生技醫藥生物資訊核心設施

Bioinformatics Core Facility for Biotechnology and Pharmaceuticals

熊昭	特聘研究員/所長	國家衛生研究院
何信瑩	教授	國立交通大學
江士昇	助研究員	國家衛生研究院
孫孝芳	教授/所長/基因體醫學中心主任	國立成功大學
呂平江	教授/副校長	國立清華大學
許聞廉	特聘研究員	中央研究院

灝湍　髂
擷昄


功能基因體及轉譯醫學
熊昭特聘研究員/所長，國家衛生研究院
轉錄體學及生醫影像分析
何信瑩教授，國立交通大學
癌症基因體學及臨床研究
江士昇助研究員，
國家衛生研究院
應用基因體醫學
孫孝芳教授/所長，
國立成功大學
結構蛋白質體學及藥物應用
呂平江教授/副校長，
國立清華大學
生醫文獻探勘及生物標記探索
許聞廉特聘研究員，中央研究院
NHRI
協調中心

核心設施簡介

生技醫藥生物資訊核心設施

使用者委員會

協調中心

專職管理團隊(NHRI)服務窗口

客製化服務、線上服務、諮詢服務、教育訓練

子計畫

子計畫一：功能基因體及轉譯醫學
子計畫二：轉錄體學及生醫影像分析
子計畫三：癌症基因體學及臨床研究
子計畫四：應用基因體醫學
子計畫五：結構蛋白質體學及藥物應用
子計畫六：生醫文獻探勘及生物標記探索

入口網站/電子報

線上資料庫25種

線上分析工具52種

教育推廣

國際研討會

客製化資料分析服務

- 高通量實驗客製化資料分析
- 生醫研究應用客製化資料分析服務
- 資料庫設計和開發
- 蛋白質-化合物與藥物設計整合型分析服務

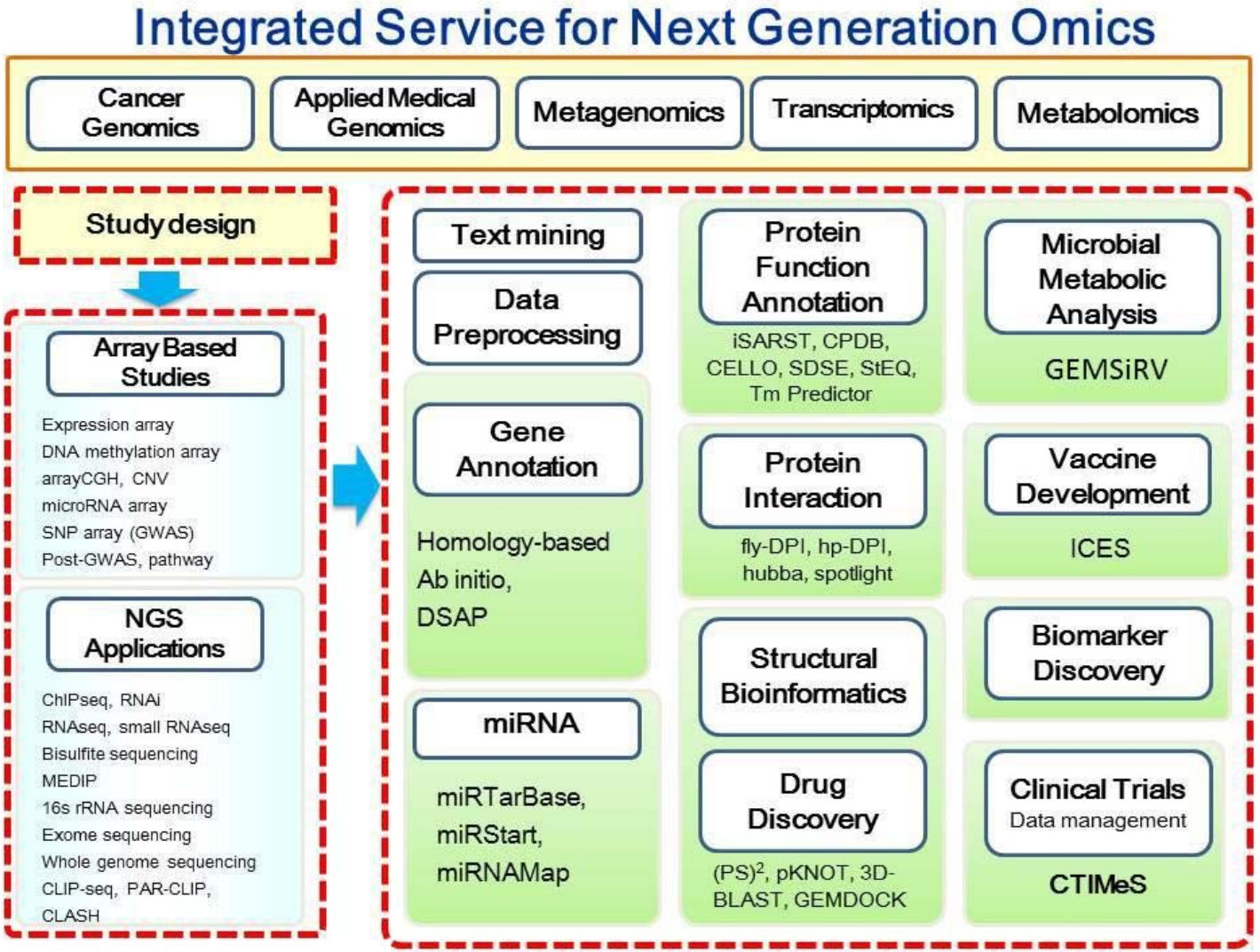
合作研究

技術研發

分析方法/分析平台
分析工具/資料庫

任務	百分比
服務	75%
技術研發	10%
教育推廣	10%
合作研究	5%

提供生技轉譯所需的生物資訊服務



一站式客製化服務

使用者

TMBD 核心: 南北設站、服務轉介、產業合作



填表申請

繳費

研究會議

取得分析結果



需求訪談

研究設計
平台選擇

實驗轉介

資料前處理

資料分析

生物意義分析
預測分析
臨床應用
結果解讀

產業合作與服務轉介



■ NGS
- 均泰生物科技
- 美基生物科技
- 康健生物科技
- 有勁生命科技
- 圖爾斯生物科技

■ NanoString
- 冷泉港生物科技

■ 人工智慧及文獻探勘
- 網際智慧
- 行動基因生物科技
- 數冠科技

■ 蛋白質結構
- 頂尖生技
- 康邁醫學科技

■ RNA定序
- 先見基因科技
- 百歐生命科技
- 基龍米克斯

■ 客製化晶片
- 華聯生物科技

■ 雲端系統開發整合
- 遺傳軌跡股份有限公司

■ 微生物鑑定
- 善笙生物科技

■ 醫藥大數據分析
- 柏瑞醫

■ 合辦研討會研習會 有勁生物科技、康健基因科技、均泰生物科技、冷泉港生物科技

核心設施執行成果

(108年5月1日至109年1月31止)

