Journal of NeuroVirology (1997) 3, 455–459

http://www.jneurovirol.com

Short Communication

C2H2-546: A zinc finger protein differentially expressed in HTLV-1 infected T cells

Paul D Drew[1], Ameer M Gado[1], Rachel D Canning[1], James W Nagle[2], Anindya M Dehejia[3], Mihael H Polymeropoulos[3], William E Biddison[1], Steven Jacobson[1] and Kevin G Becker[1]

[1]Neuroimmunology Branch and [2]Laboratory of Neurogenetics, NINDS; [3]Gene Mapping Unit, Laboratory of Genetic Disease Research, NCHGR National Institutes of Health, Bethesda Maryland 20892, USA

We report the cloning and characterization of a novel cDNA termed C2H2-546 which encodes a C2H2-type zinc finger protein. C2H2-546 RNA is expressed in the HTLV-1 infected T cells examined which were derived from HAM-TSP patients, but not in T cells derived from ATL patients. The C2H2-546 gene is conserved in humans and primates and maps to chromosome 10q11.2, a site associated with a variety of cancers. Thus, C2H2-546 is a candidate regulatory molecule important in the formation of these tumors, and may serve as an important marker to distinguish HTLV-1 infected ATL *versus* HAM-TSP T cell lineages.

Keywords: HTLV-1; transcription factor; zinc finger; molecular pathogenesis

Zinc finger domains are found in a variety of transcription factor families (Pieler and Bellefroid, 1994). C2H2-type zinc finger proteins possess two conserved cysteines which are part of an antiparallel beta-sheet and two conserved histidines which are part of an alpha-helix, coordinated by a central zinc atom to form a globular domain. C2H2 zinc finger proteins comprise the largest family of transcription factors known with an estimated 300–500 genes encoding these proteins in the human genome (Bellefroid *et al*, 1989). The protein family is typified by the RNA Polymerase II transcription factor TFIIIA (Miller *et al*, 1985) and the gap gene product Kruppel (Schuh *et al*, 1986).

The function of most C2H2-type zinc finger proteins is unknown. However, individual proteins have been shown to be important in development, tumorigenesis, RNA metabolism and chromatin assembly. C2H2 zinc finger proteins are believed to affect gene expression through sequence-specific binding to DNA and/or RNA and through protein-protein interactions, and have been shown to function as transcriptional activators and repressors. In addition to the conserved zinc finger domain, some C2H2-type zinc finger proteins contain the conserved amino acid sequence TGEKP between adjacent fingers, as well as KRAB, FAX, tramtrack, POZ and homeodomains found outside of the zinc finger domain (referenced in Becker *et al*, 1995).

Previously, we isolated greater than 100 novel cDNAs encoding C2H2-type zinc finger proteins. This was accomplished by screening a human hippocampal cDNA library from a normal 2-year-old female using a degenerate oligonucleotide specific for C2H2-type zinc finger motifs (Becker *et al*, 1995). A series of RNA dot-blot analyses were performed in order to determine the differential expression of the RNA molecules encoding these proteins in a variety of tissues and cells. Interestingly, one of these clones, C2H2-546, was expressed in human T cells including Jurkat, but not in ATL lines including HUT-102 (data not shown). C2H2-546 was sequenced in both directions following creation of nested deletions. Sequencing reactions were performed on an Applied Biosystems robotic workstation and analyzed on an Applied Biosystems 373A automated DNA sequencer using fluorescent labeled vector primers. Sequences were compared to the non-redundant databases of the National Center for Biotechnology Information (NCBI) at the National Library of Medicine using the BLAST algorithm. The sequence reveals a cDNA of 1140 nucleotides, as presented in Figure 1. The open reading frame codes for a protein of 273 amino acids. An ATG coding for a candidate initiator methionine is found at the 5′ end of the cDNA. C2H2-546 contains seven zinc finger domains,

