Research

Isolation, *in silico* characterization and chromosomal localization of a group of cDNAs from ciliated epithelial cells after *in vitro* ciliogenesis

Amit K Maiti*, Mark Jorissen† and Patrice Bouvagnet*

Addresses: *Laboratoire de Génétique Moléculaire Humaine, Faculté de Pharmacie, Université Claude Bernard Lyon 1, 8 av. Rockefeller, F-Lyon cedex 08, France. †Centrum voor Menselijke Erfelijkheid, Campus Gasthuisberg, 6 Herstraat 49, B-3000 Leuven, Belgium.

Correspondence: Amit K Maiti. E-mail: Amit.Maiti@medecine.unige.ch; amit@dplanet.ch

Published: 27 June 2001

*Genome Biology* 2001, 2(7):research0026.1–0026.9

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2001/2/7/research/0026

© 2001 Maiti et al., licensee BioMed Central Ltd

(Print ISSN 1465-6906; Online ISSN 1465-6914)

Abstract

**Background:** Immotile cilia syndrome (ICS) or primary ciliary dyskinesia (PCD) is an autosomal recessive disorder in humans in which the beating of cilia and sperm flagella is impaired. Ciliated epithelial cell linings are present in many tissues. To understand ciliary assembly and motility, it is important to isolate those genes involved in the process.

**Results:** Total RNA was isolated from cultured ciliated nasal epithelial cells after *in vitro* ciliogenesis and expressed sequenced tags (ESTs) were generated. The functions and locations of 63 of these ESTs were derived by BLAST from two public databases. These ESTs are grouped into various classes. One group has high homology not only with the mitochondrial genome but also with one or more chromosomal DNAs, suggesting that very similar genes, or genes with very similar domains, are expressed from both mitochondrial and nuclear DNA. A second class comprises genes with complete homology with part of a known gene, suggesting that they are the same genes. A third group has partial homology with domains of known genes. A fourth group, comprising 33% of the ESTs characterized, has no significant homology with any gene or EST in the database.

**Conclusions:** We have shown that sufficient information about the location of ESTs could be derived electronically from the recently completed human genome sequences. This strategy of EST localization should be significantly useful for mapping and identification of new genes in the forthcoming human genome sequences with the vast number of ESTs in the dbEST database.

**Background**

Immotile cilia syndrome (ICS) or primary ciliary dyskinesia (PCD) is a human autosomal recessive disorder with a frequency of 1 in 20,000. Patients with PCD have recurrent respiratory tract infections, bronchiectasis and often male sterility. About 50% of patients have situs inversus and hence a Kartagener syndrome. These patients show abnormalities in the beating of cilia in ciliated epithelial cells and of flagella of spermatozoa. Electron microscopic ultrastructural study of cilia and spermatozoa of patients show that this disease is extremely heterogeneous [1,2]. Ciliated epithelial cell linings are present in the upper airways of the respiratory tract, sinuses, middle ear, efferent duct of testis, Fallopian tubes, brain and spinal cord. Embryonic heart contains nodal cilia...
that produce a directional movement and it has been shown in mice that failure of the movement of these nodal cilia causes breakdown of left-right asymmetry [3]. Cilia and flagella are complex structures and ciliary assembly alone requires more than 250 different proteins [4]. Upper airway epithelial cells are also important for studying cystic fibrosis and asthma, and are often cultured in vitro for drug testing for asthma and related diseases. The identification of genes expressed in these cells may be helpful in characterizing genes involved in such diseases.

Upper airway epithelial cells have not been used previously for isolation of ESTs. We cultured ciliated epithelial cells starting from a patient’s nasal biopsy, and after in vitro degeneration and regeneration of cilia, total RNA was isolated from these cells. A catalog of the function and chromosomal location of the expressed sequence tags (ESTs) generated from the RNA was deduced by BLAST searching of the public databases (GenBank, normal and HTGS). This implies that comprehensive information about gene functions and chromosomal locations of ESTs could be derived from these databases.

**Results and discussion**

We have isolated a group of ESTs from ciliated epithelial cells after in vitro ciliogenesis starting from a patient’s nasal biopsy. The probable functional significance of these ESTs and their chromosomal locations are derived from published databases. For homology searches, two databases were considered. The first was a normal database which gives the identity of the sequence with respect to the other transcribed sequences from all organisms. The second was a ‘high throughput genome sequences’ (HTGS) database, which was used to determine the genomic clones that are homologous to these transcribed sequences. According to the known position of the sequenced clone, ESTs are placed in between the two closest markers in the chromosome (see, for example Table 1). These transcribed short sequences are divided into four subgroups according to their homology with the database.