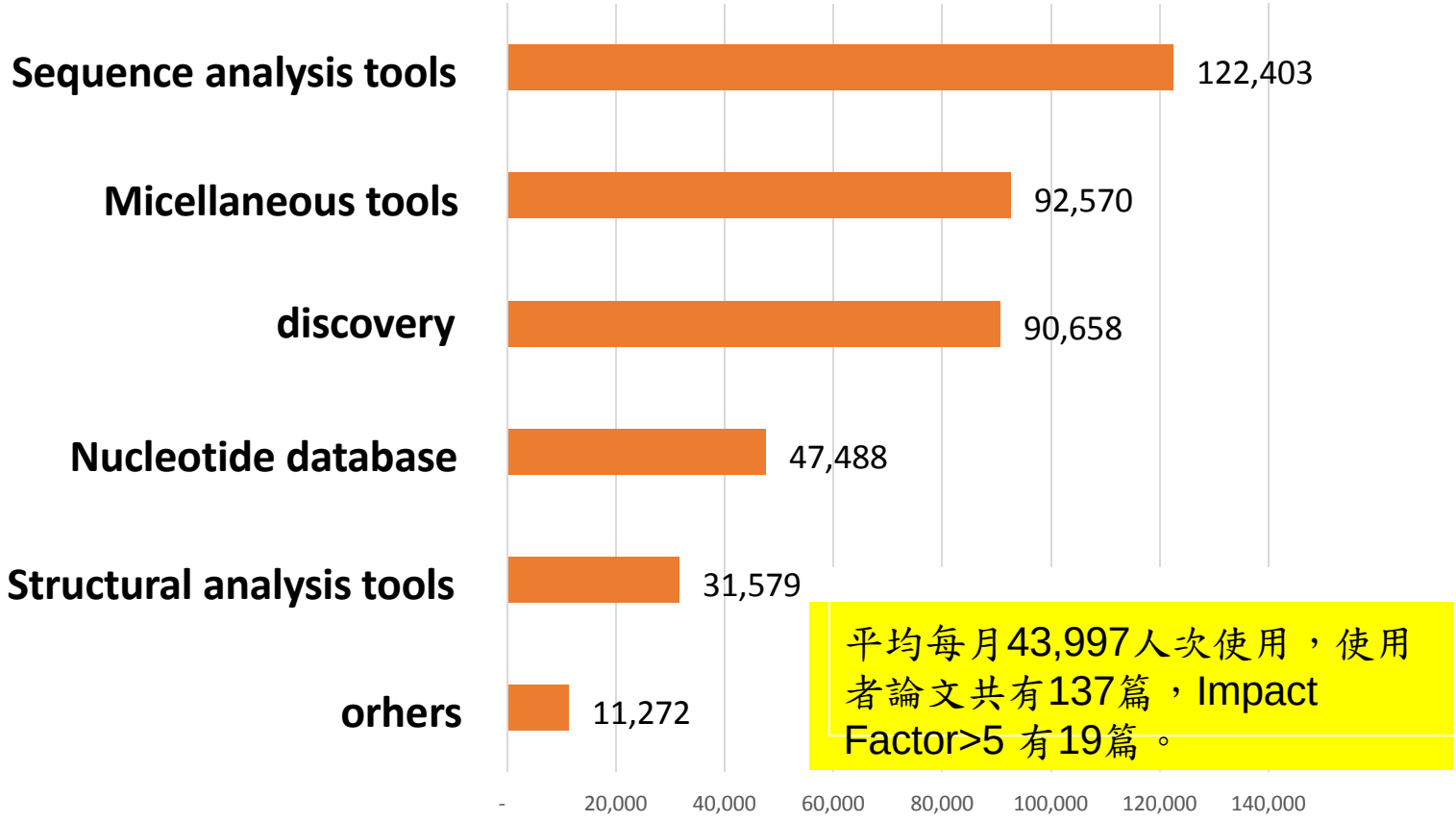
整合型線上分析工具與資料庫

將研發成果開發為分析工具與資料庫

基因體分析工具	引子/探針設計工具
NGS資料分析工具	蛋白質交互作用搜尋
microRNA分析工具	演化序列分析工具
蛋白質結構分析工具	流感病毒及腸病毒監測系統
電腦輔助藥物設計	加值資料庫
代謝體質譜定量軟體	生醫文獻探勘系統

