Comparative integromics on the breast cancer-associated gene *KIAA1632*: Clues to a cancer antigen domain

NIELS HALAMA[1], SILKE A. GRAULING-HALAMA[2], ANDREAS BEDER[3] and DIRK JÄGER[1]

[1]National Center for Tumor Diseases, University of Heidelberg, Heidelberg; [2]Institute for Microbiology and Hygiene, Clinical Faculty Mannheim, University of Heidelberg, Mannheim; [3]Department of Computer Science, Darmstadt University of Technology, Darmstadt, Germany

Received February 1, 2007; Accepted March 12, 2007

Abstract. Sequence analysis of protein-coding human genes in human breast and colorectal cancer recently identified the *KIAA1632* gene as a potential contributor to neoplastic processes in breast cancer. In the present study, we characterized the *KIAA1632* gene by computational methods: detailed investigation of the genomic structure, protein prediction, identification of orthologs in other species and phylogenetic analysis. The human *KIAA1632* gene was located within human genome sequence AC090355 and consists of 44 exons. *KIAA1632* is located on chromosome 18q12.3-q21.1. These findings were determined by aligning the potential full-length transcript NM_020964 to the genomic sequence. The existing predicted gene model could be refined based on these alignments (using additional EST data). Protein sequence was predicted from the full-length transcript and possible orthologs in other species were identified. Using rigorous phylogenetic methods we were able to draw conclusions about the evolutionary history of these proteins. Comparison of the protein sequences of *KIAA1632* and its orthologs revealed a shared 'signature motif'. This 'signature motif' was found in three well-characterized cancer antigens, raising the hypothesis of a common 'signature motif' for a certain group of cancer antigens.

Introduction

Cancer research has shed light on the molecular basis of malignancy. With the identification of the human genome sequence even more detailed identification of genetic alterations in cancer is possible (1). One recent approach was the systematic analysis of 13023 genes in 11 breast and 11 colorectal cancers (2). A group of 189 genes with significant genetic alterations was detected, the majority being not analysed previously in cancer. The identified genes are predicted to alter a broad range of cellular functions, e.g. transcription, adhesion and invasion. Many of these cancer-associated genes are well characterized but a minority have not been described in detail before. One interesting candidate is the *KIAA1632* gene on chromosome 18q22.1. No detailed information on the genomic structure and possible proteins is available. However, a report on pericentric inversions on chromosome 18 involves the genomic region where supposedly *KIAA1632* is located. These inversions are associated with psychiatric disorders and numerous genes are possible candidates for the pathomechanism of the observed condition (3). In summary, *KIAA1632* is an interesting candidate gene for the molecular analysis of breast cancer and it may even play an additional role in psychiatric disorders.

In this report we characterize the *KIAA1632* gene and its orthologs in other species by bioinformatical methods. Coding region of human *KIAA1632* (AK023817 and NM_020964) was found in the HUGO gene database (http://www.gene.ucl.ac.uk/nomenclature/). Using the sequence data from NM_020964 and available ESTs we refined the full-length *KIAA1632* sequence. Genomic structure (intron and exon boundaries), expression data, functional relevant SNPs (single nucleotide polymorphisms), domain topology and comparative protein homology are discussed. Furthermore, we identified a 'signature motif' (deduced from *KIAA1632* and its orthologs) that is also present in other cancer antigens and may lead to the identification of other cancer-associated genes. This is the first report on the characterization, phylogeny and the possible functional role of the *KIAA1632* gene.

Materials and methods

Identification of a novel gene. Automatic annotation of the human genome has produced numerous entries but in many cases these entries are automated computational predictions and are therefore often incomplete. Human genome sequences, expressed sequence tags (ESTs) and uncharacterized cDNAs were analysed with BLASTN and MegaBLAST programs (http://www.ncbi.nlm.nih.gov/BLAST) (4-6). Available partial nucleotide sequence of human *KIAA1632* was used as a query sequence (7,8). Additional databases were used for refinement of existing predictions.

Structural and chromosomal localization of the KIAA1632 gene. Exon-intron boundaries were determined by examining

Correspondence to: Dr Niels Halama, National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany
E-mail: nhalama@gmx.net

Key words: comparative integromics, breast cancer, *KIAA1632*

Table I. Exon-intron structure of human *KIAA1632*.

