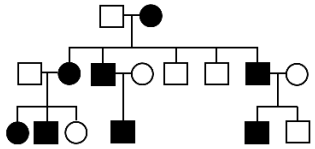
Clinical Workflow of Whole Exome Sequencing

2025/02/12

Hsin-Yu Huang MD | Department of Dermatology

In the clinic...

Phenotyping



Hereditary Pattern/ Pedigree

Sampling

Genetic Test

Member

Confidence of Dx?

Proband (patient)

+ Parents (trio)

+ Affected sibling

± Other members

Media

Blood (germline)

EDTA

Blood, saliva, tissue

± Tissue (somatic)

Formalin

± Normal tissue

IF: Michel's sol.

RNA: RNA later

EM: EM sol.

Commercially available services

MLPA, aCGH, FISH, Microarray

Sanger sequencing

Hotspots for common conditions

NextGen Sequencing

Whole exome (NTD ~10,000/case)

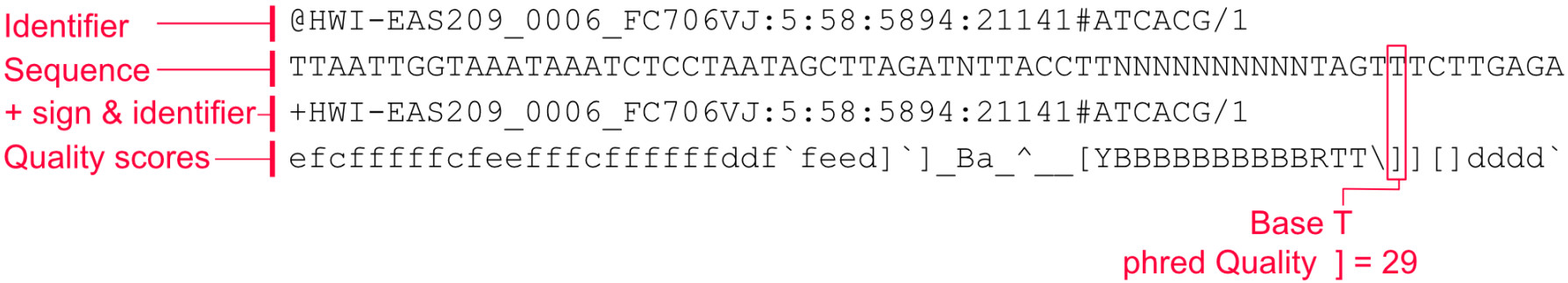
Whole genome, deep seq...

Similar technique also in RNAseq, ATACseq, scRNAseq...

Suitable for point mutation

❶ Raw sequences (paired-end fastq)