提供52種分析工具，25種資料庫，滿足多樣化的生物資訊需求。

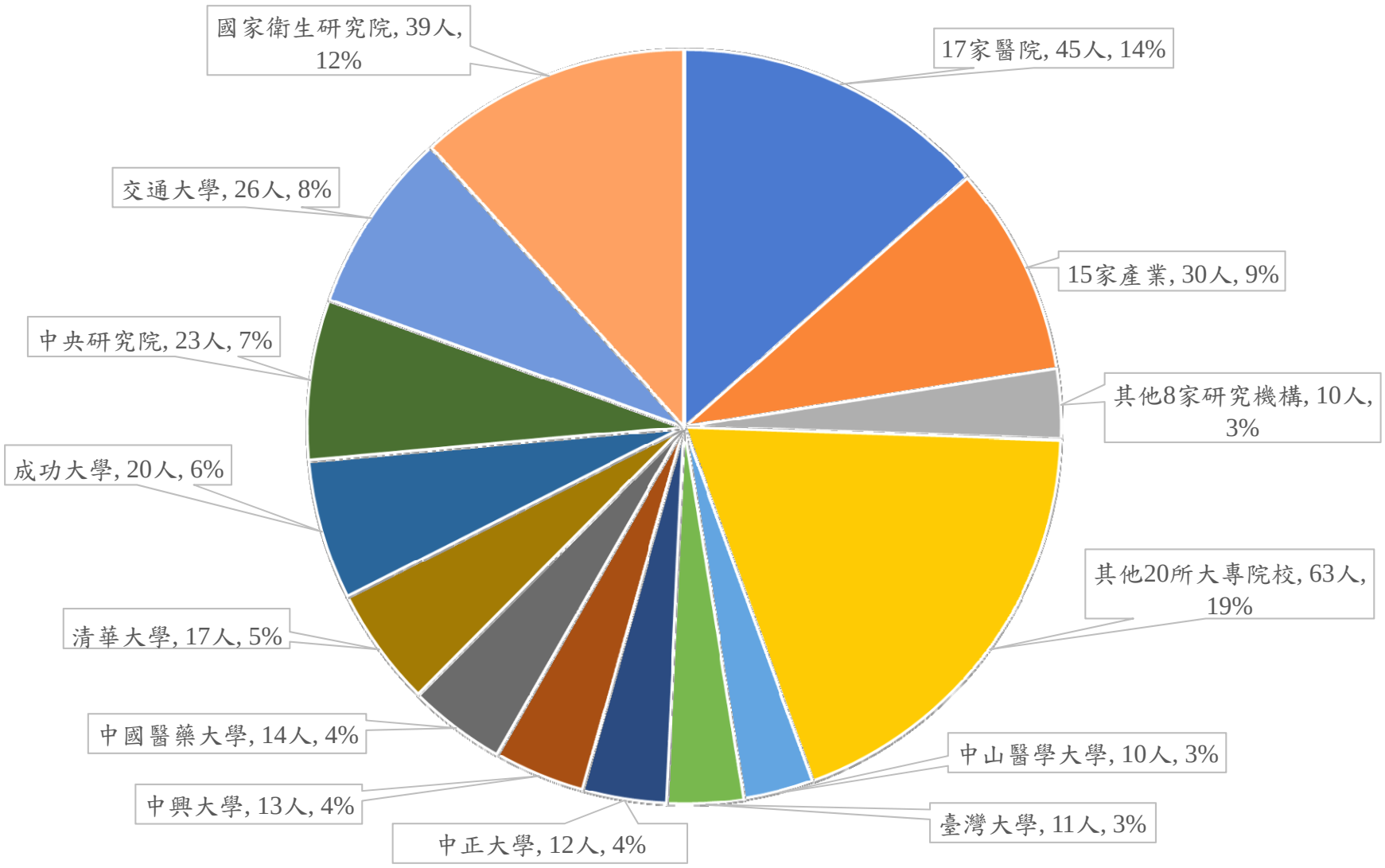
線上服務使用人次
108.05.01~109.01.31



平均每月43,997人次使用，使用者論文共有137篇，Impact Factor>5 有19篇。

教育訓練人次分佈

108.05.01~109.01.31



教育推廣

訓練及推廣活動

舉辦國際研討會
應邀至印尼峇里島、
馬來西亞舉辦生物資訊
實作課程

提升我國
生物資訊能見度



定期常態性教育訓練課程
高通量資料分析、基因體
分析暨蛋白質結構生物資
訊分析、次世代定序數據
分析 (106/5~109/4 共30場)

協助產學界
培養高階人才



跨校/跨領域/業界合辦
客製化生物資訊研習會
(106/5~109/4 共17場)

客製化內部職訓
提升產學界競爭力



跨校/跨領域/業界合辦之收費式生物資訊研習會

本核心研習會為收費制，結合各子計畫之專長、經驗和成果，朝向跨校園、跨領域及業界合作等，開課地點遍佈北部、中部、南部及東部，讓各地對於生物資訊有需求之學者也能就近參與課程。

	鶦鱳	堷岭	擛
	2018/8/28	高通量資料分析研習會:以具臨床應用性之基因表現標記分析為例	國衛院(竹南)
	2019/2/14-15	大小分子結構生物資訊分析研習會	中興大學
	2019/2/19-20	Workshop on Bioinformatics tools for NGS-data analysis	高雄醫學大學
	2019/7/10-11	次世代定序及生物資訊分析研習營	高雄醫學大學
	2019/11/7	Exome-Seq Data Analysis 全外顯子組定序數據分析研習會	中興大學
	2019/12/5	生物資訊工作坊：多體學視覺化分析平台 UCSC Xena	清華大學
	2020/1/16-17	大小分子結構生物資訊分析研習會	成功大學
	2020/1/30-31	大小分子結構生物資訊分析研習會	清華大學
	2020/4/23~	全民防疫 學習不間斷~ UCSC Xena線上課程	

收費標準與作業流程

收費標準

服務代號	服務項目	學術界	產業界
E2-1	整合型線上分析工具與資料庫	-	-
E2-2	高通量實驗資料分析	250元/小時	680元/小時
E2-3	生醫研究應用資料分析	250元/小時	680元/小時
E2-4	資料庫設計和開發	250元/小時	680元/小時
E2-5	蛋白質-化合物與藥物設計 整合型分析服務	250元/小時	680元/小時
E2-6	教育訓練與研習會	(依場地而定)	

客製化服務作業流程

1. **填表:申請人填寫客製化服務申請表，email寄送至本核心服務窗口:**

聯絡人：齊淑雯
TEL: +886-37-246-166 ext.33602
FAX: +886-37-586-410
E-mail: service@tbi.org.tw

2. 簽約:本核心洽使用者商訂服務合約書暨保密協定，雙方議定服務時數及費用，並填寫使用同意證明單，交雙方簽署。

3. 繳費:本核心以email繳費通知予申請人，請申請人依通知繳費，費用須一次付清，匯款帳號如下，請注意: 匯款後請將匯款憑證影本連同使用證明單傳真/email予收費服務窗口，以便確認收款。

(1)收費銀行:台新銀行
(2)收費帳號:004-10-070557-7-00
(3)收費戶名:財團法人國家衛生研究院
(4)開立收據(免稅發票)單位:財團法人國家衛生研究院
(5)收費服務聯繫窗口人員姓名及電話:齊淑雯小姐，037-246166*33602

4. 提供服務:本核心於收到匯款憑證後即開立發票(免稅發票)寄交申請人，並通知服務人員開始服務。服務人員將洽申請人取得資料，服務期間將視需要召開研究會議討論資料及分析結果。請注意:服務起始將以繳費後且取得資料之日起算。

5. 結案:本核心繳交分析結果，經申請人同意後結案，並寄出滿意度調查表請申請人填寫。

6. 成果發表: 申請人如有成果發表，請於Acknowledgement或Method中註明本核心之貢獻。

Thank you for your attention!



Establish a Whole Exome Sequencing Pipeline for Prenatal Healthcare

Dr. H. Sunny Sun
Director, IMM / Director, CGM
National Cheng Kung University

Services provided by NCKU Center for Prenatal Medicine

http://cgm.ncku.edu.tw/

次世代高通量定序　核酸自動定序　分子檢驗及分析　微生物快速鑑定　生物資訊分析



專業 Profession 精確 Accuracy 卓越 Excellence 創新 Innovation

高品質、客製化服務，歡迎洽詢 電話 06-2353535-3614、3632 網址 http://cgm.ncku.edu.tw

Laboratory Developed Tests

Preconception

- Preimplantation Genetic Testing -aneuploidy (PGT-A)
- Preimplantation Genetic Testing-monogenic disease (PGT-M)
- Carrier test