Exon no.	Intron (lead)	Exon	Intron (trail)	Genomic coord. (AC090355)	Length (base pairs)
1	GGAGTGGCGG	AGATTC...ACAAAG	GTGAGGATCT	136358-136520	163
2	TTGTTTTTAG	GAAAAG...GTACAG	GTATTTTGAT	123575-124519	945
3	CCTCTGGTAG	GGTATC...GCCGAG	GTATGTAAGT	121581-121824	244
4	TCATTTACAG	CGTCTA...AAATTG	GTAAGGGGAC	120283-120419	137
5	TATTGTTCAG	GTATCC...ATCCAG	GTATGATGAA	118665-118772	108
6	CTGATTTCAG	ATCAAA...TGTCAA	GTAAGTGTGC	117684-117757	74
7	TGTCCTTTAG	AAATCG...GAAGAG	GTAACTTGTC	115844-115949	106
8	TTTCCCTTAG	GATGAA...CAAAAG	GTAGACTCAT	113186-113300	115
9	TAACTTCTAG	GTGATT...TATGAT	CAGGTAATAC	112345-112492	148
10	TCCTCATGCT	CAGGAT...CAAAAG	GTAAGTAAGA	108781-108939	159
11	TTACATGCAG	ACTTGG...TCCAAG	GTACAGCCCC	104147-103990	158
12	GAACTTGCAG	ACCCAG...TATGAG	GTTAGAATAT	99857-100011	155
13	CCCCATTTAG	GTATCT...GGAATG	GTAAGAGCAC	98050-98190	141
14	TTTTAATAAG	GTTTCT...AAACAG	GTATGTGTTC	94919-95083	165
15	TCTTTAAAAG	GCTACC...AAGCAG	GTATTTCCTT	92449-92568	120
16	ATTGTTTTAG	GTTTCA...CCACAG	GTACAGTGTT	91522-91781	260
17	TTTTCCATAG	CATTGA...TGAGCA	GTAAGTGTGT	86859-86999	141
18	TGCTTGGCAG	GTTTTT...ATCCAG	GTGAGGGCCT	85618-85762	145
19	TTCTCCACAG	GCACAC...CATAAG	GTAATTAACC	85189-85386	198
20	CCCGGTGTAG	AATGCC...ACTCAG	GTAGGAAACT	84691-84801	111
21	GTCATTATAG	GTCTGG...CTGAAG	GTACAAATAG	82886-83008	21
22	TTTTTTGTAG	AAAGCC...GTATGG	GTATGCTGCC	81470-81636	167
23	GTTTTTTAAG	GTTACC...GGTGAG	GTGAGTATGG	79701-79922	222
24	TATCTTTTAG	GCTCTT...CAGCAG	GTGAAATTAT	77138-77261	124
25	TCTGTTTCAG	GATCTG...TGCTGG	GTGGGTCCTG	73153-73297	145
26	TTTCTTTCAG	CTAAAG...GGCCAG	GTAAGACCAT	70176-70347	172
27	TGTTTTTCAG	AACCGC...GTTCAG	GTAAGTGTTT	68584-68746	163
28	TTTTTGTTAG	TTTGAA...GATCAC	GTAAGTTTTG	58978-59120	143
29	TTTCCCACAG	GGCCTT...GGACAG	GTAAGGCTAT	56931-57087	157
30	TTTTTCCAAG	GTATTT...ACCAAG	GTAAGATGTA	53797-53991	195
31	TTCTTTTTAG	TTCGAT...TGCAAA	GTAAGTGTCT	51454-51667	214
32	CTGCTGGCAG	GCTCCG...AAGCAG	GTGTGTAGCC	49255-49403	149
33	TCTCTTCCAG	GTAATG...AAACAG	GTAATATACT	48193-48394	202
34	AACTTTTTAG	TGCTGA...CGAAAG	GTAAACAGTT	47556-47628	73
35	TTCTTTAAAG	CAGCCA...TTAAAG	GTAGGAAGAT	45416-45522	107
36	CACATTCTAG	ATAAGC...TTCAAA	GTAAGCCCTT	39747-39922	176
37	TTTCTTTTAG	GTGGAA...CAAACA	GTAAGTTTGG	36928-36743	186
38	TTATTTTTAG	GATTCA...CATCTT	GTAAGTTGAG	35978-36187	210
39	TTTTGGTCAG	GATGCA...CGCCCG	GTAAGTGACC	34795-34939	145
40	TTCCTTTCAG	ACATGG...AAGAAA	GTAAGAGGAG	29285-29527	243
41	ATGTTCTTAG	GCCCTC...CCCAAG	GTACGTTGGC	27747-27963	217
42	CCTCCCCTAG	CTCCGT...TAACAG	GTAAGGAAGC	27034-27249	216
43	AATGCCTAAG	GTTCCG...CAGCAG	GTCAGTATAT	4754-24868	115
44	TGTCTTTTAG	GCTCTG...TTTACA	CATGGCGGGC	16788-21830	5043