XXXX_R1.fastq + XXXX_R2.fastq

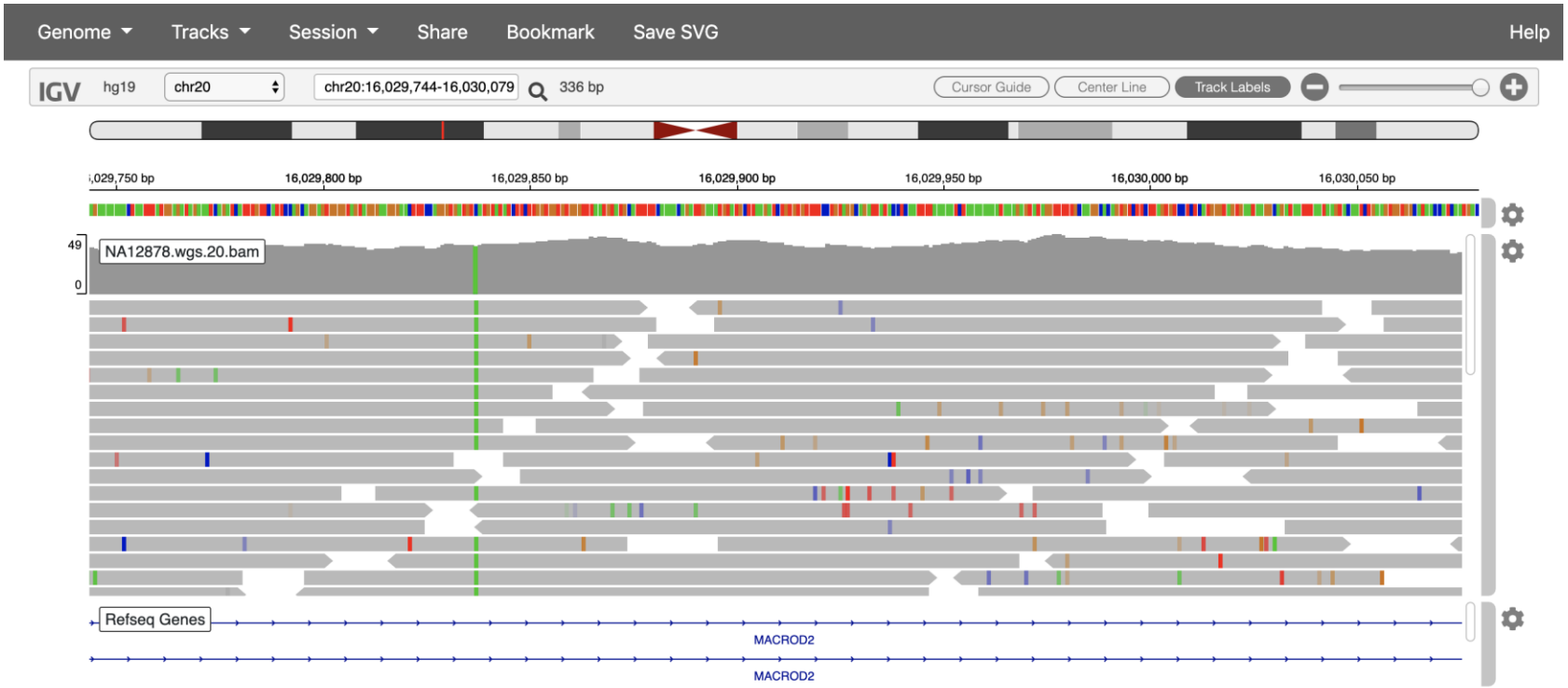


❷ Alignment and sort (*.BAM)

Align with human genome (version hg19)

With bwa mem

Sorting and index with samtools



❸ Variant calling (*.vcf)

Calibration (sometimes not performed)

With MarkDuplicate, BaseRecalibration, ApplyBQSR

Variant calling

With Haplotypecaller (germline), Mutect2 (somatic)

##bcftools_viewCommand=view -c1 -Oz -s A160251_Pro -o test.genotypecalls.A160251_Pro.vcf.gz test.genotypecalls.vcf; Date=Wed Jan 10 17:15:30 2018

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	A160251_Pro
chrM	150	.	T	C	6348.35	.	AC=2;AF=1;A	GT:AD:DP:GQ:PGT:PID:PL	1/1:0,39:39:99:.:.:989,115,0
chrM	195	.	C	T	5484.35	.	AC=2;AF=1;A	GT:AD:DP:GQ:PL	1/1:0,29:29:86:759,86,0
chrM	302	.	A	AC	454.02	.	AC=1;AF=0.7	GT:AD:DP:GQ:PL	0/1:1,5:6:9:110,0,9
chrM	410	.	A	T	1706.35	.	AC=2;AF=1;A	GT:AD:DP:GQ:PL	1/1:0,15:15:45:413,45,0
chrM	483	.	C	T	566.96	.	AC=2;AF=0.7	GT:AD:DP:GQ:PL	1/1:0,9:9:26:202,26,0

❹ Annotation (Excel files)

Defines variant pathogenicity

With ANNOVAR (may upload to 凱迪's website)