Pregnancy

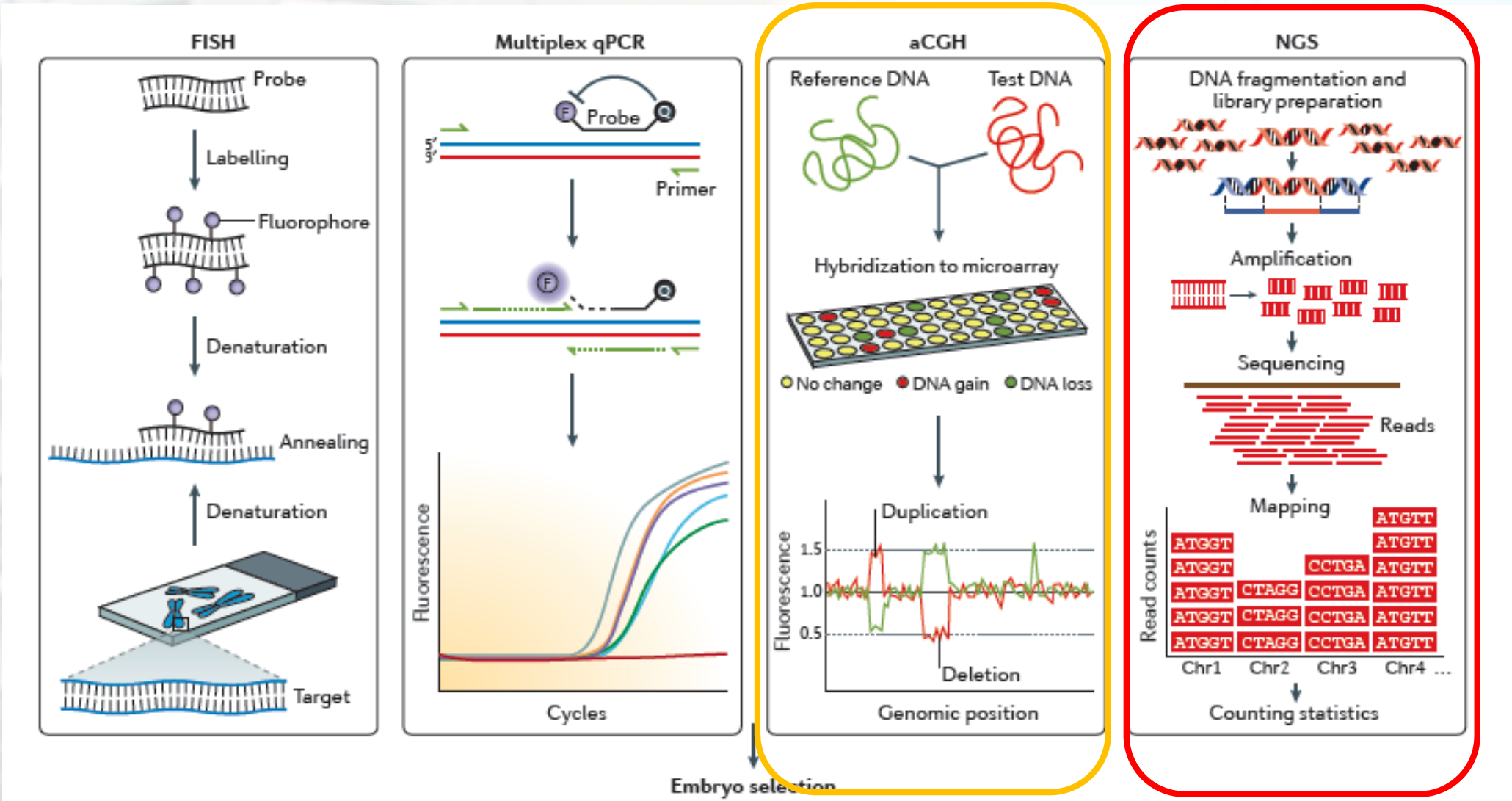
-Non-invasive Prenatal Screening (NIPS)
-Non-invasive Prenatal Testing (NIPT 2.0)
-Spinal muscular atrophy(SMN1/SMN2 Genetic testing)
-Fragile X syndrome (FMR1 Genetic testing)

Newborn

- Congenital central hypoventilation syndrome (CCHS, PHOX2B Genetic testing)
- Hearing loss (GJB2, SLC26A4, 12S rRNA Genetic testing)

Preimplantation genetic testing-aneuploidy (PGT-A)

Current techniques used in PGT



Vermeesch et al., 2016, NATURE REVIEWS | GENETICS 17:643-656.

NCKU-CGM NGS-based PGT pipeline

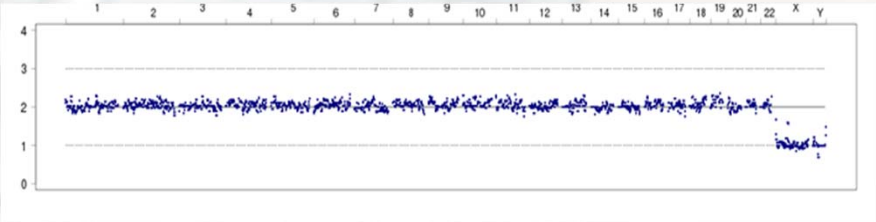


1-10 cells from the Blastocyst

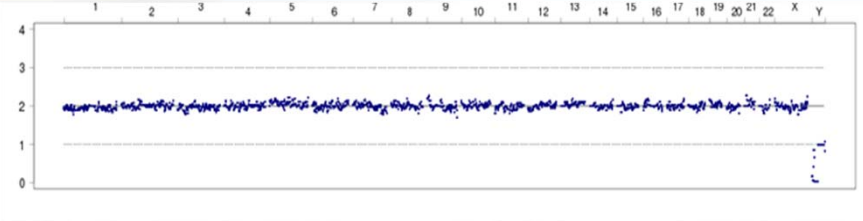
- Whole Genome Amplification ← WGA kit
- Sequencing ← Sequencing platform
- Data collection
- Data processing
- Reads in Fastq
- In house analytic pipeline ← Bioinformatics pipeline
  - Mapping, coverage analysis, normalization, copy number calling, chromosome plot