the consensus sequence of exon-intron junctions ['gt...ag' rule of intronic sequence, as described previously (9,10)] and the codon usage within the coding region. To refine the exon-intron boundaries the existing cDNAs or ESTs (expressed sequence tags) were aligned to the genomic sequence (contig accession number AC090355) using SPIDEY (http://www.

ncbi. nlm.nih.gov/IEB/Research/Ostell/Spidey/). This revealed a different genomic structure of *KIAA1632* as predicted by the automated computational analysis. These findings were then used to identify the human chromosomal localization of the genome clones in the Ensembl database (http://www.ensembl.org/) and subsequently the precise human chromosomal localization of *KIAA1632* (11,12). Existing alternative splicing databases were searched for known alternatively spliced transcripts [EBI Alternate Splicing Database http://www.ebi.ac.uk/asd/ (13-15), Alternative Splicing Database http://hazelton.lbl.gov/~teplitski/alt/ (16,17), Human Alternative Splicing Database http://www.bioinformatics.ucla.edu/~splice/HASDB/] but did not yield any information on alternative transcripts of *KIAA1632*.

Analysis of deduced amino acid sequence and phylogenetic calculations. Using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and the protein database (http://www.expasy.uniprot.org/) (18) we predicted the coding region. Translation into amino acid sequence was performed using ORF Finder, and analysis of the identified domains was performed using InterProScan (http://www.ebi.ac.uk/InterProScan/) and the Pfam program (http://pfam.wustl.edu/). Searching PROSITE (http://www.expasy.org/prosite/) for known domains did not reveal any significant new information. With PRATT (http://www.expasy.org/tools/pratt/) (19,20) however we detected 108 possible domain-homologs, for a more detailed approach we used the domain homologs found in mammals. The comparative analysis of *KIAA1632* in different species was performed using ClustalW (http://www.ebi.ac.uk/clustalw/) and the Ensembl (http://www.ensembl.org) database. Orthologs in other species were detected using PSI-BLAST (6) against the predicted amino acid sequence of *KIAA1632*. Phylogenetic analyses were performed using the MEGA3 package (21) and the DAMBE software (22). Neighbour-joining (NJ) analysis, minimum evolution (ME) calculation, maximum parsimony (MP) and UPGMA analysis were performed with bootstrapping of 1000 replicates. Additionally TreePuzzle was used to generate maximum-likelihood distance matrices with 1000 quartet puzzling steps to assess branch support (23).

Functional relevant SNP evaluation. To identify possible functional relevant SNPs (single nucleotide polymorphisms) that could disrupt ESE/ESS (exonic splicing enhancer/exonic splicing silencer) motifs (24,25) we extracted coding SNP data from Ensembl (http://www.ensembl.org) and NCBI's SNPdb (http://www.ncbi.nlm.nih.gov). Using RESCUE-ESE (26) web server software (http://genes.mit.edu/cgi-bin/rescue-ese_new.pl) we analyzed 10 bp in either direction around each SNP.

Expression profiling. The expression profiles for normal human tissues were obtained from GeneAnnot (twelve normalized tissues were hybridized against Affymetrix GeneChips HG-U95B and HG-U95E with optimal annotation quality, see http://bioinfo2.weizmann.ac.il/cgi-bin/geneannot/GA_search.pl?keyword_type=gene_symbol&keyword=*KIAA1632*) and from ArrayExpress (Affymetrix GeneChip HG-U133B). Furthermore, 'electronic Northern' analysis of NCBI's UniGene dataset was extracted from GeneCards (http://www.genecards.org).