**ESTs with homology with mitochondrial DNA**

A number of the nasal epithelial ESTs show very high homology with sequences from the human mitochondrial genome (Table 1), implying that these sequences are derived from mitochondrial DNA. It is surprising, however, that most of these ESTs are not only homologous with the mitochondrial genome but also with chromosomal DNA, and that the same region and extent of homology with mitochondrial and genomic DNA is observed in many cases. Moreover, an individual EST can have very high homology with a HTGS genomic clone from more than one chromosome (see 34-18, 36-62, 5R22 and 36-100 in Table 1). ESTs 34-18 and 9694, for example, have homology with the mitochondrial *urf4* gene and with the same genomic clone in chromosome 5.

Multiple alignment of ESTs 34-18 and 9694 suggests that these are different sequences and from different regions of the genomic clone.

It is possible that families of very similar genes (or of genes with similar domains) are expressed from mitochondrial and nuclear genomic DNA or that a massive amount of domain fusion has occurred between mitochondrial and nuclear genes. Only one chimeric *urf4* cDNA (fused mitochondrial and nuclear DNA) has previously been recovered experimentally from a viral integration, over a decade ago [5]. On the basis of the high homology of each EST with both mitochondrial and nuclear DNA, we suggest that nuclear and mitochondrial domain fusion is not an isolated phenomenon, but is rather common. This remains to be rigorously investigated.

**ESTs with complete homology with known genes**

A second group of ESTs are completely homologous with known genes in the human genome (Table 2). We assume that these are either the same gene as their genomic counterpart or a gene containing the same domain. From this evidence it is interesting to note that a number of important genes whose functions are known are also expressed in ciliated epithelial cells, although the significance of this expression is unknown. Although further rigorous experiments are needed to characterize these genes in ciliated epithelium, the probable functions of some important genes are discussed below.

Cytohesin (EST 14-49) is involved in signal transduction pathways and regulates cell adhesion [6]. Expression of this gene may play an important role in the adhesion of epithelial cells during the expansion of the cell layer.

Cyclophilin C (EST 24-51). In response to endotoxin, mice deficient in cyclophilin-associated protein overproduce interleukin-12 and interferon-gamma systemically and tumor necrosis factor-alpha locally. These are proinflammatory molecules that also promote helper T-cell responses [7]. The role of this gene in ciliated epithelial cells remains elusive, however, and an important concern for further investigation.

Epidermal growth factor receptor (EGFR) kinase (EST 6034). EGFR is an important ligand-binding protein and its level is elevated in many tumors [8]. The EGFR gene is a potential oncogene and expression of EGFR kinase may be required during formation of epidermis by epithelial cells.

The *Drosophila Staufen* genes (EST 6092) are RNA-binding proteins important for RNA transport and localization in the oocyte and neurons in *Drosophila*. The motor protein dynein (dld1c) in conjunction with Staufen and Swallow acts as an adaptor for transporting *bicoid* RNA along microtubules to their minus ends at the anterior pole of the oocyte.
## Table 1

### Homology with mitochondrial DNA

| Clone names | Homology to normal database (identity), accession number | Homologous region, extent of similarity (bp)*, identity (bp/bp)† | Chromosome | Homologous region, extent of similarity (bp)*, identity (bp/bp)† | In between markers | Distance from P-tel (kb) |
|-------------|--------------------------------------------------------|---------------------------------------------------------------|-------------|---------------------------------------------------------------|-------------------|------------------------|
| 34-18, 356 bp | L00016.1 Hs urf4 gene, mitochondrial                    | 24-336, 313/313                                              | AC021965 | 5                                                             | 35-333, 271/299  | D5S2400 & SHGC-141614 | 124831          |
| 36-105, 70 bp | NC_001807.2 Hs mitochondrion                           | 15-51, 31/36                                                  | NSH        | NA                                                            | NA                | NA                     | NA               |
| 24-16, 215 bp | NC_001807 Hs mitochondrion                            | 1-196, 195/196                                               | AC021965 | 5                                                             | 1-196, 178/196  | D5S2400 & SHGC-141614 | 124831          |
| 36-10, 56 bp  | NC_001807 Hs mitochondrion                             | 31-100, 69/70                                                | AL359496  | I                                                             | 31-100, 68/70   | NMF                    | 124831          |
| 36-16, 105 bp | NC_001807 Hs mitochondrion                            | 24-82, 57/59                                                 | AC008670  | 5                                                             | 27-82, 51/56    | D5S2056 & A007G12     | 82616           |
| 36-34, 72 bp  | NC_001807 Hs mitochondrion                             | 18-55, 37/38                                                 | AC021914  | I                                                             | 21-52, 30/32    | sTG4656 & sTG46623    | 26100           |
| 36-36, 87 bp  | AF134583 Hs mitochondrial DNA-like                     | 19-67, 49/49                                                 | AL049739  | 6                                                             | 26-46, 21/21    | S7853 & sTG46623      | 39448           |
| 9694, 150 bp  | L00016.1 Hs urf4 gene, mitochondrial                  | 19-123, 114/114                                              | AC021965  | 5                                                             | 19-121, 95/103  | D5S2400 & SHGC-141614 | 124831          |
| 34-47, 103 bp | NC_001807.2 Hs mitochondrion                           | 16-82, 67/67                                                 | AC022223  | 5                                                             | 16-82, 67/67    | RH12923 & WI-18379    | 95113           |