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Chr	Start	End	Ref	Alt	Func.refGene	Gene.refGene	GeneDetail.refGeneWithVer	ExonicFunc.refGeneWithV	AAChange.refGeneWithVer	Xref.refGeneWith	cytoBand	AF
2	1	948921	948921	T	C	UTR5	ISG15	NM_005101.4:c.-33T>C	.	.	Immunodeficiency	1p36.33	0.9457
3	1	1404001	1404001	G	T	UTR3	ATAD3C	NM_001039211.3:c.*91G>T	.	.	.	1p36.33	0.0559
4	1	5935162	5935162	A	T	splicing	NPHP4	NM_001291594.2:exon17:c.1282-2T	.	.	Nephronophthisis	1p36.31	0.8264
5	1	162736463	162736463	C	T	intronic	DDR2	.	.	.	Spondylometaepi	1q23.3	.
6	1	84875173	84875173	C	T	intronic	DNASE2B	.	.	.	.	1p31.1	.
7	1	13211293	13211294	TC	-	intergenic	PRAMEF36P;	dist=11566;dist=116902	.	.	.	1p36.21	.
8	1	11403596	11403596	-	AT	intergenic	UBIAD1;DISP	dist=43968;dist=135616	.	.	.	1p36.22	.
9	1	105492231	105492231	A	ATAAA	intergenic	LOC1001291	dist=872538;dist=640085	.	.	.	1p21.1	.
10	1	67705958	67705958	G	A	exonic	IL23R	.	nonsynonymous SNV	IL23R:NM_144701.3:exon9:c.G1142A:p.R381Q	.	1p31.3	0.0422
11	2	234183368	234183368	A	G	exonic	ATG16L1	.	nonsynonymous SNV	ATG16L1:NM_198890.2:exon5:c.A409G:p.T137A,ATG16L1:N	.	2q37.1	0.4532
12	16	50745926	50745926	C	T	exonic	NOD2	.	nonsynonymous SNV	NOD2:NM_001293557.2:exon3:c.C2023T:p.R675W,NOD2:N	Blau syndrome, A	16q12.1	0.0261
13	16	50756540	50756540	G	C	exonic	NOD2	.	nonsynonymous SNV	NOD2:NM_001293557.2:exon7:c.G2641C:p.G881R,NOD2:N	Blau syndrome, A	16q12.1	0.0113
14	16	50763778	50763778	-	C	exonic	NOD2	.	frameshift insertion	NOD2:NM_001293557.2:exon10:c.2936dupC:p.L980Pfs*2,N	Blau syndrome, A	16q12.1	0.015
15	13	20763686	20763686	G	-	exonic	GJB2	.	frameshift deletion	GJB2:NM_004004.6:exon2:c.35delG:p.G12Vfs*2	Bart-Pumphrey sy	13q12.11	0.006
16	13	20797176	21105944	0	-	exonic	CRYL1;GJB6	.	startloss	GJB6:NM_001110219.3:exon1:c.1_786del:p.M1?,GJB6:NM_	.	13q12.11	.
17	8	8887543	8887543	A	T	exonic	ERI1	.	stoploss	ERI1:NM_001354635.2:exon7:c.A815T:p.X272L,ERI1:NM_15	.	8p23.1	.
18	8	8887539	8887539	A	T	exonic	ERI1	.	stopgain	ERI1:NM_001354635.2:exon7:c.A811T:p.K271X,ERI1:NM_15	.	8p23.1	.
19	8	8887536	8887537	AG	GATT	exonic	ERI1	.	stopgain	ERI1:NM_001354635.2:exon7:c.808_809delinsGATT:p.R270	.	8p23.1	.
20	8	8887540	8887540	G	GGAA	exonic	ERI1	.	nonframeshift substitutio	ERI1:NM_001354635.2:exon7:c.812delinsGGAA:p.R270_K2	.	8p23.1	.
21	5	1295288	1295288	G	A	upstream	TERT	dist=105	.	.	.	5p15.33	.
22	chr14	95602958	95602958	A	C	splicing	DICER1	NM_001271282.3:exon1:UTR5	.	.	Goiter, multinodu	14q32.13	.

❺ Filtering (Excel files)

Work on pathogenic variants

"Rare, damaging and clinically reported?"

Rare	Functional damaging	Splicing damaging	Clinically damaging
Minor Allele Freq. (MAF)	Exome or near splice-site	Spidex (dpsi)	Clinvar
gnomAD <1%	CADD >10	dbscsnv	LOVD (凱迪ver.)
1000genome	DANN	Varseak (外部)	HGMD (凱迪ver.)
ExAC	MutationTaster		
TWbiobank	MetaSVM/LR		

Newer damaging prediction database available in VEP and other annotation program (etc., SpliceAI)

❺ Filtering (Excel files)

Work on pathogenic variants

"Rare, damaging and clinically reported?"