Results

Human KIAA1632 gene. The data on *KIAA1632* genomic structure in available databases (ENSEMBL, AltSpliceDB, etc.) shows a gene with 24-26 exons. The BLAST analysis however showed a longer transcript with more exons. According to our findings the full-length transcript is NM_020964 (confirming the HUGO entry) and the novel corresponding genomic structure shows a gene with 44 exons (for details see Table I). *KIAA1632* was localized to the genomic region of chromosome 18q12.3-q21.1 due to the mapping of the human genome clone AC090355 to this region. Boundaries of exon-intron junctions were determined through alignment of ESTs to the genomic sequence and through observation of consensus sequences at exon-intron junctions. Human *KIAA1632* cDNA consists of a 7737-bp coding region, a 4860-bp 3'-UTR and it encodes a 2579-aa *KIAA1632* protein.

Human KIAA1632 protein. The putative transcript of the *KIAA1632* gene encodes a protein of 2579 amino acids. Further analysis with InterProScan (http://www.ebi.ac.uk/InterProScan/) did not reveal any known domains. The full protein sequence (of the full-length transcript) can be found at http://www.halama.org/KIAA1632prot.htm.

Comparative proteomics of KIAA1632 orthologs and phylogeny. Data mining revealed putative orthologs of human *KIAA1632* in different species. Orthologs were found in mammals, fish, flies and worms. These proteins were analyzed with MEME (27) which showed three distinct motifs. These motifs are depicted in Fig. 1, whereas motifs 'A' and 'B' are only found in mammals. The distribution of motifs across different species is shown in Fig. 2. Phylogenetic analysis was carried out using multiple phylogenetic models. The best supported phylogenic tree is shown in Fig. 3.

Possible functions of KIAA1632. Analysis with the PRATT algorithm identified a common motif for human KIAA1632 and its orthologs in other species. The consensus pattern for this domain is: E-x(1,3)-Q-[AEQS]-x(2)-L-x(3)-L-x(0,2)-L-[DER]. Using this pattern for a search in the PROSITE database showed several proteins containing similar domains. Interestingly three other cancer antigens were detected with this approach. Cancer antigens were tumor antigen se2-2 (28), NY-REN-58 (29) and CNG1 (30). The complete list of detected proteins can be found at http://www.halama.org/KIAA1632pratt.htm. The majority of mammalian domain homologs detected with PRATT are involved in cell-cell interaction and adhesion.

SNP analysis. From the data of three available SNPs (rs1893523, rs3744996 and rs3744998) in the genomic region of *KIAA1632* we identified three as functionally relevant, i.e. one of the available alleles disrupted an existing exonic splicing enhancer or the SNP resulted in a non-synonymous amino acid change. Sequences and available allele frequencies for rs3744996 are shown in Fig. 4. An amino acid change from Valin to Alanin was reported at amino acid residue 343 for rs1893523. An identical amino acid change was also reported for rs3744998 at residue 1059.