*Extent of similarity, the number corresponds to the starting and ending base pair in the respective homologous gene with the EST. †Identity, the number corresponds to the identical base pair of EST/homologous gene. NA, not applicable; Hs, Homo sapiens; Dm, Drosophila melanogaster; Dr, Drosophila radiodurans. NSH, no significant homology (identity less than 20 bases); NMF, no matches found.

[9,10]. As dynein genes are highly expressed in ciliated epithelial cells [11] a potential interaction of dynein with the human Staufen homolog could be deduced.

Decay-accelerating factor (DAF, CD55) (EST 9661) protects host cells from the activation of autologous complement on their surfaces. It functions to disable the C3 convertases, the
Table 2
Complete homologies with known genes

| Clone names | Homology to normal database (identity), accession number | Homologous region, extent of similarity (bp)\(^1\) | Homology to human clone (HTGS) accession number | Chromosome | Homologous region, extent of similarity (bp)\(^2\), identity (bp/bp)\(^1\) | In between markers | Distance from P-cel (kb) |
|-------------|---------------------------------------------------------|-----------------------------------------------|-----------------------------------------------|------------|----------------------------------------------------------------|-----------------|------------------------|
| 14-49, 336 bp | NM_004228.2 Hs coiled/coil domain 2 (cytohesin-2)     | 22-317, 296/296                                | AC073131                                      | 19         | 110-225,116/116 223-317, 95/95                                   | D19S902 & sTsg58178 | 68477                  |
| 24-51, 234 bp | NM_000943.1 Hs peptidylprolyl isomerase C cyclophilin | 33-220, 186/186                                | AC012424,                                     | 5          | 99-220,121/122                                                  | RH101603 & RH103740 | 150395                 |
| 3s-1, 121 bp | NM_006870.2 Hs (actin depolymerizing factor)          | 20-109, 90/90                                  | AI132765,                                     | 20         | 20-109, 90/90                                                   | WI-22195 & RH123144 | 25762                  |
| 6034, 82 bp  | NM_004447.1 Hs epidermal growth factor receptor kinase | 18-66, 49/49                                   | NSH                                           | NA         | NA                                                               | NA               | NA                     |
| 6086, 115 bp | M27024 Hs heat shock protein                          | 19-115, 91/97                                  | AL133223.3                                    | 14         | 19-115, 91/97                                                   | H14a433 & D14S305 | 119103                 |
| 6092, 71 bp  | NM_004602.1 Hs staufen(STAU) (Dm RNA-binding protein) | 29-71, 41/43                                   | AC068845                                      | 19         | 29-71, 41/43                                                    | NMF              | NA                     |
| 9661, 257 bp | M31516.1 Hs decay-accelerating factor mRNA             | 17-236, 218/220                                | AL355527                                      | 1          | 17-236, 218/220                                                  | NMF              | NA                     |
| 968, 104 bp  | AF203815 Hs alpha gene sequence                       | 23-87, 63/65                                   | AP000769                                      | 11         | 23-87, 63/65                                                    | NMF              | NA                     |

\(^1\)Extent of similarity, the number corresponds to the starting and ending base pair in the respective homologous gene with the EST. \(^2\)Identity, the number corresponds to the identical base pair of EST/homologous gene. NA, not applicable; Hs, Homo sapiens; Dm, Drosophila melanogaster; Dr, Drosophila radIODURANS. NSH, no significant homology (identity less than 20 bases); NMF, no matches found.

Central amplification enzymes of the complement cascade [12]. Expression of this gene in nasal epithelial cells could be explained by the need for protection against antigen-induced complement activation.

**ESTs with partial homology to domains of known genes**

A number of ESTs (Table 3) are partially homologous (that is in part of the EST sequence) to domains of known genes and could be of interest. However, further investigation of complete cDNAs and their functions may reveal the true identity of these genes. A few ESTs which are partly homologous with the domains of important genes are discussed below.

**EST 24-17** (H3 pseudogene). A stretch of 21 base pairs (bp) of this EST has homology with the pseudogene of histone H3 but not with the normal histone H3 gene. Histone H3 is an important housekeeping protein involved in chromatin packing [13]. It is possible that another H3 RNA is transcribed, which may be different from both the normal H3 gene and the pseudogene.

