Analyse mit Exomiser:

1. yml Datei erstellen mit Pfad zur ped, vcf, proband sample name und phänotypen

```
## Exomiser Analysis Template.
# These are all the possible options for running exomiser. Use this as a template for
# your own set-up.
---
analysis:
  ped: /PATH/TO/PED
  proband: PROAND_SAMPLE_NAME
  vcf: /PATH/TO/VCF
    # AUTOSOMAL_DOMINANT, AUTOSOMAL_RECESSIVE, X_RECESSIVE or UNDEFINED
    inheritanceModes: {
            AUTOSOMAL_DOMINANT: 0.1,
            AUTOSOMAL_RECESSIVE_HOM_ALT: 3.0,
            AUTOSOMAL_RECESSIVE_COMP_HET: 5.0,
            X_DOMINANT: 0.1,
            X_RECESSIVE_HOM_ALT: 3.0,
            X_RECESSIVE_COMP_HET: 5.0,
    }
    #FULL, SPARSE or PASS_ONLY
    analysisMode: PASS_ONLY
    #hpoIds: ['HP:0100280']
         hpoIds: []
    #Possible frequencySources:
    #Thousand Genomes project http://www.1000genomes.org/
    #   THOUSAND_GENOMES,
    #ESP project http://evs.gs.washington.edu/EVS/
    #   ESP_AFRICAN_AMERICAN, ESP_EUROPEAN_AMERICAN, ESP_ALL,
    #ExAC project http://exac.broadinstitute.org/about
    #   EXAC_AFRICAN_INC_AFRICAN_AMERICAN, EXAC_AMERICAN,
    #   EXAC_SOUTH_ASIAN, EXAC_EAST_ASIAN,
    #   EXAC_FINNISH, EXAC_NON_FINNISH_EUROPEAN,
    #   EXAC_OTHER
    frequencySources: [
        THOUSAND_GENOMES,
        TOPMED,
        UK10K,
        ESP_AFRICAN_AMERICAN, ESP_EUROPEAN_AMERICAN, ESP_ALL,
        EXAC_AFRICAN_INC_AFRICAN_AMERICAN, EXAC_AMERICAN,
        EXAC_SOUTH_ASIAN, EXAC_EAST_ASIAN,
        EXAC_FINNISH, EXAC_NON_FINNISH_EUROPEAN,
        EXAC_OTHER,
        GNOMAD_E_AFR,
        GNOMAD_E_AMR,
#        GNOMAD_E_ASJ,
        GNOMAD_E_EAS,
        GNOMAD_E_FIN,
        GNOMAD_E_NFE,
        GNOMAD_E_OTH,
        GNOMAD_E_SAS,
        GNOMAD_G_AFR,
        GNOMAD_G_AMR,
#        GNOMAD_G_ASJ,
        GNOMAD_G_EAS,
        GNOMAD_G_FIN,
        GNOMAD_G_NFE,
        GNOMAD_G_OTH,
        GNOMAD_G_SAS
    ]
    #Possible pathogenicitySources: POLYPHEN, MUTATION_TASTER, SIFT, CADD, REMM
    #*WARNING* if you enable CADD, ensure that you have downloaded and installed the CADD tabix files
    #and updated their location in the application.properties. Exomiser will not run without this.
    pathogenicitySources: [POLYPHEN, MUTATION_TASTER, SIFT]
    #this is the standard exomiser order.
    #all steps are optional
    steps: [
        #intervalFilter: {interval: 'chr10:123256200-123256300'},
        #genePanelFilter: {geneSymbols: []},
        #failedVariantFilter: {},
        qualityFilter: {minQuality: 20.0},
        variantEffectFilter: {remove: [UPSTREAM_GENE_VARIANT,
            INTERGENIC_VARIANT,
            NON_CODING_TRANSCRIPT_INTRON_VARIANT,
            SYNONYMOUS_VARIANT,
            DOWNSTREAM_GENE_VARIANT]},
        #knownVariantFilter: {}, #removes variants represented in the database
        frequencyFilter: {maxFrequency: 5.0},
        pathogenicityFilter: {keepNonPathogenic: true},
        #inheritanceFilter and omimPrioritiser should always run AFTER all other filters have completed
        #they will analyse genes according to the specified modeOfInheritance above- UNDEFINED will not be analysed.
        inheritanceFilter: {},
        #omimPrioritiser isn't mandatory.
        omimPrioritiser: {},
        #priorityScoreFilter: {minPriorityScore: 0.4},
        #Other prioritisers: Only combine omimPrioritiser with one of these.
        #Don't include any if you only want to filter the variants.
        hiPhivePrioritiser: {},
        # or run hiPhive in benchmarking mode:
```

```
        #hiPhivePrioritiser: {runParams: 'mouse'},
        #phivePrioritiser: {}
        #phenixPrioritiser: {}
        #exomeWalkerPrioritiser: {seedGeneIds: [11111, 22222, 33333]}
    ]
outputOptions:
    outputPassVariantsOnly: false
    #numGenes options: 0 = all or specify a limit e.g. 500 for the first 500 results
    numGenes: 0
    #outputPrefix options: specify the path/filename without an extension and this will be added
    # according to the outputFormats option. If unspecified this will default to the following:
    # {exomiserDir}/results/input-vcf-name-exomiser-results.html
    # alternatively, specify a fully qualifed path only. e.g. /users/jules/exomes/analysis
    outputPrefix: analysis/results
    #out-format options: HTML, TSV-GENE, TSV-VARIANT, VCF (default: HTML)
    outputFormats: [TSV-GENE, TSV-VARIANT, VCF, HTML]
```