Genome aneuploidy analysis

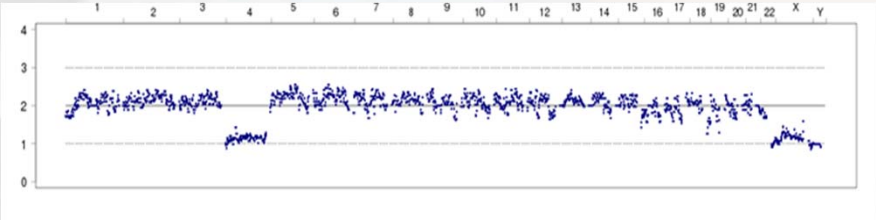
46,XY



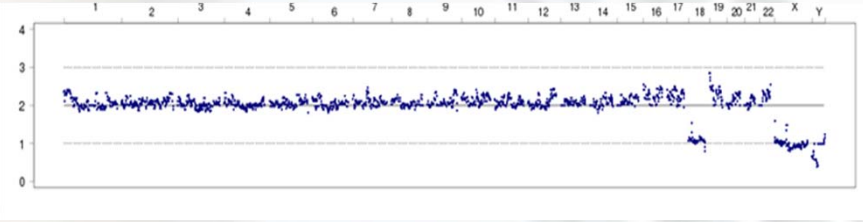
46,XX



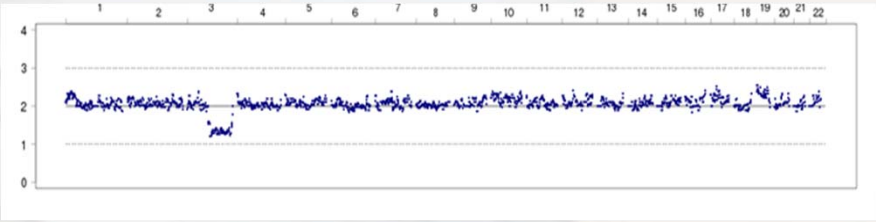
46,XY, -4



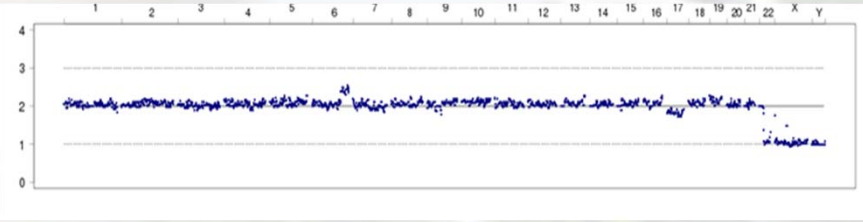
46,XY, -18



46,XX (mosaicism)

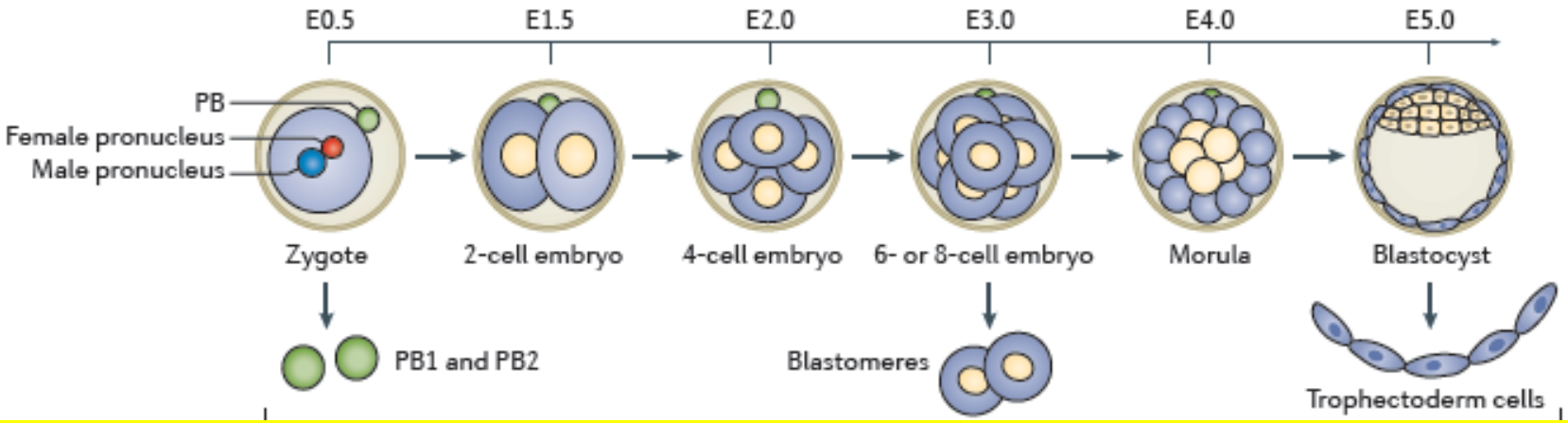


46,XY, -22, (mosaicism)