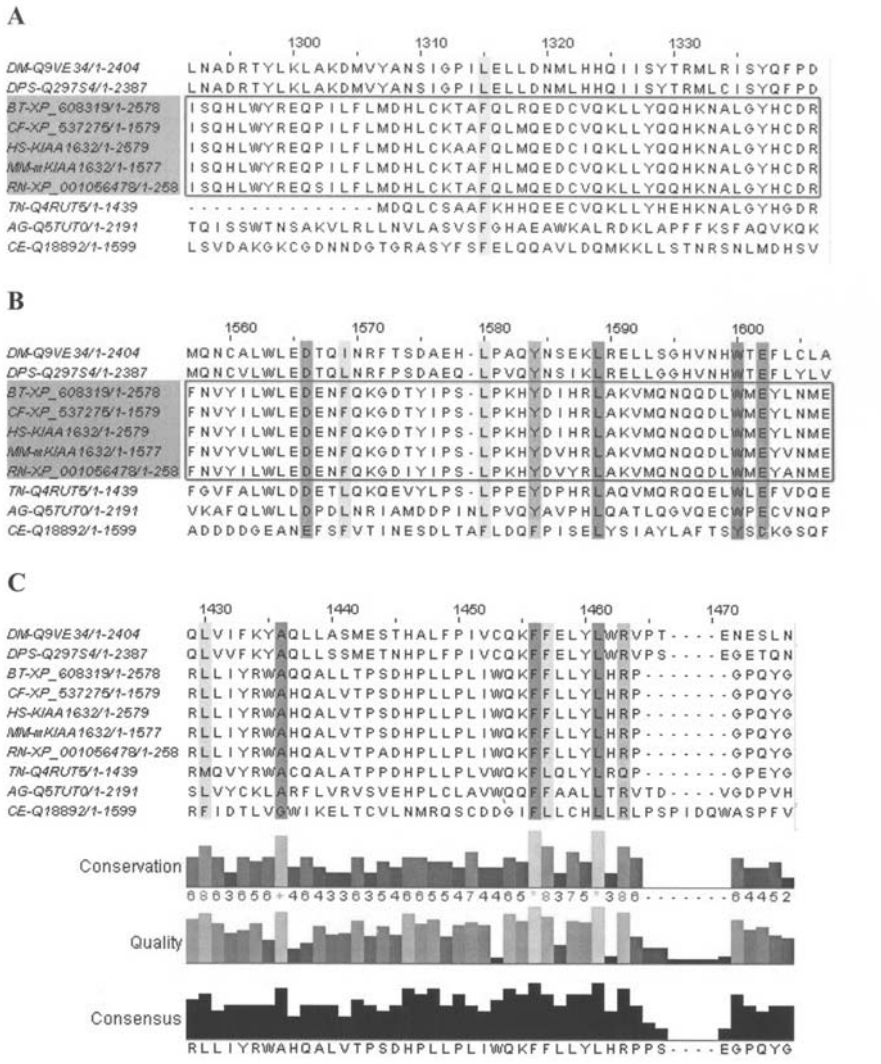


Figure 1. Three common motifs found in the investigated proteins. Motifs 'A' and 'B' are only present in mammals as shown by boxed sequence and horizontal shaded protein names and accession numbers. Vertical shaded amino acids represent conserved residues. For motif 'C' the consensus sequence, the conservation (position specific for every amino acid) and quality are shown.

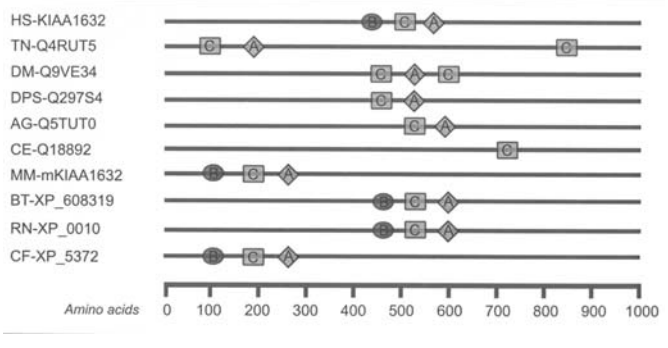


Figure 2. Distribution of motifs across different species. Symbols represent motifs and motifs are named according to Fig. 1. Grey lines present amino acids of the corresponding protein. Species and accession numbers are shown on the left. HS, *Homo sapiens*, *CF*, *Canis familiaris*; *MM*, *Mus musculus*; RN, *Rattus norvegicus*; BT, *Bos taurus*; TN, *Tetraodon nigroviridis*; AG, *Anopheles gambia*; DM, *Drosophila melano*gaster; DPS, *Drosophila pseudoobscura*; CE, *Caenorhabditis elegans*.

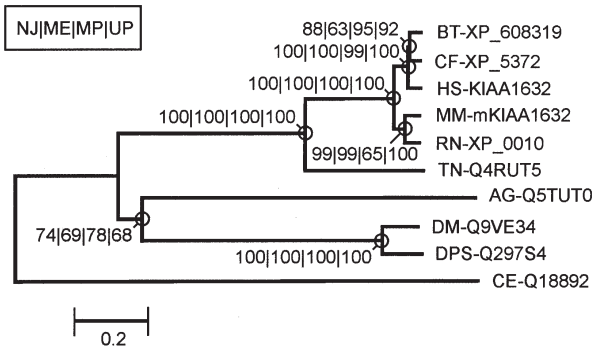
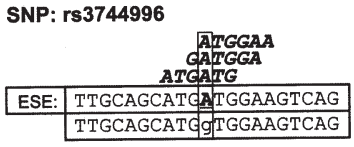


Figure 3. Phylogenetic analysis of human KIAA1632 protein and homologs. Unambiguously aligned amino acids were analysed by distance (Protdist + NJ), maximum parsimony (MP) minimal evolution (ME) and clustering likelihood (UPGMA: unweighted pair-group method using arithmetic averages) methods. The tree shown is the NJ distance topology. Numbers at the nodes represent the percent of bootstrap replicates in support of each group for the respective method (the small box in the upper left shows the order of presented numbers: NJ, neighbour-joining; MP, maximum parsimony; ME, minimal evolution and UPGMA). The numbers for the quartet puzzling steps (as generated with TreePuzzle) confirm the presented model (data not shown). HS, *Homo sapiens*; CE, *Canis familiaris*; MM, *Mus musculus*; RN, *Rattus norvegicus*; BT, *Bos taurus*; TN, *Tetraodon nigroviridis*; AG, *Anopheles gambia*; DM, *Drosophila melanogaster*; DPS, *Drosophila pseudoobscura*; CE, *Caenorhabditis elegans*.