**EST 36-5** (retinoic acid responder). As retinoic acid plays an important role in *in vitro* cilogenesis, expression of the retinoic acid responder domain is not unexpected in ciliated epithelial cells, where it may modulate a number of dynein heavy-chain genes during cilogenesis [14].

**EST 9010** (Attractin precursor). The protein attractin is secreted by activated T cells and has also been detected in the central nervous system [15]. It is suggested to be involved in immunity, obesity and pigmentation [16]. The mouse *mahogany* mutation is caused by a mutation in the attractin gene [17]. A portion (192 bp out of 194 bp) of EST 9010 has a very high similarity or identity with the attractin gene, suggesting that this EST is derived from a gene containing an attractin precursor domain. Expression of such a gene in ciliated epithelial cells is of unknown significance.
Table 3
Partial homologies with known genes

| Clone names | Homology with normal database (identity), accession number | Homologous region, extent of similarity (bp)*, identity (bp/bp)† | Homology with human clone (HTGS), accession number | Chromosome | Homologous region, extent of similarity (bp)*, identity (bp/bp)† | In between marker | Distance from P-tel (kb) |
|-------------|----------------------------------------------------------|---------------------------------------------------------------|---------------------------------------------------|------------|---------------------------------------------------------------|-----------------|------------------------|
| 24-17, 389 bp | AL137022 Histone (H3) pseudogene                          | 125-145, 21/21                                               | NSH                                               | NA         | NA                                                             | NA              | NA                     |
| 24-4, 323 bp | AF078904 Zeta globin gene                                  | 84-105bp, 22/22                                             | AC058816                                          | 6          | 1-135,121/122                                                  | stSN21216 and D651442 | 7798                   |
| 26-1, 138 bp | AE003672 Dm genome scaffold binding protein                | 56-78, 22/23                                                | AC022237                                          | 15         | 1-20, 20/20                                                    | NMF             | NA                     |
| 3R1-2f, 429 bp | AE001826 Dr R1 megaplasmid MP1                             | 6-32, 26/27                                                 | NSH                                               | NA         | NA                                                             | NA              | NA                     |
| 36-5, 225 bp | NM_002888.1Hs retinoic acid receptor responder             | 52-190, 136/139                                            | AC025033                                          | 3          | 51-190,137/140                                                 | stSG335337 and R95445 | 197037                 |
| 36-50, 104 bp | NM_014666.1 Hs KIAA071 gene product                        | 32-84, 48/53                                                | AC026407                                          | 5          | 32-84, 48/53                                                    | D5S1853 and RH101108 | 123950                 |
| 36-98, 195 bp | AF068299 A. thaliana gamma glutamylcysteine synthetase gene | 170-193, 24/25                                              | AC069530                                          | 3          | 164-182 19/19                                                    | NMF             | NA                     |
| 3R1-32, 182bp | AE003819 Dm genomic scaffold                               | 102-121, 20/20                                             | NSH                                               | NA         | NA                                                             | NA              | NA                     |
| 5R29, 506bp  | AE003568 Dm genomic scaffold                               | 465-491,26/27                                              | NSH                                               | NA         | NA                                                             | NA              | NA                     |
| 5R5, 83 bp   | AE003650 Dm genomic scaffold                               | 22-42, 21/21                                                | AC027364                                          | 6          | 19-66, 48/48                                                   | RH112849 & stSG47852 | 6789                   |
| 9010, 440 bp | AF218906.1 Hs attractin precursor (ATRN)                   | 215-408,192/194                                            | AC015847                                          | 17         | 215-408,192/194                                                | NMF             | NA                     |
| 9014, 396 bp | AF095856 Hs asthmatic clone 4 mRNA                         | 373-394, 22/22                                              | AL133245.2                                        | 2          | 109-358,247/250                                                | stSG60109 and RH120618 | 36134                  |
| 905, 485 bp  | X73004.1Hs EWS gene                                        | 173-415,237/239                                            | AF121897                                          | 21         | 176-415,238/240                                                | NMF             | NA                     |
| 906, 130 bp  | AF144028.1Hs MDM2 gene                                     | 19-72, 54/54                                                | AC019009                                          | 14         | 19-67, 49/49                                                   | NMF             | NA                     |
| 907, 391 bp  | AB026436 Hs for dual specificity phosphatase MKPS          | 12-64, 52/53                                                | AL049696.9                                        | 6          | 92-371,279/280                                                 | D651762 and D651856 | 88638                  |
| 9640, 271 bp | AB024935 Mus musculus SId3177 mRNA                         | 106-200, 89/96                                              | AC073620                                          | 12         | 13-249, 237/237                                                | RH44840 and RH83752 | 9861                   |
| 9646, 249 bp | NM_014928.1 Hs KIAA1046 protein                            | 17-124, 107/108                                             | AC006207                                          | 12         | 17-235, 216/219                                                | B568G1/T7 and D1252049 | 1911                   |
| 9667, 262 bp | AF119664 Hs Transcriptional regulator protein HCN6G mRNA   | 130-216, 86/87                                              | AC019214                                          | 17         | 1-179, 177/179                                                 | D175609 and D1751769 | 87186                  |
| c2s-3, 348 bp | AF207550 Hs protein translocase                            | 49-86, 34/38                                                | AC019099                                          | Y          | 4-286, 252/254                                                 | DYS215 and DYS197 | 21795                  |