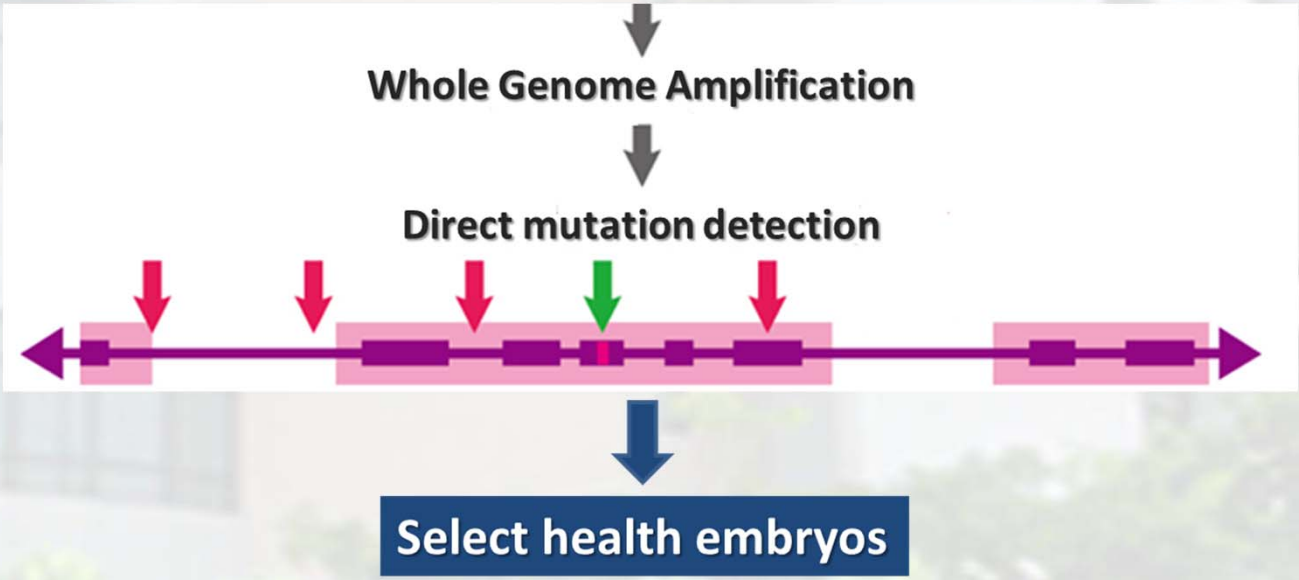


Preimplantation genetic testing-Monogenic PGT-M

PGT-M for known mutations

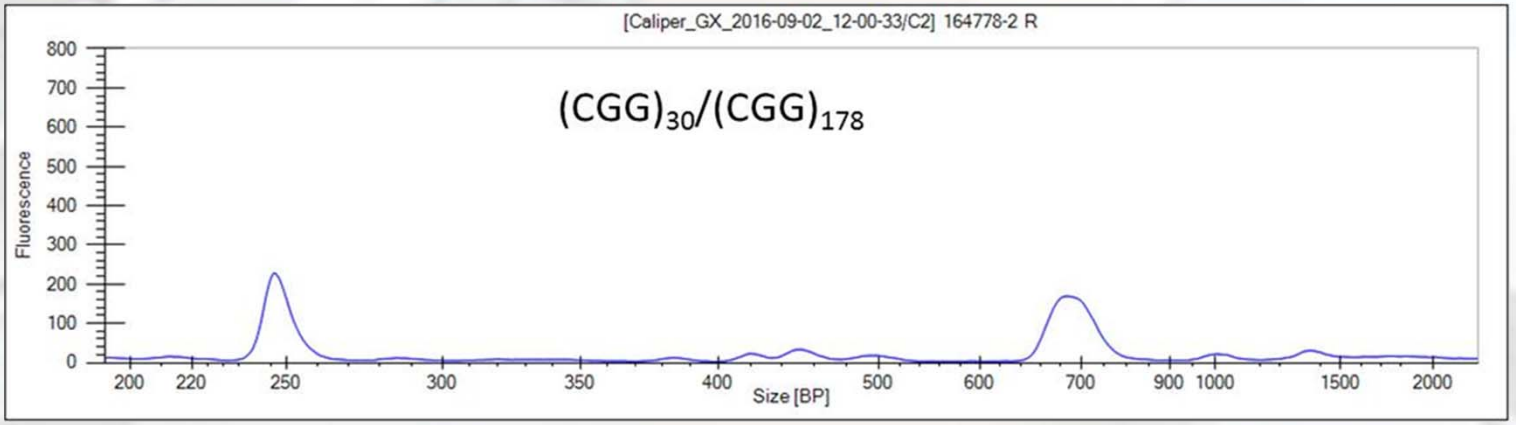


For the diseases with known causative mutations



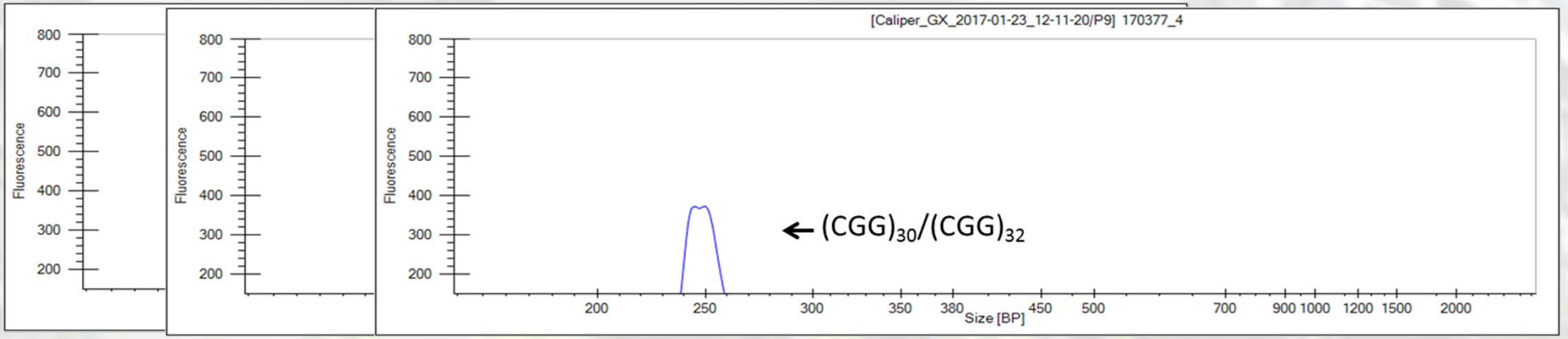
PGT-M for FMR1

- Mother is a FMR1 pre-mutation carrier





170377-1 170377-3 170377-4

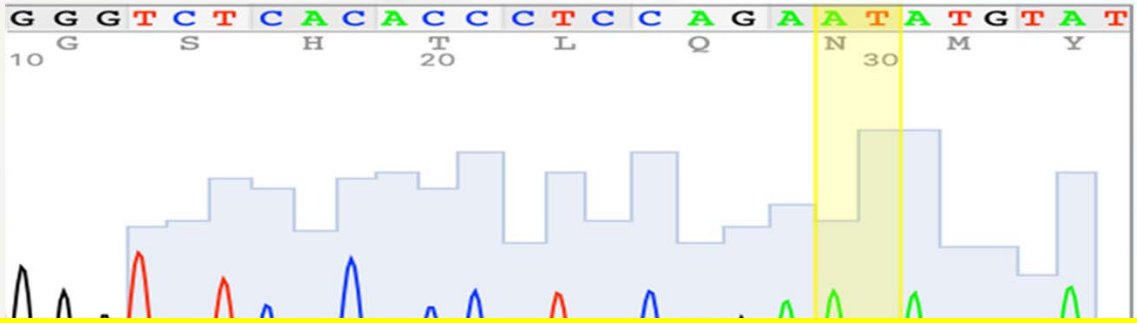


PGT-M for HLA-B27

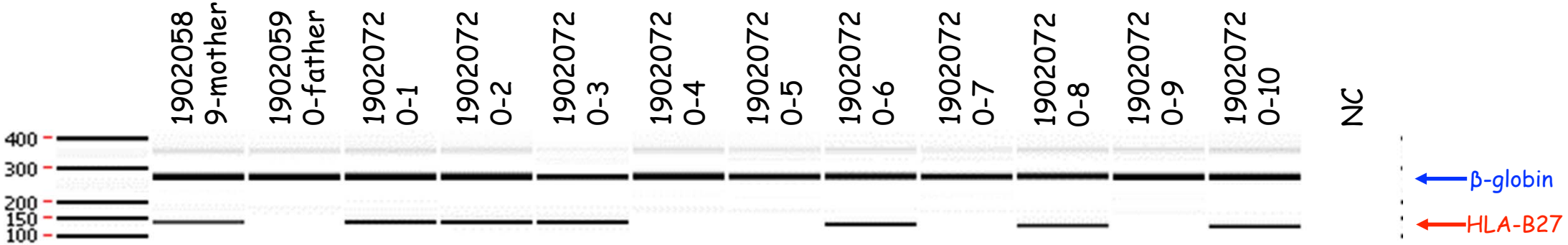
ASO-PCR
c.385-c.386AT

Parents

Mother



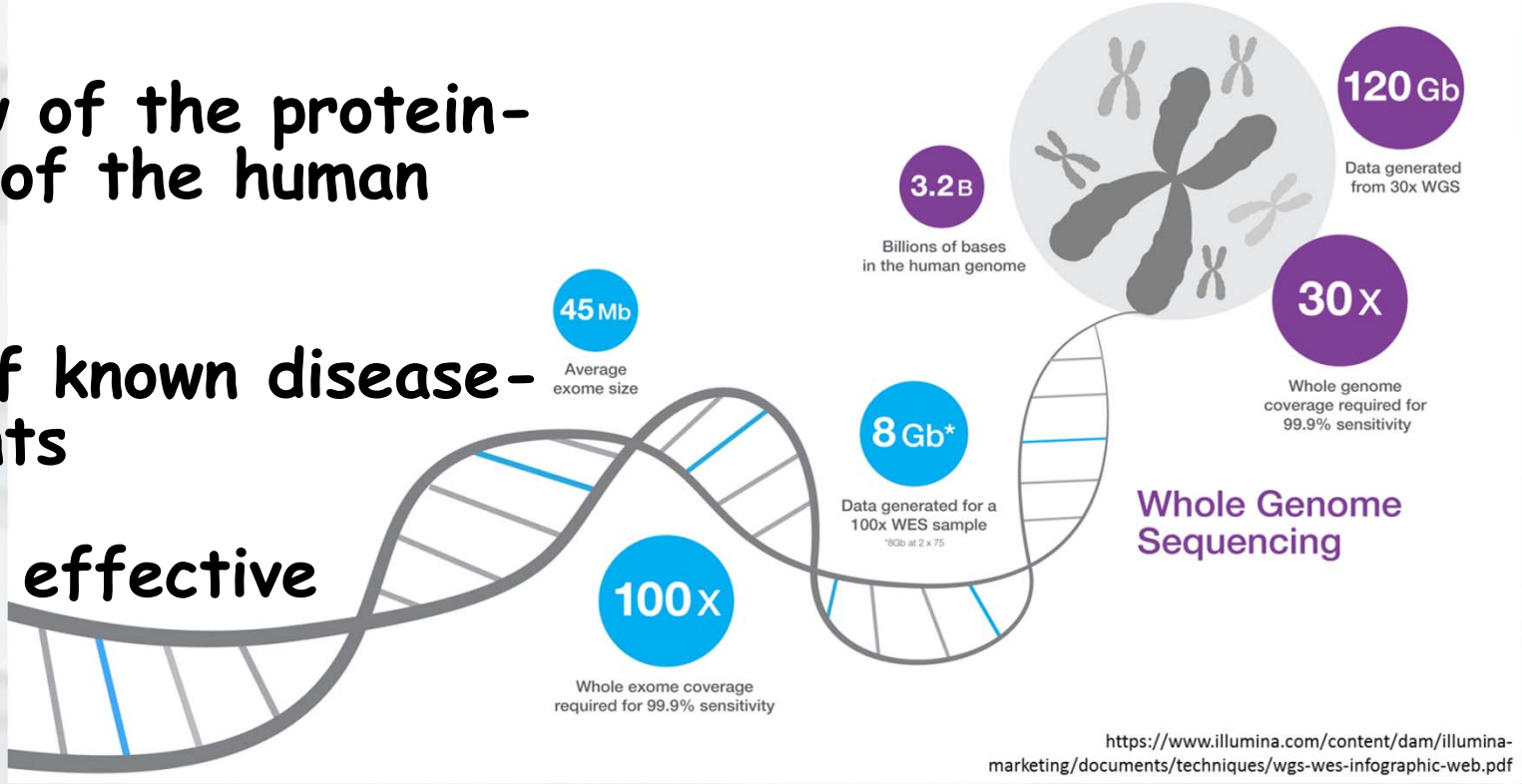
PGT-M effectively prevents inheritance of genetic defects



For the diseases with unknown causative mutations?

Whole Exome Sequencing (WES)

- Targeted view of the protein-coding region of the human genome (1%)