Expression profiles of human KIAA1632. The investigation of available microarray experiments and the results of the 'virtual Northern blot' showed a predominant expression of



Allele frequencies per population

Population	A	G
European	0.312	0.688
African-American	0.217	0.783
Asian	0.250	0.750

Figure 4. Single nucleotide polymorphism (SNP) disrupting an exonic splicing enhancer (ESE). ESE sequences shown in bold above the exonic sequence.

KIAA1632 in the thymus, the central nervous system, muscle cells (heart and skeletal muscle), liver and prostate tissue and in kidney and lung tissue. Literature research showed that *KIAA1632* is also found in breast and cervical cancer. Additional information was obtained from the GNF SymAtlas (http://symatlas.gnf.org/SymAtlas/), showing an elevated expression in cells of the immune system (e.g. $CD4^+$, $CD8^+$ and natural killer cells).

Discussion

The human *KIAA1632* gene, consisting of 44 exons is located at human chromosome 18q12.3-q21.1. We analyzed the human KIAA1632 gene in this study by refinement of the genomic structure and characterization of possible phylogenetic relations to orthologous proteins in other species. To reveal the genomic structure of human *KIAA1632* we identified the complete coding sequence. *KIAA1632* genomic sequence was determined by assembling the full-length transcript (accession number NM_020964) to the genomic region (genome sequence AC090355). The putative protein sequence of *KIAA1632* showed orthologs in other species, all belonging to the phylogenetic group of bilateria. Based on the phylogenetic findings the protein originated one billion years ago. This finding suggests a functional importance for higher eukaryotic species.

In silico expression analyses revealed that human *KIA1632* mRNA is expressed in a variety of tissues: in cells of the immune system, thymus, the central nervous system, muscle cells (heart and skeletal muscle), liver and prostate tissue, kidney and lung tissue and in breast and cervical cancer.

Analysis of the amino acid sequence of the putative *KIAA1632* protein revealed no known domains. Interestingly the comparative analysis revealed a common 'signature motif' found in all *KIAA1632* homologous proteins. A database search detected this 'signature motif' in two well-known cancer antigens: tumor antigen se2-2 (Nephrocystin-6) and NY-REN-58 antigen (renal cancer antigen). Antibodies against se2-2/Nephrocystin-6 are present in sera from patients with cutaneous T-cell lymphomas (31) and NY-REN-58 is an antigen recognized by autologous antibodies in patients with renal-cell carcinoma (29).

Furthermore, we detected this 'signature motif' in the 'Rod photoreceptor cGMP-gated channel subunit α' (CNG1) protein, a protein of a new class of cancer antigens in melanoma: novel photoreceptor proteins that are responsible for visual transduction and its regulation (30).

This raises the question: is there a common 'signature motif' for a certain type of tumor antigen? What possible function in cancer have those other (mammalian) proteins which have this motif? Subsequently *KIAA1632* could also be an immunogenic cancer antigen. The function of the motif-related domain is unknown, but the relative abundance of proteins involved in cell-cell interaction or adhesion hints to a possible structural role.

In summary, this is the first report on comparative integromics on *KIAA1632* and its orthologs, revealing clues to a 'signature motif' of a cancer antigen domain.