*Extent of similarity, the starting and ending base pair in the respective homologous gene with the EST. †Identity, the number corresponds to the identical base pair of EST/homologous gene, number of base pair identical in EST with homologous gene. NA, not applicable. Hs, Homo sapiens; Dm, Drosophila melanogaster; Dr, Drosophila radiodurans. NSH, no significant homology (identity less than 20 bases); NMF, no matches found.
EST 906 (MDM2). MDM2 is an oncogene and the MDM2 oncoprotein binds to the p53 protein, inhibiting p53’s function as a transcription factor and inducing its degradation. An MDM2-p53 autoregulatory feedback loop regulates the function of the p53 tumor suppressor gene [18]. The significance of the expression of MDM2 or of a gene carrying an MDM2 domain in ciliated epithelial cells remains to be investigated.

EST 907 (Mkp5 dual-specificity phosphatase). Mkp5 is a member of the mitogen-activated kinase (MAP kinase) family (10 genes in total) and has an important role in apoptosis, tumor progression and immune responses [19]. MAP kinases contain a docking motif that increases the efficiency of the reaction [20]. Part of EST 907 (52 bp out of 391 bp) has high similarity with the docking motif, possibly implying that another member of the MAP family could exist.

EST 9667 (transcription regulator protein; HCNGP). This transcriptional regulator has been isolated from adrenal gland (C. Jiang, J Shi, C Huang, S Ren, Y Li, J Zhou, Y Yu, S Xu, Y Wang, G Fu, et al., unpublished data; GenBank accession number AF119664). A part of EST 9667 (86 bp out of 87 bp) is highly similar to the HCNGP; expression of this gene in ciliated epithelial cells is of unknown significance.

EST C2s-3 (translocase). Protein translocation across the cytoplasmic membrane has not been studied extensively in mammalian cells. It is also not known how many genes are involved in this process [21]. It is interesting that this EST shows homology (34 bp out of 38 bp) with a translocase gene and may carry out a translocase-like function.

EST 9640 (Sid3177). Sid3177 is a unique gene isolated from the mouse and assumed to be a part of the inactive progesterone receptor complex (N Seki, A Hattori, A Hayashi, S Kozuma, M Muramatsu, T Saito, unpublished data; GenBank accession number AB024935). The high homology of EST 9640 with a domain of Sid3177 suggests that a similar gene(s) is present in humans and is expressed in ciliated epithelial cells.

ESTs 3R1-32, 5r29, 5R5 (genomic scaffold-binding protein). Parts of these ESTs have high homology with the Drosophila scaffold-binding protein [22]. Scaffold-binding proteins are important in the replication and segregation of chromosomes. However, human counterparts of the complete genes have yet to be isolated.

EST 9014 (asthmatic clone). Part of this EST has homology with a cDNA isolated from asthma patients and that appears to be expressed differentially in asthmatics (IC Kilty, PJ Vickers, unpublished data; GenBank accession number AF095856). Expression of such a gene in the upper airway epithelium is of important in the context of identifying genes responsible for asthma.

ESTs with no significant homology to sequences in the database
Twenty ESTs (33%) have no significant homology (Table 4) to sequences in normal databases. The identities of these genes are not known even after 2 million sequences have been accumulated in the dbEST database. BLAST searches against the HTGS database, however, reveals that most of them are highly homologous with known genomic clones. These ESTs have been mapped electronically and their chromosomal locations derived. In recent years, there has been an exponential rise in the number of sequences available in the public databases. Despite this, a high percentage of partial sequences of cDNAs (ESTs) submitted to the databases remain unrecognized (anonymous ESTs). This lack of similarity could be explained [23] in several ways. One explanation is that a different part of the transcript is present in GenBank; second, the transcript represents a novel gene not yet isolated; third, there is alternative splicing of the same gene in different species; fourth, inaccurate sequence data; and/or fifth, the sequence of the transcript has diverged to an extent that it is not recognized as an ortholog.