- Covers 85% of known disease-related variants
- Fast and cost effective



Whole-Exome Sequencing workflow

NGS

DNA (10-100ng)

Library prep

Whole Exome capture by probes

Capture size: 47Mb
Coverage:
99.8% CCDS,
99.9% ClinVar
99.7% OMIM

Deep Sequencing

Bioinformatics





Variant Calling

GATK

VCF file

Variant annotation

dbSNP, SIFT, PolyPhen2, dbscSNV, TaiwanBio, 1000G, ExAC, gnomAD, ESP, COSMIC, ClinVar, dbNSFP, interVar.

Identification of pathogenic variants for Skeletal dysplasia analysis

Rare inherited disorders

Skeletal dysplasia (SD)

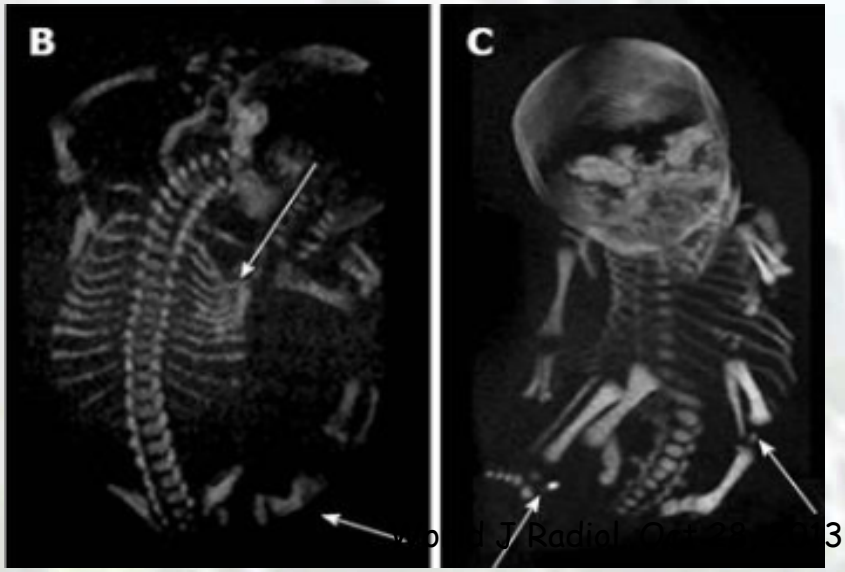
- There are 247 genes involved in skeletal dysplasia.
- The incidence rate: non-lethal SD 1/4,000~5,000
  Lethal SD 1/10,500



