Whole Exome Sequencing
Gene package Very early onset inflammatory bowel disease / congenital diarrhea, version 4, 30-9-2020

Technical information
DNA was enriched using Agilent SureSelect DNA + 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. Duplicate and 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 Alissa Interpret 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).

Erasmus MC
University Medical Center Rotterdam

Dept. Clinical Genetics

HGNC approved gene symbol	OMIM gene ID (active link to omim.org)	median depth	% covered >10x	% covered >20x	% covered >30x
ACTG2	102545	169	100	100	100
ADAM17	603639	98	99	94	88
AIRE	607358	248	100	100	100
ALPI	171740	390	100	100	100
ANKZF1	617541	126	100	100	95
ARPC1B	604223	266	100	100	100
ARX	300382	100	92	89	84
BACH2	605394	212	100	100	100
CARD8	609051	107	100	100	99
CARMIL2	610859	217	100	98	96
CD19	107265	226	100	100	99
CD40	109535	158	100	100	100
CD40LG	300386	61	100	92	77
CD55	125240	79	100	95	90
CLMP	611693	155	100	100	100
COG6	606977	66	94	83	74
CR2	120650	122	99	99	96
CTLA4	123890	153	100	100	100
CYBA	608508	164	100	98	96
CYBB	300481	65	100	98	91
CYBC1	No ID	No coverage data			
DEF6	610094	209	100	100	99
DGAT1	604900	341	99	96	94
DOCK2	603122	118	100	100	100
DOCK8	611432	118	100	99	96
DUOX2	606759	183	100	100	100

HGNC approved gene symbol	OMIM gene ID (active link to omim.org)	median depth	% covered >10x	% covered >20x	% covered >30x
EGFR	131550	156	100	100	99
EPCAM	185535	69	100	94	83
FOXP3	300292	157	100	100	100
GPIHBP1	612757	208	100	100	100
GUCY2C	601330	91	100	97	92
HPS1	604982	250	100	100	100
HPS4	606682	365	100	99	98
HPS6	607522	246	100	100	100
ICOS	604558	70	100	95	87
IKBKG	300248	35	36	26	26
IL10	124092	156	100	100	100
IL10RA	146933	181	100	100	100
IL10RB	123889	86	100	98	96
IL21	605384	53	93	80	68
IL21R	605383	271	100	100	100
IL2RA	147730	139	100	100	100
IL2RB	146710	229	100	100	100
ITCH	606409	61	96	90	80
ITGB2	600065	241	100	100	100
LCT	603202	230	100	100	100
LRBA	606453	90	99	96	91
MALT1	604860	75	97	88	78
MEFV	608107	158	100	100	100
MPI	154550	228	100	100	100
MVK	251170	207	100	100	100
MYO5B	606540	165	100	100	99
NCF1	608512	62	63	58	56
NCF2	608515	213	100	100	100
NCF4	601488	170	100	100	100
NEUROG3	604882	384	100	100	100
NFKBIA	164008	168	100	100	100
NLRC4	606831	127	100	97	94
NOD2	605956	268	100	100	100
OTULIN	615712	81	98	91	78
PCSK1	162150	114	100	100	98
PLA2G4A	600522	69	98	93	83
PLCG2	600220	146	100	100	98
PLVAP	607647	255	100	100	100
PNLIP	246600	96	100	100	98
POLA1	312040	47	94	84	70
RAC2	602049	233	100	100	100
RFX6	612659	109	98	95	91
RIPK1	603453	112	100	96	95
SAR1B	607690	66	99	88	88

HGNC approved gene symbol	OMIM gene ID (active link to omim.org)	median depth	% covered >10x	% covered >20x	% covered >30x
SH2D1A	300490	81	100	95	85
SI	609845	65	94	87	79
SKIV2L	600478	190	100	100	100
SLC10A2	601295	106	100	100	98
SLC26A3	126650	78	100	95	88
SLC2A2	138160	76	100	98	94
SLC37A4	602671	193	100	100	100
SLC39A4	607059	284	100	100	100
SLC5A1	182380	136	100	100	98
SLC7A7	603593	146	100	100	100
SLC9A3	182307	252	100	100	100
SPINT2	605124	169	100	100	100
STAT1	600555	91	97	95	90
STAT3	102582	135	100	100	99
STAT5B	604260	157	100	100	100
STX3	600876	99	100	100	99
STXBP2	601717	220	100	100	100
TCN2	613441	190	100	100	100
TGFB1	190180	274	100	100	100
TMPRSS15	606635	61	88	82	77
TNFAIP3	191163	169	100	99	97
TNFRSF13B	604907	241	100	100	100
TNFRSF13C	606269	135	100	98	90
TOM1	604700	215	100	100	100
TRIM22	606559	119	100	100	99
TTC37	614589	65	99	96	91
TTC7A	609332	192	100	100	100
TYMP	131222	204	100	100	100
WAS	300392	113	100	100	100
WIPF1	602357	186	100	100	99
WNT2B	601968	191	100	100	100
XIAP	300079	63	94	89	81
ZAP70	176947	312	100	100	100

- 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