Conclusions
We have isolated a group of cDNAs that are expressed in ciliated epithelial cells in the upper airway of the human respiratory tract. These short cDNAs may be extremely helpful for isolating and characterizing the complete genes and for studying their expression pattern in the human body. We also noted that a number of ESTs are highly homologous to genes that are involved in cancers and immune reaction pathways. Expression of these genes in ciliated epithelial cells in the upper respiratory tract is of unknown significance. In addition, mapping these genes may be helpful for retrieving and characterizing complete genes. Subsequently, it may help in cloning those disease genes by the positional candidate gene approach. A number of the ESTs can be mapped electronically from the human genome sequence (HTGS database) and their probable function could be derived from the normal database. This shows that a large number of ESTs in the dbEST database could be mapped electronically by BLAST and a comprehensive EST map could be generated that may be helpful for characterizing a large number of genes in the human genome.

Materials and methods
Human epithelial cell culture
Human nasal epithelial cells were enzymatically dissociated from biopsies using a protease type XIV (Pronase) digestion overnight at 4°C. Pronase was inactivated by adding fetal calf serum (FCS) or NU-Serum (10% final concentration) and the cell suspension was washed three times in culture medium (Ham’s F12- DME 1/1 with cholera toxin (10 ng/ml), streptomycin (50 μg/ml), penicillin (50 IU/ml) and 2% Ultroser G). The cell suspension was preplated on
Table 4

No significant homology in the normal database

| Clone names | Homology to normal database (identity), accession number | Homologous region, extent of similarity (bp)*, identity (bp/bp) | Homology to human clone (HTGS), accession number | Chromosome | Homologous region, extent of similarity (bp)*, identity (bp/bp) | In between markers | Distance from P-tel (kb) |
|-------------|----------------------------------------------------------|---------------------------------------------------------------|-------------------------------------------------|-------------|---------------------------------------------------------------|-------------------|--------------------------|
| 123-12, 324 bp | NSH NA | NSH | NA | NA | NA | NA | NA |
| 24-16n, 401 bp | NSH NA | AC015927 | 9 | 4-390, 378/395 | NMF | NA | NA |
| 24-37, 170 bp | NSH NA | NA | NA | NA | NA | NA | NA |
| 26-16, 226 bp | NSH NA | NA | NA | NA | NA | NA | NA |
| 5r9, 310 bp | NSH NA | AC009554 | 15 | 22-114, 91-93 | WI-14756 and D15S553 | 65614 |
| 26-6, 280 bp | NSH NA | AC009086 | 16 | 30-146, 114/117 | NMF | NA | NA |
| 36-39, 72 bp | NSH NA | AC008670 | 5 | 30-77, 45/49 | D5S2056 and 1007G12 | 82630 |
| 34-2, 175 bp | NSH NA | NA | NA | NA | NA | NA | NA |
| 26-6bis, 188 bp | NSH NA | AC009086 | 16 | 30-146, 114/117 | NMF | NA | NA |
| 26-16, 226 bp | NSH NA | AC009086 | 16 | 30-146, 114/117 | NMF | NA | NA |
| 24-37, 170 bp | NSH NA | AC009086 | 16 | 30-146, 114/117 | NMF | NA | NA |
| 5r9, 310 bp | NSH NA | AC009554 | 15 | 22-114, 91-93 | WI-14756 and D15S553 | 65614 |

*Extent of similarity, the number corresponds to the starting and ending base pair in the respective homologous gene with the EST. †Identity, the number corresponds to the identical base pair of EST/homologous gene. NA, not applicable; Hs, Homo sapiens; Dm, Drosophila melanogaster; Dr, Drosophila radiodurans. NSH, no significant homology (identity less than 20 bases); NMF, no matches found.

plastic for 1 h at 37°C to remove most of the contaminating fibroblasts [24,25]. Cells were plated in T75 tissue culture flasks on 0.2% collagen gel for monolayer culture and kept at 37°C at 5% CO2 atmosphere in a biological oxygen demand (BOD) incubator. Culture medium was changed three times a week. After three weeks of exponential growth the cultures reach confluence.

Cells were then released from the collagen gel using 200 IU/ml collagenase type IV. Cell clusters, aggregates and cell sheets were washed three times in culture medium to eliminate collagenase and then placed in culture medium at 37°C on a gyratory shaker at 80 rpm to avoid attachment of the cells to the culture flask. During the first week, the medium (the same as used earlier) was changed every day. On the second day, the 2% Ultroser G was replaced by 10% NU-serum [24,25]. After 1 week stable aggregates, spheroids and vesicles were formed and showed no tendency to adhere to the culture flask; the culture was kept stationary for another few weeks. Generally, cilia appear in 2 weeks and cells were used to isolate total cellular RNA after the third or fourth week.