Correspondence: PD Drew, Neuroimmunology Branch, National Institute of Neurological Disease and Stroke, NIH, Building 10, Rm 5B16, Bethesda, Maryland 20892, USA
Received 17 February 1997; revised 14 July 1997; accepted 7 August 1997

```
1   CTGGCGCTGCTGGGGCTCGGCGNCGGCCTTTGTCTGCGGGCACGGCCGCTGCGGTGCTCA 60
    GGAACAGCCCATGGAAGAATCATATGAAGAGGTGGTGACTGAGGTCGTAAGCAGGAGTGG 120
    ACATGTTTGGATTTCCAACAGCTACCCTGCTGGACTGTCATGGAAGATATGCCCAGAATG 180
1     M  F  G  F  P  T  A  T  L  L  D  C  H  G  R  Y  A  Q  N  V

    TAGCGTTCTTCAATGTGATGACTGAAGCCCACCACAAATATGACCACTCTGAGGCTACAG 240
21    A  F  F  N  V  M  T  E  A  H  H  K  Y  D  H  S  E  A  T  G

    GATCCTCAAGCTGGGATATCCAAAATTCTTTCAGAAGAGAGAAGCTGGAACAAAAATCCC 300
41    S  S  S  W  D  I  Q  N  S  F  R  R  E  K  L  E  Q  K  S  P

    CAGATTCGAAGACACTACAGGAAGATTCACCTGGAGTGAGACAAAGGGTCTATGAGTGCC 360
61    D  S  K  T  L  Q  E  D  S  P  G  V  R  Q  R  V  Y  E  C  Q

    AGGAGTGTGGAAAATCCTTCCGGCAAAAAGGTAGTCTAACGTTACATGAGAGAATCCACA 420
81    E  C  G  K  S  F  R  Q  K  G  S  L  T  L  H  E  R  I  H  T

    CTGGTCAAAAGCCTTTTGAGTGCACCCACTGTGGAAAAAGCTTCAGGGCCAAAGGCAATC 480
101   G  Q  K  P  F  E  C  T  H  C  G  K  S  F  R  A  K  G  N  L

    TTGTTACACATCAACGGATACACACGGGAGAGAAGCCTTATCAGTGCAAGGAGTGTGGGA 540
121   V  T  H  Q  R  I  H  T  G  E  K  P  Y  Q  C  K  E  C  G  K

    AAAGCTTCAGTCAACGAGGTAGTCTCGCTGTCCACGAGAGACTCCACACTGGACAGAAAC 600
141   S  F  S  Q  R  G  S  L  A  V  H  E  R  L  H  T  G  Q  K  P

    CCTACGAGTGTGCTATTTGTCAGAGAAGCTTCAGGAATCAGAGTAACCTTGCTGTTCACA 660
161   Y  E  C  A  I  C  Q  R  S  F  R  N  Q  S  N  L  A  V  H  R

    GGAGAGTTCACAGTGGTGAGAAGCCCTATAGATGTGATCAGTGTGGAAAAGCCTTCAGTC 720
181   R  V  H  S  G  E  K  P  Y  R  C  D  Q  C  G  K  A  F  S  Q

    AGAAAGGAAGCTTAATTGTTCACATCAGAGTCCACACAGGCCTGAAGCCCTATGCCTGTA 780
201   K  G  S  L  I  V  H  I  R  V  H  T  G  L  K  P  Y  A  C  T

    CCCAGTGCAGGAAGAGTTTCCACACCAGGGGGGaATTGTATTCTGCATGGCAAAATCCACA 840
221   Q  C  R  K  S  F  H  T  R  G  N  C  I  L  H  G  K  I  H  T

    CAGGAGAGACACCCTATCTGTGCGGCCAGTGTGGAAAAAGCTTCACCCAGAGAGGGAGTC 900
241   G  E  T  P  Y  L  C  G  Q  C  G  K  S  F  T  Q  R  G  S  L

    TGGCTGTGCACCAGCGAAGCTGCTCACAGAGGCTCACCCTTTGACCACTTTCCTGAAGAG 960
261   A  V  H  Q  R  S  C  S  Q  R  L  T  L  *  273

    AAGTTCTCTTTATGAATTAAGAGTACAAAATCCTCTGAGATGAAGCAACCTATCCAGTTC 1020
    TATGGAATGAATGGAGAATCTTTCAGAAAGACCATCATTGGGTAGGGCAAACTGATTTTT 1080
    TTCCTTTCCCCCAAAAGAGTATGAAAAATAAAGGTCTTGTTTATTATCaTtAAAAAAAAA 1140
```