References

1. Futreal PA, Coin L, Marshall M, *et al*: A census of human cancer genes. Nat Rev Cancer 4: 177-183, 2004.
2. Sjoblom T, Jones S, Wood LD, *et al*: The consensus coding sequences of human breast and colorectal cancers. Science 314: 268-274, 2006.
3. Pickard BS, Malloy MP, Clark L, *et al*: Candidate psychiatric illness genes identified in patients with pericentric inversions of chromosome 18. Psychiatr Genet 15: 37-44, 2005.
4. Katoh M: Paradigm shift in gene-finding method: from bench-top approach to desk-top approach (Review). Int J Mol Med 10: 677-682, 2002.
5. Katoh M and Katoh M: FGFR2 and WDR11 are neighboring oncogene and tumor suppressor gene on human chromosome 10q26. Int J Oncol 22: 1155-1159, 2003.
6. Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ: Basic local alignment search tool. J Mol Biol 215: 403-410, 1990.
7. Katoh Y and Katoh M: Comparative integromics on FAT1, FAT2, FAT3 and FAT4. Int J Mol Med 18: 523-528, 2006.
8. Katoh Y and Katoh M: Comparative integromics on VEGF family members. Int J Oncol 28: 1585-1589, 2006.
9. Katoh M and Katoh M: Identification and characterization of human FMNL1, FMNL2 and FMNL3 genes *in silico*. Int J Oncol 22: 1161-1168, 2003.
10. Katoh M and Katoh M: CLDN23 gene, frequently down-regulated in intestinal-type gastric cancer, is a novel member of CLAUDIN gene family. Int J Mol Med 11: 683-689, 2003.
11. Katoh Y and Katoh M: Comparative integromics on angiopoietin family members. Int J Mol Med 17: 1145-1149, 2006.
12. Katoh Y and Katoh M: Comparative integromics on BMP/GDF family. Int J Mol Med 17: 951-955, 2006.
13. Stamm S, Riethoven J, Le Texier V, *et al*: ASD: a bioinformatics resource on alternative splicing. Nucleic Acids Res 34: D46-D55, 2006.
14. Thanaraj TA, Stamm S, Clark F, Riethoven J, Le Texier V and Muilu J: ASD: the Alternative Splicing Database. Nucleic Acids Res 32: D64-D69, 2004.
15. Clark F and Thanaraj TA: Categorization and characterization of transcript-confirmed constitutively and alternatively spliced introns and exons from human. Hum Mol Genet 11: 451-464, 2002.
16. Dralyuk I, Brudno M, Gelfand MS, Zorn M and Dubchak I: ASDB: database of alternatively spliced genes. Nucleic Acids Res 28: 296-297, 2000.
17. Gelfand MS, Dubchak I, Dralyuk I and Zorn M: ASDB: database of alternatively spliced genes. Nucleic Acids Res 27: 301-302, 1999.
18. Bairoch A, Apweiler R, Wu CH, *et al*: The Universal Protein Resource (UniProt). Nucleic Acids Res 33: D154-D159, 2005.
19. Jonassen I, Collins JF and Higgins DG: Finding flexible patterns in unaligned protein sequences. Protein Sci 4: 1587-1595, 1995.
20. Jonassen I: Efficient discovery of conserved patterns using a pattern graph. Comput Appl Biosci 13: 509-522, 1997.
21. Kumar S, Tamura K and Nei M: MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5: 150-163, 2004.
22. Xia X and Xie Z: DAMBE: software package for data analysis in molecular biology and evolution. J Hered 92: 371-373, 2001.
23. Schmidt HA, Strimmer K, Vingron M and von Haeseler A: TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics 18: 502-504, 2002.
24. Fairbrother WG, Holste D, Burge CB and Sharp PA: Single nucleotide polymorphism-based validation of exonic splicing enhancers. PLoS Biol 2: E268, 2004.

25. Fairbrother WG, Yeh R, Sharp PA and Burge CB: Predictive identification of exonic splicing enhancers in human genes. Science 297: 1007-1013, 2002.
26. Fairbrother WG, Yeo GW, Yeh R, Goldstein P, Mawson M, Sharp PA and Burge CB: RESCUE-ESE identifies candidate exonic splicing enhancers in vertebrate exons. Nucleic Acids Res 32: W187-W190, 2004.
27. Bailey TL, Williams N, Misleh C and Li WW: MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34: W369-W373, 2006.
28. Sayer JA, Otto EA, O'Toole JF, *et al*: The centrosomal protein nephrocystin-6 is mutated in Joubert syndrome and activates transcription factor ATF4. Nat Genet 38: 674-681, 2006.
29. Scanlan MJ, Gordan JD, Williamson B, *et al*: Antigens recognized by autologous antibody in patients with renal-cell carcinoma. Int J Cancer 83: 456-464, 1999.
30. Bazhin AV, Schadendorf D, Willner N, *et al*: Photoreceptor proteins as cancer-retina antigens. Int J Cancer 2006.
31. Eichmuller S, Usener D, Dummer R, Stein A, Thiel D and Schadendorf D: Serological detection of cutaneous T-cell lymphoma-associated antigens. Proc Natl Acad Sci USA 98: 629-634, 2001.