2. Starten des Durchlaufs

Auf dem cluster mit:

```
module load IKMB Java/1.8.0 && sbatch --exclude=rzcl236 -p ikmb_a --qos ikmb_a -J exomiser --mem=16G -t 2:0:0 --
wrap="java -Xms8g -Xmx16g -jar /ifs/data/nfs_share/ikmb_repository/software/exomiser/10.0.0/exomiser-cli-10.0.0.jar
--analysis analysis/config_exomiser.yml"
```

3. Ausführen von run_pipieline.py

Parameter:

```
exomiser_path = '/Users/broder/projects/IBD-DACH_FT_3/exomiser/analysis'
annovar_path = '/Users/broder/projects/IBD-DACH_FT_3/variants.merged.normalized.vcf.gz.annovar.hg19_multianno.txt'
output_path = '/Users/broder/projects/IBD-DACH_FT_3/result.tsv'
```

- Merged die Varianten von den verschiedenen Exomiser Inheritance Dateien zusammen

	H
	FILTER
2	AD_PASS;AR_PASS
2	AR_PASS
L	AR_PASS
4	AD_PASS
7	AR_PASS
2	AR_PASS
5	AR_PASS
2	AR_PASS
5	AR_PASS
5	AR_PASS
3	AR_PASS
7	AR_PASS
7	AR_PASS
2	AD_PASS;AR_PASS
5	AD_PASS;AR_PASS

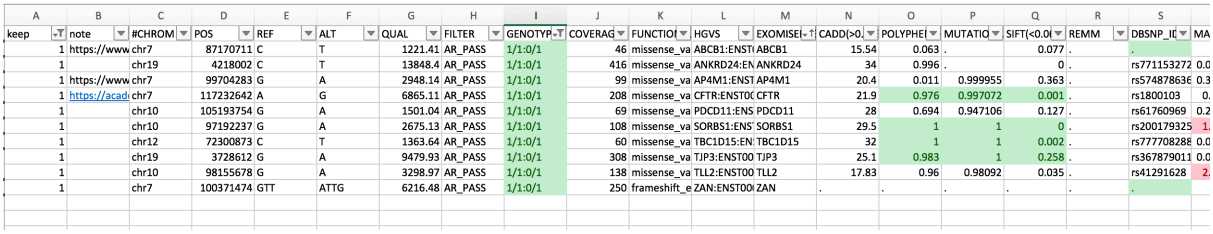
- Wendet die verschiedenen Genlisten auf die Varianten an

GeneFilter
ClinVar;imm
pedIbd
pedIbd
ClinVar;DisGeNET
ClinVar;DisGeNET
ClinVar;imm
DisGeNET
DisGeNET
DisGeNET
DisGeNET
DisGeNET
DisGeNET
DisGeNET
DisGeNET
DisGeNET
ClinVar;pedIbd
ClinVar;imm
ClinVar;imm
ClinVar;pedIbd;imm
ClinVar;pedIbd;imm
ClinVar;imm
ClinVar;imm