Figure 1 Nucleotide and predicted amino acid sequences of C2H2-546. The nucleotide and amino acid sequences are numbered on the right and left, respectively. The zinc finger motifs are underlined. The nucleotide sequence appears in the GenBank sequence database (accession number U69645).

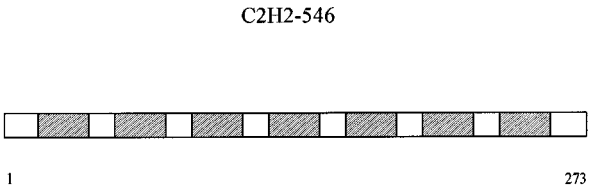

Figure 2 Schematic diagram of C2H2-546. Zinc finger domains are designated by hatched bars. Amino acids are numbered.

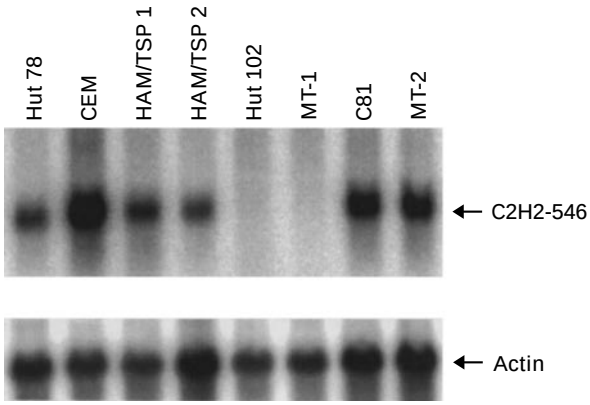

Figure 3 RNA expression of C2H2-546. Total RNA was isolated from the indicated cells, and Northern blots (30 microgram/lane) were prepared as described in the text. Blots were probed with full-length ^{32}P-labeled C2H2-546 cDNA.

approximately evenly spaced within the protein. The first six domains closely resemble the TFIIIA prototype C2H2 type zinc finger domain (Miller *et al*, 1985), conforming to the zinc finger consensus C-X2-C-X3-F-X5-L-X2H-X3-H (except finger six in which L is substituted). Interestingly, the carboxy-terminal zinc finger domain of C2H2-546 is a C2HC rather than a C2H2 zinc finger. To our knowledge, the presence of C2H2 and C2HC zinc fingers domains in the same cDNA has not been described previously in higher vertebrates. Mutations in the gene encoding the C2HC-type zinc finger protein longitudinals lacking (lola) lead to defects in axonal development in central and peripheral nervous system in *Drosophila* (Griniger *et al*, 1994). All zinc fingers are linked by the conserved H/C linker and spaced by seven amino acids (Schuh *et al*, 1986). C2H2-546 is represented schematically in Figure 2.

As stated previously, C2H2-546 RNA was originally observed in RNA dot-blot analyses to be present in T cells including Jurkat but absent from ATL lines including HUT-102 (data not shown). These data were supported by Northern blot analyses (performed as described previously by Drew *et al*, 1993) which indicated that ^{32}P-labeled C2H2-546 cDNA bound a single RNA species of approximately 1.2 kb in control HTLV-1 negative T cells including HUT-78 and CEM, but not the HTLV-1 expressing primary ATL derived T cell lines HUT-102 and MT-1 (Figure 3). This suggested that C2H2-546 RNA expression may be repressed in HTLV-1 infected T cells. However, C2H2-546 RNA is present in the HTLV-1 infected non-leukemic CD4$^+$ T cells derived from two different HAM-TSP patients designated HAM/TSP 1 and HAM/TSP 2. In addition, C2H2-546 RNA was present in two HTLV-1 expressing T cell lines, C81 and MT-2, derived by co-cultivation of ATL leukemic cells and non-leukemic CD4$^+$ T cells.