**Primer design, RT-PCR amplification, cloning and sequencing of clones**

We used degenerate primers for the reverse transcription and PCR amplification (RT-PCR). Primers are initially designed [11] to clone the dynein heavy-chain genes from ciliated epithelial cells and are taken from the P-loop region (ATP hydrolysis region) of dynein heavy-chain genes. Reverse transcription was done with primers 4, 6 and 3R. PCR amplifications were done in round-robin fashion with
primers 1 and 4, 2 and 4, 3 and 4, 1 and 6, 2 and 6, 3 and 6, 1 and 3R, 2 and 3R, 3 and 3R, and so on. Primers [11] used in these studies were 1, 5'-TAY GGN TTY GAR TAY YTN GG-3'; 2, 5'-G47/G45/G4E/G45/G42/G41/G4E/G4B/G20/G4D/G41/G50/G92/G20/G61/G6E/G64/G20/G74/G68/G65/G20/G63/G6C/G6F/G6E/G65/G20/G63/G61/G72/G72/G79/G69/G6E/G67/G20/G66/G6C/G61/G6E/G6B/G69/G6E/G67/G20/G6D/G61/G72/G6B/G65/G72/G73

Total RNA was isolated from the cultured cells by the method described in [26,27]. RT-PCR was carried out with GeneAmp RT-PCR Kit (Perkin Elmer). Each sample of RNA was routinely treated with DNase I for 6 h at 37°C to remove any genomic DNA contamination. Two micrograms of RNA were reverse transcribed by the downstream primer in 20 μl at 42°C and PCR amplified with the addition of upstream primer in a 100 μl volume. In all cases, PCR conditions were: for denaturation, 94°C, 4 min; for amplification, 94°C for 1 min; 50°C for 1 min; 72°C for 1 min for 40 cycles; and for elongation, 72°C for 10 min. RNA without reverse transcriptase and water without RNA (plus reverse transcriptase) did not yield any product in any of the PCR reactions.

PCR products were cloned in PCR2.1 vector of the TA-cloning Kit (Invitrogen) and were subjected to blue/white selection. White colonies were checked by PCR for the presence of insert with the vector-specific primers (M13 forward and reverse). Approximately 400 clones were sequenced on an ABI377 Automated Fluorescence Sequencer (Perkin Elmer). Sequences were screened with BLAST for the identity of these clones. Along with the cloning of nine dynein heavy-chain genes [11], a number of non-dynein cDNAs were recovered which were studied in detail. From 400 clones, 63 were selected as unique by the following procedures. The dynein heavy-chain genes (82 clones) were ignored; only one sequence was selected when two or more clones containing similar sequences were obtained; very small sized clones (below 70 bases) were ignored. Only in one case was a 56 bp clone (5R16) selected, as this sequence was not obtained repeatedly.

Electronic mapping of ESTs and derivation of their probable function from database searching

ESTs were BLASTed against the normal database (in NCBI BLAST page normal database (GenBank) designated as ‘nr’) and highest similarities with the known genes were taken into account. In cases of homologies with more than one gene, only the gene with the highest homology (number of base pairs, highest similarity and identity) was taken as the homologous gene and functional characterization has been done on the basis of the function of the known gene.

EST sequences were also BLASTed against the HTGS databases and the accession number and chromosome number of the highest-similarity clone were noted. Each clone was searched by accession number in Locus Link [28] in ‘GENEBANK MAP’ and the clone carrying flanking markers in a particular chromosome was assigned in ‘STS MAP’. The distance from P-tel of the chromosome was taken as the map position of the EST.

Accession numbers

All ESTs are deposited in databases under the dbEST accession numbers 8451921 to 8451980 and 8452140 to 8452142 and the Genbank accession numbers BG673720 to BG673779 and BG687691 to BG687693.

Acknowledgements

We thank Tom Willems and Sandra Deckker for helping with the cell culture. A.K.M. is supported by a postdoctoral fellowship by a grant from the European Union (BIOMED2). This work was supported by a grant from the European Union (Biomed2, BMH4 CT96-0302).

