Whole Exome Sequencing
Gene package Pulmonary Arterial Hypertension (PAH)
version 1, 31-1-2020

Erasmus MC
University Medical Center Rotterdam

Dept. Clinical Genetics

Technical information
DNA was enriched using Agilent SureSelect DNA + SureSelect OneSeq 300kb CNV Backbone + Human All Exon V7 capture and paired-end sequenced on the Illumina platform (outsourced). The aim is to obtain 10 Giga base pairs per exome with a mapped fraction of 0.99. The average coverage of the exome is ~50x. Duplicateand non-unique 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).

HGNC approved gene symbol	OMIM gene ID (active link to omim.org)	median depth	% covered >10x	% covered >20x	% covered >30x
ABCC8	600509	105	100	100	100
ACVRL1	601284	102	100	100	95
AQP1	107776	112	100	100	98
ATP13A3	610232	48	100	97	87
BMP10	608748	98	100	100	100
BMPR1B	603248	72	100	100	100
BMPR2	600799	65	100	100	98
CAV1	601047	95	100	100	100
CBLN2	600433	67	100	100	100
CPS1	608307	61	100	100	96
EIF2AK4	609280	73	100	99	95
ENG	131195	98	100	100	99
GDF2	605120	122	100	100	100
GGCX	137167	82	100	100	100
KCNA5	176267	147	100	100	99
KCNK3	603220	185	100	100	99
KLF2	602016	90	91	84	77
KLK1	147910	135	100	100	100
MMACHC	609831	133	100	100	100
NOTCH3	600276	111	99	95	91
SARS2	612804	86	100	100	99

HGNC approved gene symbol	OMIM gene ID (active link to omim.org)	median depth	% covered >10x	% covered >20x	% covered >30x
SMAD1	601595	87	100	100	99
SMAD4	600993	79	100	100	98
SMAD5	603110	77	100	100	99
SMAD9	603295	85	100	100	99
SOX17	610928	157	100	100	100
TBX4	601719	145	100	100	100
TOPBP1	607760	57	100	97	89

- Gene symbols according HGCN
- OMIM release used: 8-9-2019

- The statistics above are based on a set of 100 samples
- Median depth is the median of the mean sequence depth over the protein coding exons (±10bp flanking introns) of the longest transcript
- % Covered 10x , 20x and 30 x describes the percentage of a gene's coding sequence (±10bp flanking introns) that is covered at least 10x, 20x or 30x