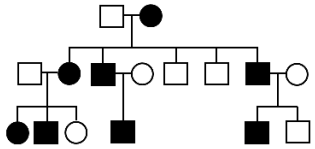
Rare	Functional damaging	Splicing damaging	Clinically damaging
Minor Allele Frequency (MAF)	Functional prediction	Splicing prediction	Clinical
gnomAD <1%	splice-site	dbscsnv	LOVD (凱迪ver.)
1000genome	CADD >10		HGMD (凱迪ver.)
ExAC	DANN		
TWbiobank	MutationTaster		
	MetaSVM/LR		

Still, thousands of candidates remained.
How to even narrow down?

Newer damaging prediction database available in VEP and other annotation program (etc., SpliceAI)

In the clinic...

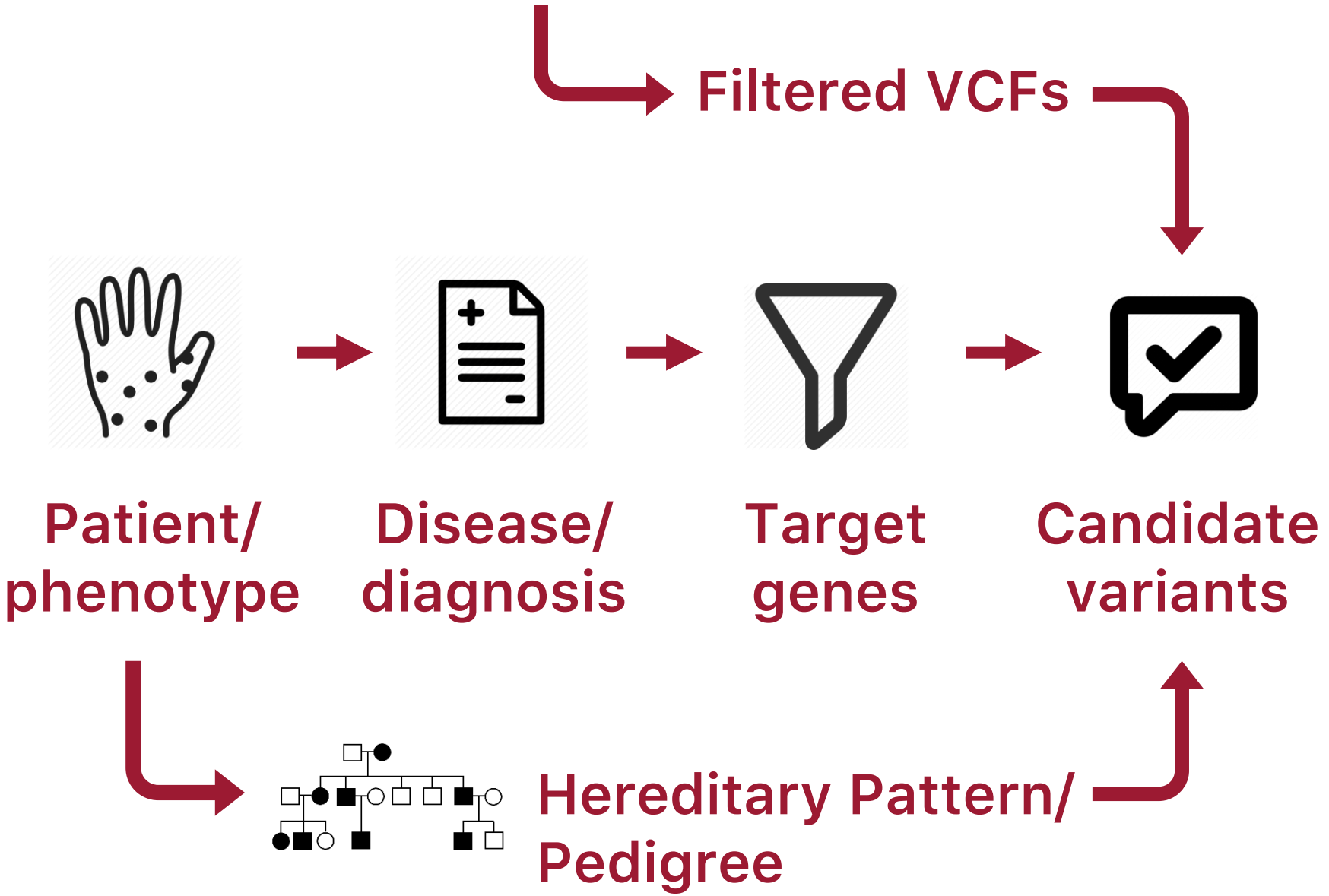
Phenotyping

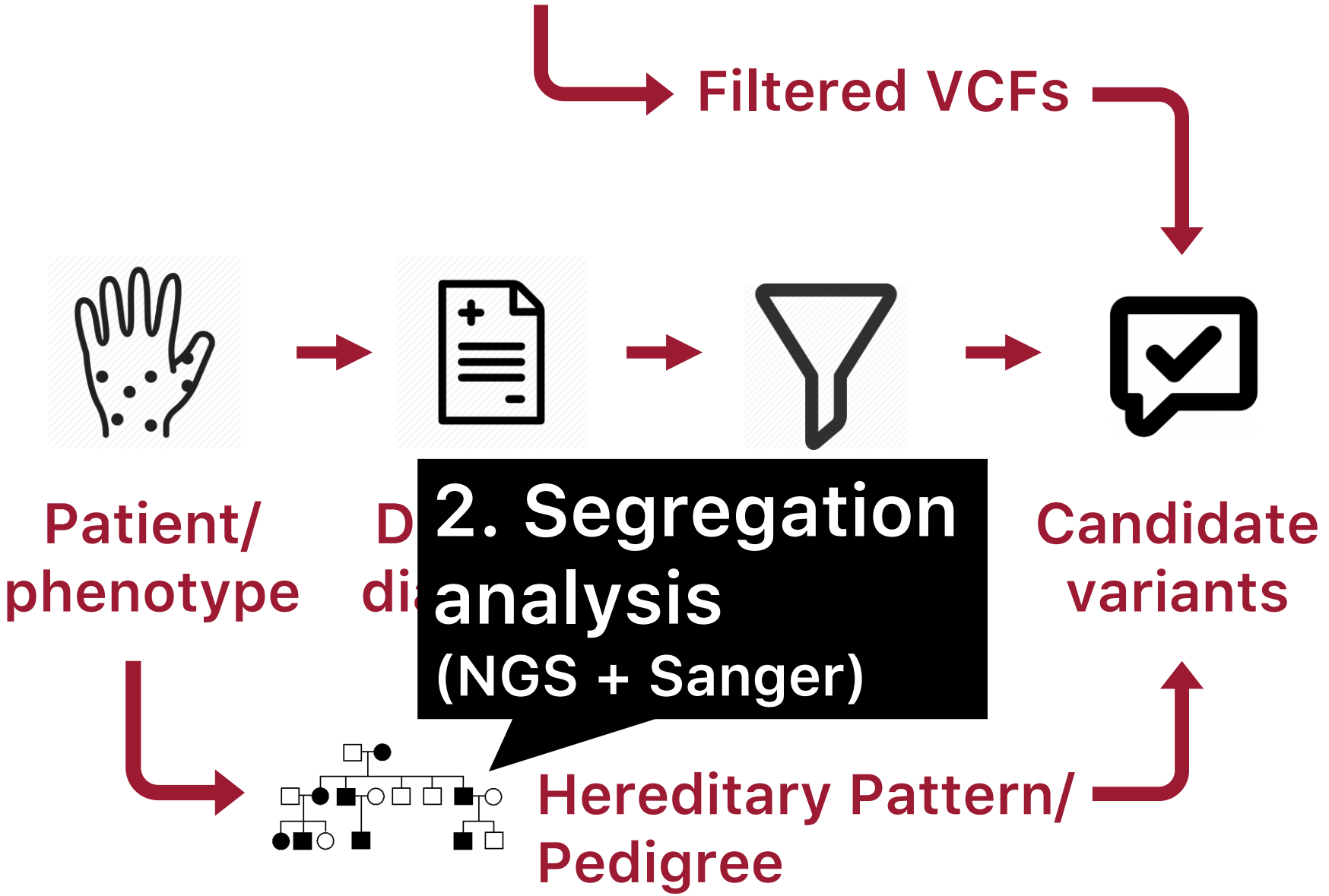


Hereditary Pattern/ Pedigree

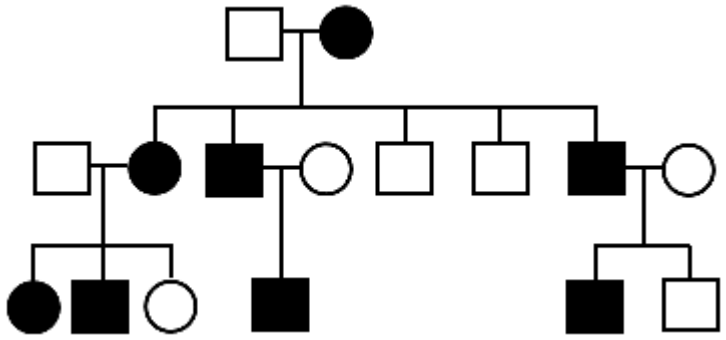
Sampling

Genetic Test


Filtered VCFs
Patient/ phenotype
Disease/ diagnosis
Target genes
Candidate variants
Hereditary Pattern/ Pedigree


