

## Whole Exome Sequencing

### Gene package Idiopathic Pulmonary Fibrosis, version 1.1, 22-11-2017



#### Technical information

DNA was enriched using Agilent SureSelect Clinical Research Exome V2 capture and paired-end sequenced on the Illumina platform (outsourced). The aim is to obtain 8.1 Giga base pairs per exome with a mapped fraction of 0.99. The average coverage of the exome is ~50x. Duplicate reads are excluded. Data are demultiplexed with bcl2fastq Conversion Software from Illumina. Reads are mapped to the genome using the BWA-MEM algorithm (reference: <http://bio-bwa.sourceforge.net/>). Variant detection is performed by the Genome Analysis Toolkit HaplotypeCaller (reference: <http://www.broadinstitute.org/gatk/>). The detected variants are filtered and annotated with Cartagenia software and classified with Alamut Visual. It is not excluded that pathogenic mutations are being missed using this technology. At this moment, there is not enough information about the sensitivity of this technique with respect to the detection of deletions and duplications of more than 5 nucleotides and of somatic mosaic mutations (all types of sequence changes).



Dept. Clinical Genetics

HGNC approved gene symbol	Phenotype description including OMIM phenotype ID(s)	OMIM gene ID	Transcript	median depth	% covered >10x	% covered >20x	% covered >30x
ABCA3	Surfactant metabolism dysfunction, pulmonary, 3, 610921	601615	NM_001089.2	70	100	98	94
AP3B1	Hermansky-Pudlak syndrome 2, 608233	603401	NM_003664.4	50	100	95	71
COPA	{Autoimmune interstitial lung, joint, and kidney disease}, 616414	601924	NM_001098398.1	53	100	100	91
DKC1	Dyskeratosis congenita, 305000	300126	NM_001363.4	45	100	96	83
FAM111B	Poikiloderma, hereditary fibrosing, with tendon contractures, myopathy, and pulmonary fibrosis, 615704	615584	NM_198947.3	52	100	100	97
GFRA1	No OMIM phenotype	601496	NM_005264.4	55	100	98	85
MUC5B	{Pulmonary fibrosis, idiopathic, susceptibility to}, 178500	600770	NM_002458.2	93	100	99	98
NKX2-1	Chorea, hereditary benign, 118700 Choreoathetosis, hypothyroidism, and neonatal respiratory distress, 610978 {Thyroid cancer, monomedullary, 1}, 188550	600635	NM_003317.3	54	100	100	96
PARN	Dyskeratosis congenita 6, 616353 Pulmonary fibrosis and/or bone marrow failure, telomere-related, 4, 616371	604212	NM_002582.3	43	100	96	77
SFTPA2	Pulmonary fibrosis, idiopathic, 178500	178642	NM_001098668.2	127	100	100	100
SFTPC	Surfactant metabolism dysfunction, pulmonary, 2, 610913	178620	NM_003018.3	62	100	100	100
TERC	{Aplastic anemia}, 614743 Dyskeratosis congenita 1, 127550 {Pulmonary fibrosis, idiopathic, susceptibility to}, 614743	602322	NR_001566.1	No coverage data			
TERT	{Dyskeratosis congenita 2}, 613989 {Dyskeratosis congenita 4}, 613989 {Leukemia, acute myeloid}, 601626 {Melanoma, cutaneous malignant, 9}, 615134 {Pulmonary fibrosis and/or bone marrow failure, telomere-related, 1}, 614742	187270	NM_198253.2	85	100	100	94

HGNC approved gene symbol	Phenotype description including OMIM phenotype ID(s)	OMIM gene ID	Transcript	median depth	% covered >10x	% covered >20x	% covered >30x
TINF2	Dyskeratosis congenita 3, 613990 Revesz syndrome, 268130	604319	NM_001099274.1	94	100	100	100
TMEM173	STING-associated vasculopathy, infantile-onset, 615934	612374	NM_198282.2	79	100	100	100

- Gene symbols according HGNC
- OMIM release used: 2-6-2017
- "No OMIM phenotypes" indicates a gene without a current OMIM association
- OMIM phenotypes between "[ ]", indicate "nondiseases," mainly genetic variations that lead to apparently abnormal laboratory test values
- OMIM phenotypes between "{}", indicate risk factors
- OMIM phenotypes with a question mark, "?", before the disease name indicates an unconfirmed or possibly spurious mapping
- The statistics above are based on a set of 96 samples
- Median depth is the median of the mean sequence depth over the protein coding exons ( $\pm 10$ bp flanking introns) of the longest transcript
- % Covered 10x describes the percentage of a gene's coding sequence ( $\pm 10$ bp flanking introns) that is covered at least 10x
- % Covered 20x describes the percentage of a gene's coding sequence ( $\pm 10$ bp flanking introns) that is covered at least 20x