The present studies indicate that C2H2-546 RNA is present in human T cell lines, but absent from primary T cell lines (Hut-102 and MT-1) derived from the leukemic CD4$^+$ T cells of two HTLV-1 infected ATL patients (Poisz *et al*, 1980; 1981; Hinuma *et al*, 1981). This is supported further by the ability to detect C2H2-546 RNA in an HTLV-1 infected T cell line (C81) which was not a primary ATL isolate but rather a co-culture of ATL leukemic cells and non-HTLV-1 infected CD4$^+$ T cells. (Mitsuya *et al*, 1983; T Waldmann, personal communication). Of interest was the presence of C2H2-546 RNA in the MT-2 cell line which, at the time we obtained this line, was unclear whether this HTLV-1 infected ATL line was derived from a primary ATL isolate or a long-term co-cultivation. This would suggest that the HTLV-1 infected MT-2 cell line is likely to be a co-cultivation based upon

the expression of C2H2-546 RNA which could not be detected in two well defined HTLV-1 infected long-term T cell lines derived from two ATL patients, HUT-102 and MT-1, but detected in an HTLV-1 expressing long-term T cell line, C81, known to be derived from co-cultivation. Since HTLV-1 infected $CD4^+$ T cell lines derived from the PBL of two HAM-TSP patients (Jacobson *et al*, 1988) express C2H2-546 RNA, this indicates that viral infection or viral regulatory proteins do not repress C2H2-546 RNA in all infected cells. In fact, the disparate expression of C2H2-546 RNA may mark an important distinction between HTLV-1 infected ATL and HAM-TSP cells.

Southern blot analyses were performed to determine evolutionary conservation of the C2H2-546 gene. In these analyses, DNA isolated from the indicated species (Clontech) were digested with *Eco*RI. The DNA (10 g/lane) was run on 1% agarose gels in 1 × TBE. Following denaturation with NaOH and neutralization, DNA was transferred to Nylon membranes (MSI, Westboro, MA) as described previously (Ausebel *et al*, 1987). Blots were hybridized with C2H2-546 cDNA radio-labeled with ^{32}P by random priming (Prime-It, Stratagene) at 68°C in Quikhyb (Stratagene) as described by the manufacturer. Blots were washed twice at room temperature in 2 × SSC, 0.1% SDS, followed by two washes at 60°C in 1 × SSC, 0.1% SDS, and autoradiography was performed. Southern blot analysis indicated that the C2H2-546 gene is present in humans and primates, but not in lower vertebrates

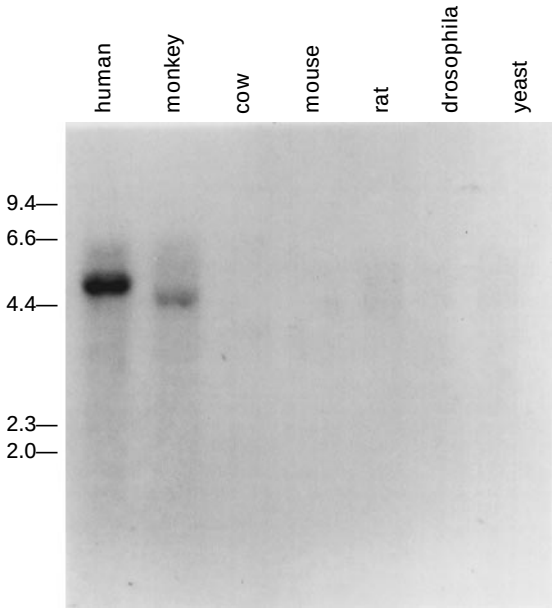

Figure 4 Genomic Southern blot analysis of C2H2-546. Genomic DNA (10 microgram/lane) from the indicated species was digested with *Eco*RI and Southern blots were prepared as described in the text. Blots were probed with full-length ^{32}P-labeled C2H2-546 cDNA. The position of molecular weight markers are indicated.