Filtered VCFs
Patient/
phenotype
D
di
2. Segregation
analysis
(NGS + Sanger)
Candidate
variants
Hereditary Pattern/
Pedigree

In same family
Of same condition

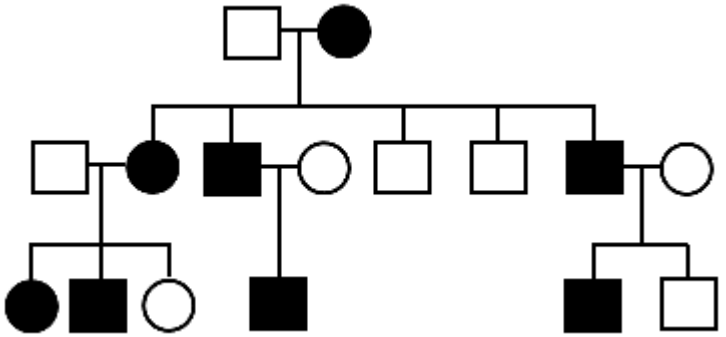


Same variant

e.g.
KRT14:exon1:c.123A>C:a.41G>Y

- Autosomal Dominant
  - HET variant
  - *De novo* mutation
- Recessive
  - HOM variant
  - Compound HET
- X-linked
- Semidominant