References

1. Afzelius BA: A human syndrome caused by immotile cilia. Science 1976, 193:317-319.
2. Afzelius BA, Mossberg B: Immotile cilia syndrome including Kartagener syndrome. In The Metabolic and Molecular Basis of Inherited Disease, Vol III. 7th edn. Edited by Scriver CR, Beaudet AL, Sly WS, Valle D. New York: McGraw-Hill; 1995: 3943-3954.
3. Wagner MK, Yost HJ: Left-right development: the roles of nodal cilia. Curr Biol 2000, 10:R149-R151.
4. Dutcher KS: Flagellar assembly in two hundred and fifty easy to follow steps. Trends Genet 1995, 11:398-404.
5. Koch I, Hofschneider PH, Koshy R: Expression of a hepatitis B virus transcript containing fused mitochondrial-like domains in human hepatoma cells. Virusolgy 1989, 170:591-594.
6. Geiger C, Nagel W, Boehm T, van Kooijk Y, Fidhor CG, Kremmer E, Hogg N, Zeitlin I, Liers C, Weber K, Kolanus W: Cytoskeleton regulates beta-2 integrin-mediated adhesion through both ARF-GEF function and interaction with LFA-1. EMBO J 2000, 19:2525-2536.
7. Trahey M, Weissman IL: Cyclophilin C-associated protein: a normal secreted glycoprotein that down-modulates endotoxin and proinflammatory responses in vivo. Proc Natl Acad Sci USA, 1999 96:3006-3011.
8. Sibilla M, Fleischmann A, Behrens A, Stirling L, Carroll J, Watt FM, Schlessinger J, Wagner EF: The EGF receptor provides an essential survival signal for SOS-dependent skin tumor development. Cell 2000, 102:211-220.
9. Roegiers F, Jan YN: Staufen: a common component of mRNA transport in oocytes and neurons? Trends Cell Biol 2000, 10:220-224.
10. Schnorrer F, Bohmann K, Nusslein-Volhard C: The molecular motor dynein is involved in targeting swallowing and bicoid RNA to the anterior pole of Drosophila oocytes. Nat Cell Biol 2000, 2:185-190.
11. Maiti AK, Jorissen M, Volz A, Mattei MG, Ziegler A, Bouvagnet P: Identification, tissue specific expression and chromosomal localization of dynein heavy chain genes. Eur J Hum Genet 2000, 3006-3011.
12. Chen S, Caragine T, Cheung NK, Tomlinson S: CD59 expressed on a tumor cell surface modulates decay-accelerating factor expression and enhances tumor growth in a rat model of human neuroblastoma. Cancer Res 2000, 60:3013-3018.
13. Rea S, Eisenhaber F, O’Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CF, Allis CD, Jennewein T: Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 2000, 406:593-599.
14. Andrews KL, Nettesheim, P, Asai DJ, Ostrowsky LE: Identification of seven rat axonemal dynein heavy chain genes: Expression during ciliated cell differentiation. Mol Biol Cell 1996, 7:71-79.
15. Duke-Cohan JS, Gu J, McLaughlin DF, Xu Y, Freeman GJ, Schlossman SF: Attractin (DPPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions. Proc Natl Acad Sci USA 1998, 95:11336-11341.
16. Lu Xy, Gunn TM, Shieh Kr, Barsh GS, Akil H, Watson SJ: Distribution of Mahogany/Attractin mRNA in the rat central nervous system. FEBS Lett 1999, 462:101-107.
17. Jackson IJ: The mahogany mouse mutation: further links between pigmentation, obesity and the immune system. Trends Genet 1999, 15:429-431.
18. Zhang, Wang H: MDM2 oncogene as a novel target for human cancer therapy. Curr Pharm Des 2000, 6:393-416.
19. Trotta R, Fettucciai K, Azzoni L, Abebe B, Puorro KA, Eisenlohr LC, Perussia B: Differential role of p38 and c-Jun N-terminal kinase 1 mitogen-activated protein kinases in NK cell cytotoxicity. J Immunol 2000, 165:1782-1789.
20. Tanoue T, Adachi M, Moriguchi T, Nishida E: A conserved docking motif in MAP kinases common to substrates, activators and regulators. Nat Cell Biol 2000, 2:110-116.
21. Manting EH, Driessen AJ: Escherichia coli translocase: the unravelling of a molecular machine. Mol Microbiol 2000, 37:226-238.
22. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, et al.: The genome sequence of Drosophila melanogaster. Science 2000, 287:2185-2195.
23. Cirera S, Winter M, Fredholm M: Why do we still find ESTs. Mammalian Genome 2000, 11:689-693.
24. Jorissen M, der Schueren V, den Berghe V, Cassiman JJ: The preservation and regeneration of cilia on human nasal epithelial cells cultured in vitro. Arch Otorhinolaryngol 1989, 246:308-314.
25. Jorissen M, Bessems A: Normal ciliary beat frequency after ciliogenesis in nasal epithelial cells cultured sequentially as monolayer and in suspension. Acta Otolaryngol 1995, 115:66-70.
26. Chomczynsky P, Sacchi N: Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987, 162:156-159.
27. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K: Current Protocols in Molecular Biology. New York: John Wiley and Sons; 1987.
28. Locus Link [http://www.ncbi.nlm.nih.gov/LocusLink/]