or invertebrates. (Figure 4). Southern analysis also indicated that some novel C2H2-type zinc finger genes were conserved in all vertebrate species analyzed (data not shown). This further supports the contention that the C2H2-546 gene is present only in higher vertebrates.

The chromosomal localization of C2H2-546 was determined by PCR analysis using somatic cell hybrid panels, radiation hybrid panels, and mega-YAC libraries. Oligonucleotides were designed to amplify a 161 bp product from the 3′ UTR of C2H2-546 (5′-TGTCACTGACAGTTTCTGAGGCA G-3′, 5′-GTCAGAGAGGAAGACTCAGACTAT-3′) in the PCR analysis. The NIGMS human-rodent panel #1 was used in somatic cell hybrid analysis (Polymeropoulos *et al*, 1993). The Genebridge 4 radiation hybrid panel (Walter *et al*, 1994) was used in radiation hybrid analysis. Statistical analysis of the data was performed using the RHMAPPER software package (D Slonim, L Stein, L Kruglyak and E Lander, unpublished software). The YAC pools screened were from the CEPH 'B' human mega-YAC library (Research Genetics, Huntsville, AL) (Berry *et al*, 1995). Microsatellite markers on or near the YAC address and the position of the YAC relative to known markers was determined using the program 'yacsr' (MH Polymeropoulos *et al*, unpublished). A cytogenetic location was determined by searching the cytogenetic location of nearby markers (MIT/Whitehead and Genome Database on-line database). C2H2-546 is contained on the YACs 895−D−3, 897−H−4, and 941−B−10, all of which are on the contig 10.3. This contig contains the marker WI-9393, whose cytogenetic location has been determined to be 10q11.2. Radiation hybrid mapping placed C2H2-546 9.87cR telomeric to the marker WI-7098, and between this marker and D10S196, both of which have been cytogenetically mapped to 10q11.2. The data vector was as follows: 021020200 122000101200010100000010100000101020002000000000200200010010201000211021002001102201. Thus, we have localized the C2H2-546 gene to chromosome 10q11.2 by somatic cell hybrids, radiation hybrids, and mega YAC libraries. The amino-terminus of C2H2-546 is homologous to a partial sequence of the uncharacterized clone ZNF 32 (KOX 30). ZNF 32 was mapped by *in situ* hybridization to human chromosome 10q23-q24. (Cannizzaro *et al*, 1993). The reason for this discrepancy is unclear, but our detailed analysis indicates that C2H2-546 maps to chromosome 10q11.2. In addition to C2H2-546, other C2H2-type zinc finger encoding genes including ZNF 11, ZNF 22, and ZNF 25 have been mapped to chromosome 10q11.2 (Rousseau-Merck *et al*, 1992). However, the function of this cluster of related genes has not been elucidated. A variety of tumors have been associated with the human chromosome 10q locus. Examples include melano-

ma (Walker *et al*, 1995), prostate cancer (Gray *et al*, 1995), Non-Hodgkins lymphoma (Speaks *et al*, 1992), acute lymphoblastic leukemia (Dube *et al*, 1991), adult T cell leukemia (Miyamoto *et al*, 1984), and glioblastoma (Fults and Pedone, 1993). Thyroid carcinomas (Herrmann *et al*, 1991) and multiple endocrine neoplasia which is frequently associated with mutations in the RET oncogene (Mulligan *et al*, 1993) are specifically associated with mutations at chromosome 10q11.2. Interestingly, C2H2-type zinc finger proteins have previously been associated with tumorigenesis (Call *et al*, 1990; Gessler *et al*, 1990; Chen *et al*, 1993; Kerckaert *et al*, 1993). This suggests the possibility that C2H2-546 may play a role in cancers associated with mutations at chromosome 10q11.2.