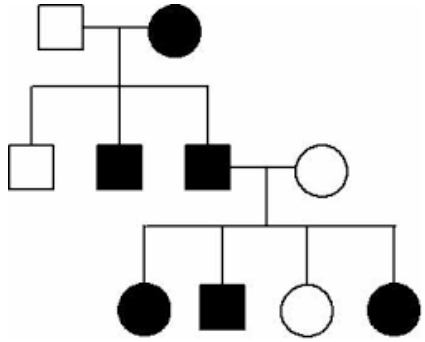
Same condition



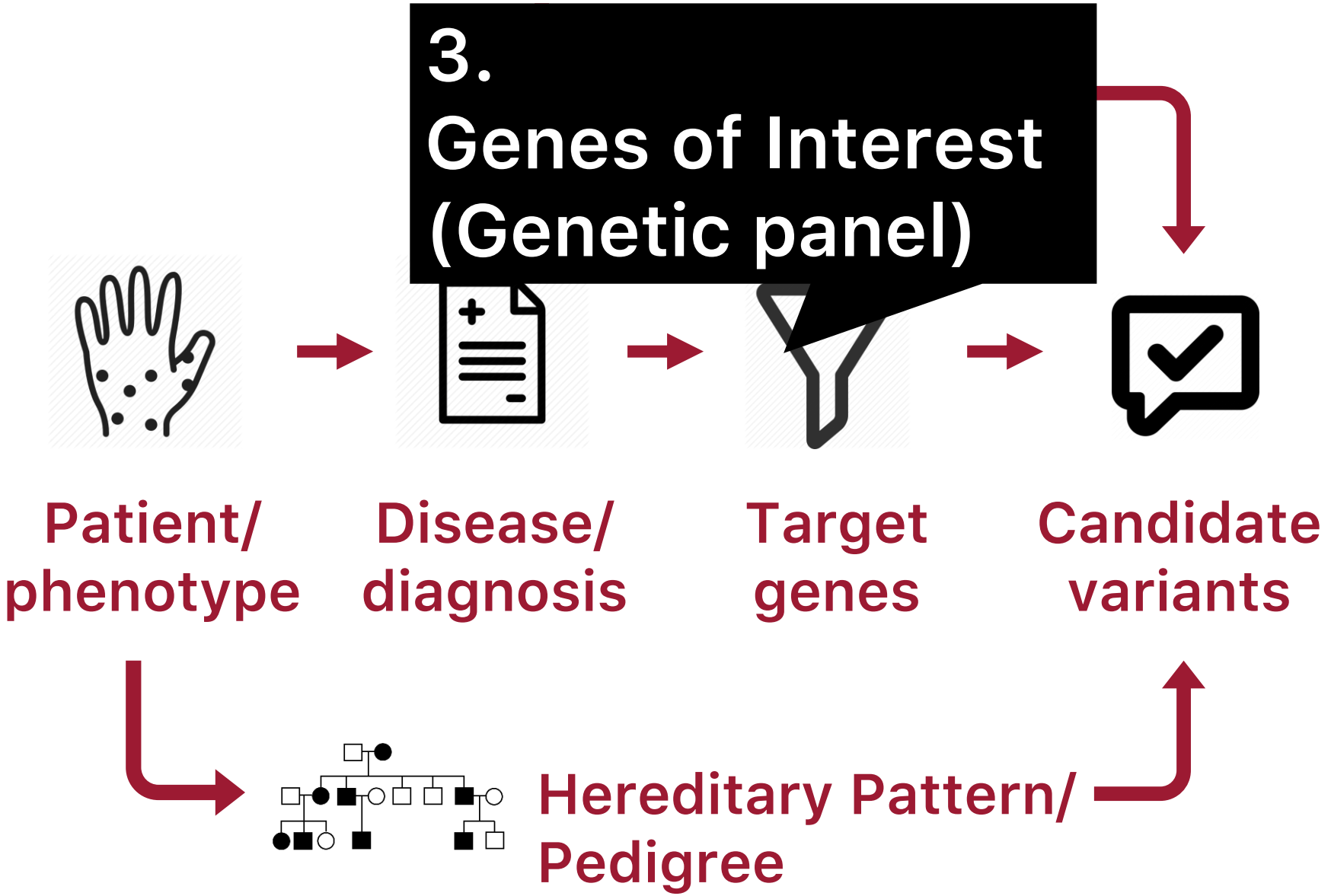
Same variant

Same gene

e.g. KRT14



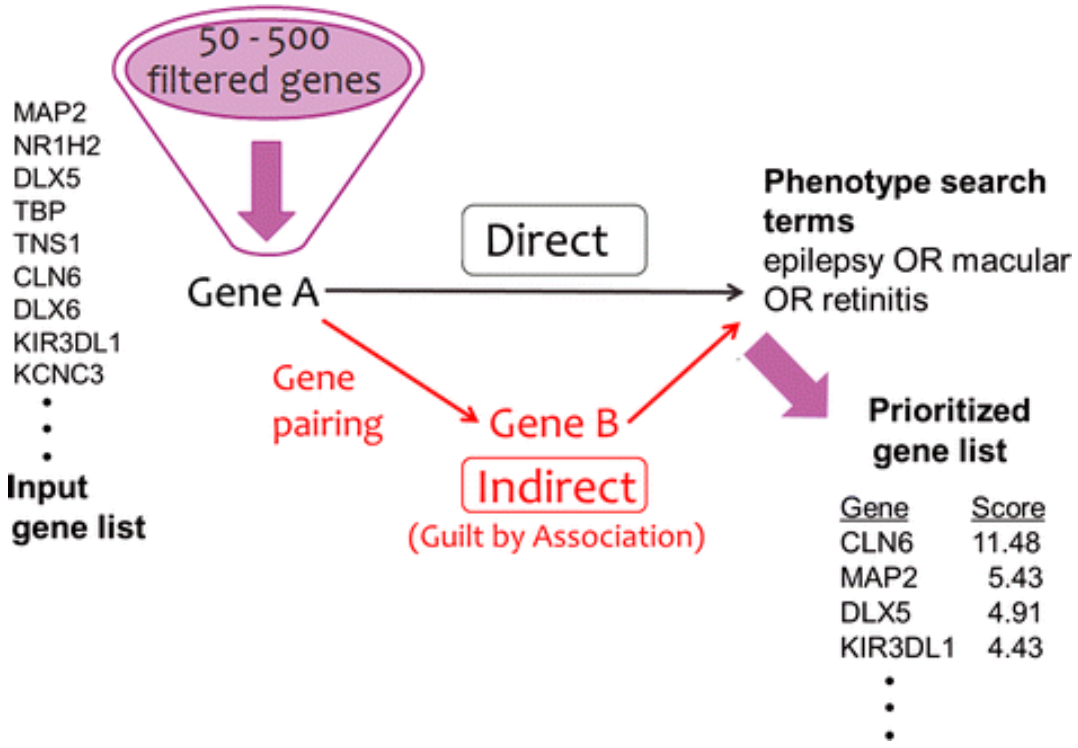
Same variant


3.
Genes of Interest
(Genetic panel)
Patient/
phenotype
Disease/
diagnosis
Target
genes
Candidate
variants
Hereditary Pattern/
Pedigree

VarElect: Phenotyper

Fast screening; no specific candidates

Phenotype + Gene → Association scores



In-house Panel of Genodermatoses

EB, ED, PPK/ PC, Ichthyosis, Inflammatory, Pigmentary, Epidermal, Tumoral, Miscellaneous…

Human Phenotype Ontology (HPO)

Online Mendelian Inheritance in Man (OMIM)

Human Gene Mutation Database (HGMD)

Genetic Testing Registry (GTR)

GO/ GSEA/ Ingenuity pathway explorer

Google/ PubMed

Clinical Workflow of Whole Exome Sequencing

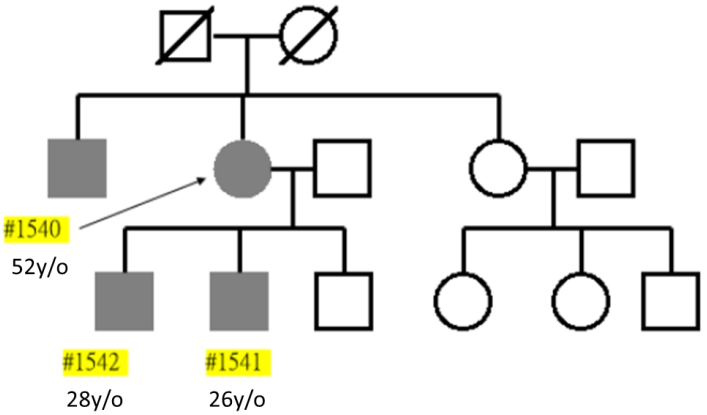
2025/02/12

Hsin-Yu Huang MD | Department of Dermatology

What to show in the report?

- Phenotype, pedigree
  - Possible diagnosis/ genes of interest
- Sample + methods
- Lists of variants and segregation
  - Include gnomAD, CADD, clinvar, etc...
