identify high scoring heptamers including 5 bp of the retained DNA and up to 50 bp of the deleted DNA. A high scoring heptamer with strong homology to RAG consensus cleavage sites (RSS>8.55) is underlined. (b) Table of frequency of high scoring RAG heptamers in in the breakpoint window of CTCL and other cancers. The prevalence of heptamers in the breakpoint window was compared between CTCL and other cancer types. P-value was established using one-sided Fisher's test. (c) Identification of high-scoring heptamers

(RSS>7.5) in the 55 bp window at the published breakpoints of oncogenic NF κ B2 C-terminal truncations of two B-cell lymphomas (RC655, LB40)⁷.

Supplementary Figure 17. RSS scores of heptamers in the endogenous human TCR loci.



RAG signal sequences at the endogenous human T cell receptor gene loci were obtained from RSSsite¹⁰. The heptamers at the recombination signal sequences were assessed for similarities to the consensus sequence (CACAGTG) using FIMO¹¹. 31 of 228 heptamers at RSS sites (13.6%) had an RSS score between 7.5 and 8.55.

References for Supplementary Information:

1. Sakata-Yanagimoto, M. *et al.* Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. *Nat Genet* **46**, 171-5 (2014).
2. Khosravi-Far, R., Solski, P.A., Clark, G.J., Kinch, M.S. & Der, C.J. Activation of Rac1, RhoA, and mitogen-activated protein kinases is required for Ras transformation. *Mol Cell Biol* **15**, 6443-53 (1995).
3. Wei, Y. *et al.* Crystal structure of RhoA-GDP and its functional implications. *Nat Struct Biol* **4**, 699-703 (1997).
4. Vaque, J.P. *et al.* PLCG1 mutations in cutaneous T-cell lymphomas. *Blood* (2014).
5. Hicks, S.N. *et al.* General and versatile autoinhibition of PLC isozymes. *Mol Cell* **31**, 383-94 (2008).
6. McNicholas, S., Potterton, E., Wilson, K.S. & Noble, M.E. Presenting your structures: the CCP4mg molecular-graphics software. *Acta Crystallogr D Biol Crystallogr* **67**, 386-94 (2011).
7. Neri, A., Fracchiolla, N.S., Migliazza, A., Trecca, D. & Lombardi, L. The involvement of the candidate proto-oncogene NFKB2/lyt-10 in lymphoid malignancies. *Leuk Lymphoma* **23**, 43-8 (1996).