References

Ausebel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (eds) (1987). In: *Current Protocols in Molecular Biology*. John Wiley and Sons: New York.

Becker KG, Nagle JW, Canning RD, Biddison WE, Ozato K, Drew PD (1995). Rapid isolation and characterization of 118 novel C2H2-type zinc finger cDNAs expressed in human brain. *Hum Mol Genet* **4:** 685–691.

Bellefroid E, Lecocq PJ, Benhida A, Poncelet DA, Belayew A, Martial JA (1989). The human genome contains hundreds of genes coding for zinc finger proteins of the *Krüppel* type. *DNA* **8:** 377–387.

Berry R, Walter NAR, Stevens TJ, Rubano T, Wilcox AS, Hopkins JA, Sikela JM (1995). A strategy for the identification of candidate genes for alcohol-related phenotypes and other human disorders using rapid polymerase chain reaction mapping of gene-based sequence-tagged sites. *Alcohol Clin Exp Res* **19:** 821–823.

Call KM, Glaser T, Ito CY and Buckler AJ (1990). Isolation and characterization of a zinc-finger polypeptide gene at the human chromosome 11 Wilms' tumor locus. *Cell* **60:** 509–520.

Cannizzaro LA, Aronson MM, Thiesen HJ (1993). Human zinc finger gene ZNF 23 (Kox 16) maps to a zinc finger gene cluster on chromosome 16q22, and ZNF 32 (Kox 30) to chromosome region 10q23-q24. *Hum Genet* **91:** 383–385.

Chen Z, Brand NJ, Chen A, Chen SJ, Tong JH, Wang ZY, Waxman S, Zelent A (1993). Fusion between a novel Kruppel-like zinc finger gene and the retinoic acid receptor-alpha locus due to a variant t(11;17) translocation associated with acute promyelocytic leukaemia. *EMBO J* **12:** 1161–1167.

Drew PD, Lonergan M, Goldstein ME, Lampson LA, Ozato K, McFarlin DE (1993). Regulation of MHC class I and B2-microglobulin gene expression in human neuronal cells: Factor binding to conserved cis-acting regulatory sequences correlates with expression of the genes. *J Immunol* **150:** 3300–3310.

Dube ID, Kamel-Reid S, Yuan CC, Lu M, Wu X, Corpus G, Raimondi SC, Crist WM, Carroll AJ, Minowada J, Baker JB (1991). A novel human homeobox gene lies at the chromosome 10 breakpoint in lymphoid neoplasias with chromosomal translocation t(10;14). *Blood* **78:** 2996–3003.

Fults D, Pedone C (1990). Deletion mapping of the long arm of chromosome 10 in glioblastoma multiforme. *Genes, Chromosomes and Cancer* **7:** 173–177.

Gessler M, Poustka A, Cavenee W, Neve RL, Orkin SH, Bruns GAP (1990). Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chromosomal jumping. *Nature* **343:** 774–778.

Gray IC, Phillips SMA, Lee SJ, Neoptolemos JP, Weissenbach J, Spurr NK (1995). Loss of the chromosomal region 10q23-25 in prostate cancer. *Cancer Res* **55:** 4800–4803.

Griniger E, Tietje K, Jan LY, Jan YN (1994). lola encodes a putative transcription factor required for axon growth and guidance in Drosophila. *Development* **120:** 1385–1398.

Herrmann MA, Hay ID, Bartelt DH Jr, Ritland SR, Dahl RJ, Grant CS, Jenkins RB (1991). Cytogenetic and molecular genetic studies of follicular and papillary thyroid cancers. *J Clin Invest* **88:** 1596–1604.