  - Validation with Sanger sequencing or Integrative Genomics Viewer (IGV)

Palmoplantar keratoderma, teeth anamolies, nail dystrophy



Favor autosomal dominant

Dx: PPK & Pachyonychia congenita

Blood of Proband + affected case

Germline mutation for PPK panel

> 28 genes of PPK/PC panel → WES

AAGAB, AQP5, CFTR, COL14A1, COL20A1, CTSC, DSC2, DSG1, DSP, ENPP1, FAM83G, GJB2, GJB3, GJB4, GJB6, JUP, KRT1, KRT16, KRT17, KRT6A, KRT6B, KRT6C, KRT9, LOR, MTTS1, NLRP1, RHBDF2, RSPO1, SERPINB7, SLURP1, TRPV3

An Example of Variant/ Report

Location	Gene	RS ID	MAF	Genotype (#ref/#alt)	Evidence	Domain	Pathogenicity	Splicing effect	OMIM
chr17:39768567T>C NM_005557:exon1:c.A374G:p.N125S	KRT16	rs60723330	gnomAD: 0.0 1000G : 0.0 TW Biobank: 0.0	het (35,45)	Clinvar: Pathogenic(2★) LOVD: .	.	Summary : (7/8) Polyphen2_HVAR: D SIFT: D VEST3: 0.416 MutationTaster: D MetaSVM: D MetaLR: D CADD: 24.4 DANN: 0.998	Summary : (0/0) dbscsnv ADA score: . dbscsnv RF score: . SPIDEX zscore: .	Palmoplantar keratoderma, nonepidermolytic, focal, 613000 (3)(AD) 148067

Family members confirmed by Sanger seq