- Fügt die AnnoVar Annotation für jede Variante hinten an die Exomiser Outputdatei

4. Kopieren der Ergebnisse in das Result Template

5. Filtern



keep	note	#CHROM	POS	REF	ALT	QUAL	FILTER	GENOTYP	COVERAG	FUNCTIO	HGVS	EXOMISE	CADD(>0.	POLYPHE	MUTATIO	SIFT(<0.0	REMM	DBSNP_ID	MA
1	https://www	chr7	87170711	C	T	1221.41	AR_PASS	1/1:0/1	46	missense_va	ABCB1:ENSTI	ABCB1	15.54	0.063	.	0.077	.	.	
1		chr19	4218002	C	T	13848.4	AR_PASS	1/1:0/1	416	missense_va	ANKRD24:EN	ANKRD24	34	0.996	.	0	.	rs771153272	0.0
1	https://www	chr7	99704283	G	A	2948.14	AR_PASS	1/1:0/1	99	missense_va	AP4M1:ENST	AP4M1	20.4	0.011	0.999955	0.363	.	rs574878636	0.3
1	https://acad	chr7	117232642	A	G	6865.11	AR_PASS	1/1:0/1	208	missense_va	CFTR:ENST0	CFTR	21.9	0.976	0.997072	0.001	.	rs1800103	0.
1		chr10	105193754	G	A	1501.04	AR_PASS	1/1:0/1	69	missense_va	PDCD11:ENS	PDCD11	28	0.694	0.947106	0.127	.	rs61760969	0.2
1		chr10	97192237	G	A	2675.13	AR_PASS	1/1:0/1	108	missense_va	SORBS1:ENS	SORBS1	29.5	1	1	0	.	rs200179325	1.
1		chr12	72300873	C	T	1363.64	AR_PASS	1/1:0/1	60	missense_va	TBC1D15:EN	TBC1D15	32	1	1	0.002	.	rs777708288	0.0
1		chr19	3728612	G	A	9479.93	AR_PASS	1/1:0/1	308	missense_va	TJP3:ENST00	TJP3	25.1	0.983	1	0.258	.	rs367879011	0.0
1		chr10	98155678	G	A	3298.97	AR_PASS	1/1:0/1	138	missense_va	TLL2:ENST00	TLL2	17.83	0.96	0.98092	0.035	.	rs41291628	2.
1		chr7	100371474	GTT	ATTG	6216.48	AR_PASS	1/1:0/1	250	frameshift_e	ZAN:ENST00	ZAN	.	.	.	.	.	.	

6. Extrahieren der gefilterten Varianten zu einem vcf

```
Sub toVcf()

Dim sh As Worksheet
Dim rw As Range
Dim keep As Integer


outputFile = FreeFile
Open "/Users/broder/tmp/vcf_output.vcf" For Output As #outputFile
Print #outputFile, "#CHROM" & vbTab & "POS" & vbTab & "ID" & vbTab & "REF" & vbTab & "ALT" & vbTab & "QUAL" & vbTab
& "FILTER" & vbTab & "INFO"

Set sh = ActiveSheet
start = 3
For Each rw In sh.Rows
  If sh.Cells(rw.Row, 1).Value = "1" Then
    chromTmp = sh.Cells(rw.Row, start).Value
    chrom = Replace(chromTmp, "chr", "")
    pos = sh.Cells(rw.Row, start + 1).Value
    Id = sh.Cells(rw.Row, start + 16).Value
    ref = sh.Cells(rw.Row, start + 2).Value
    alt = sh.Cells(rw.Row, start + 3).Value
    qual = sh.Cells(rw.Row, start + 4).Value
    filterVal = sh.Cells(rw.Row, start + 5).Value
    info = "."
    Print #outputFile, chrom & vbTab & pos & vbTab & Id & vbTab & ref & vbTab & alt & vbTab & qual & vbTab &
filterVal & vbTab & info
  ElseIf sh.Cells(rw.Row, start).Value = "" Then
    Exit For
  End If

Next rw
Close #outputFile
End Sub
```

-> Einstellen in VarWatch