Hinuma Y, Nagata K, Hanaoka M, Nakai M, Matsumoto T, Kinoshita K-I, Shirakawa S, Miyoshi I (1981). Adult T-cell leukemia: Antigen in an ATL cell line and detection of antibodies to the antigen in human sera. *Proc Natl Acad Sci USA* **78:** 6476–6480.

Jacobson S, Raine CS, Mingiola ES, McFarlin DE (1988). Isolation of an HTLV-1-like retrovirus from patients with tropical spastic paraparesis. *Nature* **331:** 540–543.

Kerckaert JP, Deweindt C, Tilly H, Quief S, Lecocq G, Bastard C (1993). LAZ3, a novel zinc-finger encoding gene, is disrupted by recurring chromosome 3q27 translocations in human lymphomas. *Nature Genet* **5:** 66–70.

Miller J, McLachlan AD, Klug A (1985). Repetitive zinc-binding domains in the protein transcription factor TFIIIA from *Xenopus laevis*. *EMBO J* **4:** 1609–1614.

Mitsuya H, Matis LA, Megson M, Bunn PA, Murray C, Mann DL, Gallo RC, Broder S (1983). Generation of an HLA-restricted cytotoxic T cell line from a patient infected with human T cell leukemia/lymphoma virus. *J Exp Med* **158:** 994–999.

Miyamoto K, Tomita N, Ishii A, Nonaka H, Kondo T, Tanaka T, Kitajima K (1984). Chromosome abnormalities of leukemia cells in adult patients with T-cell leukemia. *JNCI* **73:** 353–362.

Mulligan LM, Kwok JBJ, Healey CS, Elsdon MJ, Eng C, Gardner E, Love DR, Mole SE, Moore JK, Papi L, Ponder MA, Telenius H, Tunnacliffe A, Ponder BAJ (1993). Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. *Nature* **363:** 458–460.

Pieler T, Bellefroid E (1994). Perspectives on zinc finger function and evolution-An update. *Mol Biol Reports* **20:** 1–8.

Poiesz BJ, Ruscetti FW, Gazdar AF, Bunn PA, Minna JD, Gallo RC (1980). Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci USA* **77:** 7415–7419.

Poiesz BJ, Ruscetti FW, Reitz MS, Kalyanaraman VS, Gallo RC. (1981). Isolation of a new type-C retrovirus (HTLV) in primary uncultured cells of a patient with Sezary T-cell leukemia. *Nature* **294:** 268–271.

Polymeropoulos MH, Xiao H, Sikela JM, Adams MD, Venter JC, Merril CR (1993). Chromosomal distribution of 320 genes from a brain cDNA library. *Nature Genet* **4:** 381–386.

Rousseau-Merck MF, Tunnacliffe A, Berger R, Ponder BAJ, Thiesen HJA (1992). Cluster of expressed zinc finger protein genes in the periocentromeric region of human chromosome 10. *Genomics* **13:** 845–848.

Schuh R, Aicher W, Gaul U, Cote S, Preiss A, Maier D, Seifert E, Nauber U, Schroder C, Kemler R, Jackle H (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a *Drosophila* segmentation gene. *Cell* **47:** 1025–1032.

Speaks SL, Sanger WG, Masih AS, Harrington DS, Hess M, Armitage JO (1992). Recurrent abnormalities of chromosome bands 10q23-25 in non-Hodgkin's lymphoma. *Genes, Chromosomes, and Cancer* **5:** 239–243.

Walter MA, Spillet DJ, Thomas P, Weissenbach J, Goodfellow PN (1994). A method of constructing radiation hybrid maps of whole genomes. *Nature Genet* **7:** 22–28.

Walker GJ, Palmer JM, Walters MK, Hayward NK (1995). A genetic model of melanoma tumorigenesis based on allelic losses. *Genes, Chromosomes and Cancer* **12:** 134–141.