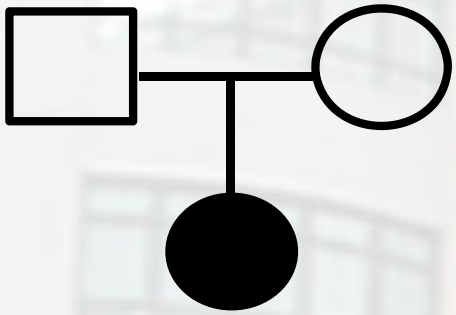


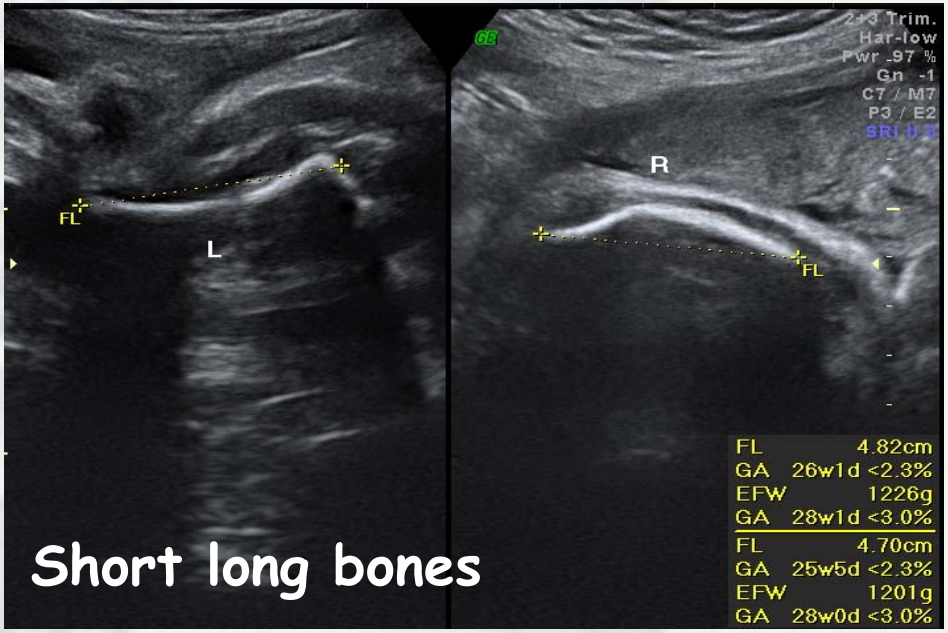






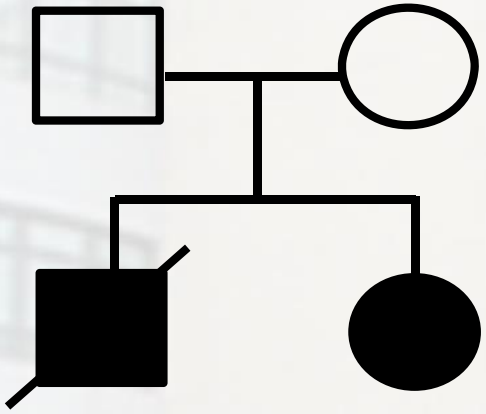
Pedigree of the family with skeletal dysplasia



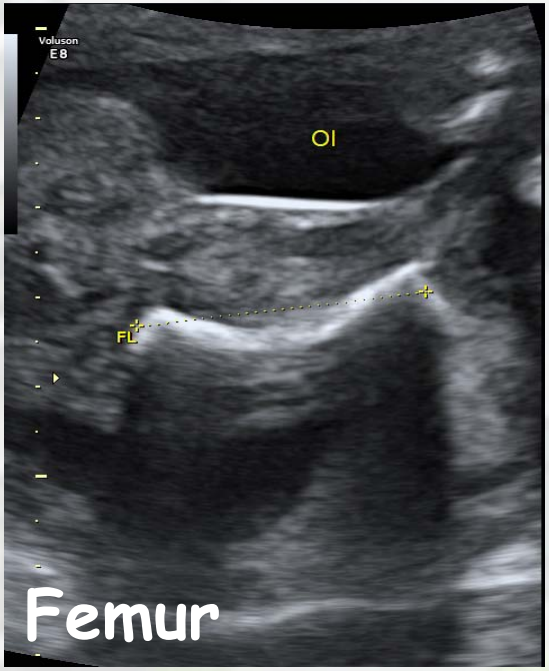


Diagnosed as Ostegenesis imperfecta (OI) after birth

A family of skeletal dysplasia with unknown genetic defect



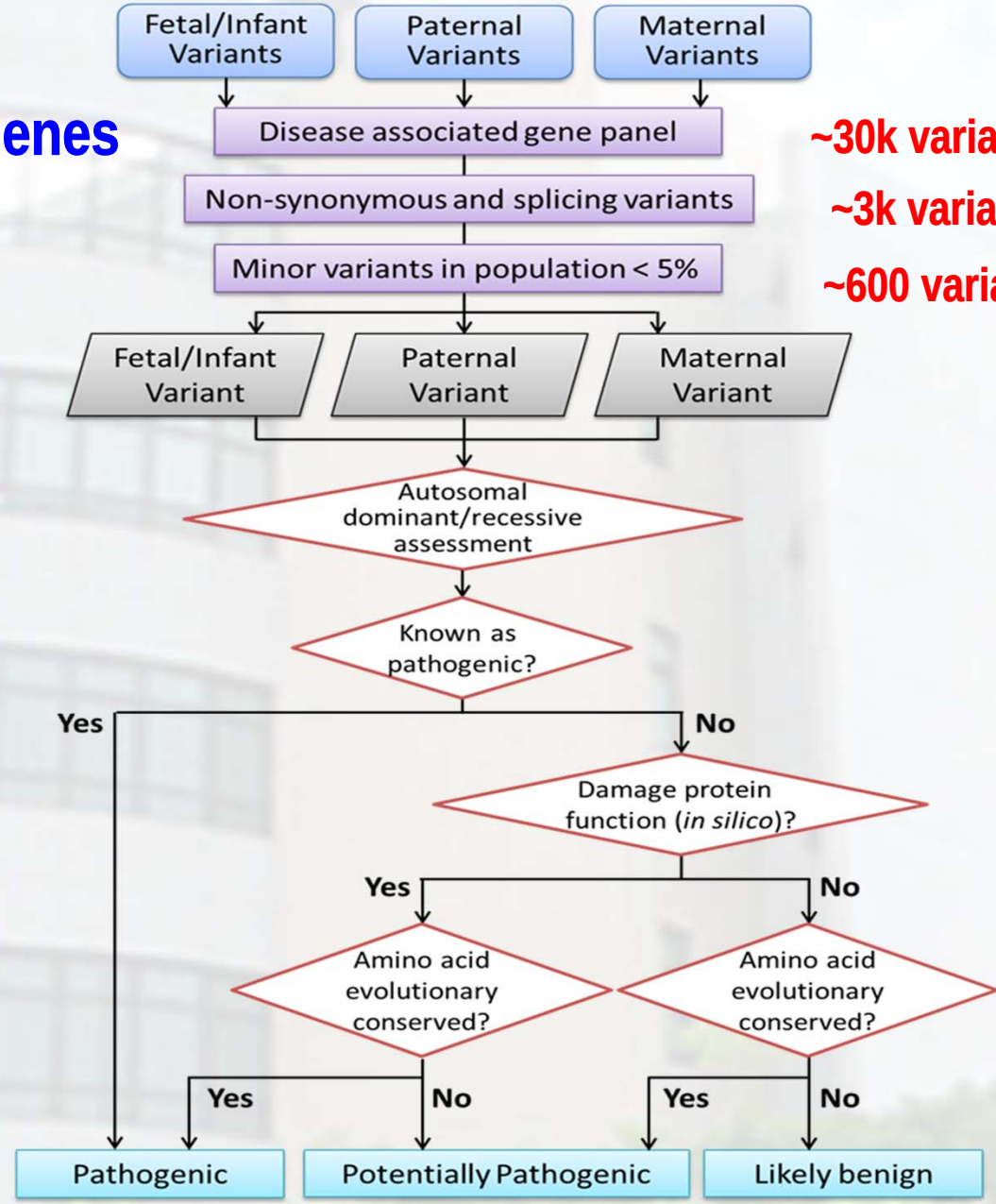
- More severe symptom: short long bones and multiple fractures
- Most common autosomal dominant genes:
  1. COL1A1
  2. COL1A2
  3. COL2A1
  4. FGFR2
  5. FGFR3





Pathogenic variants identification

~4100 genes



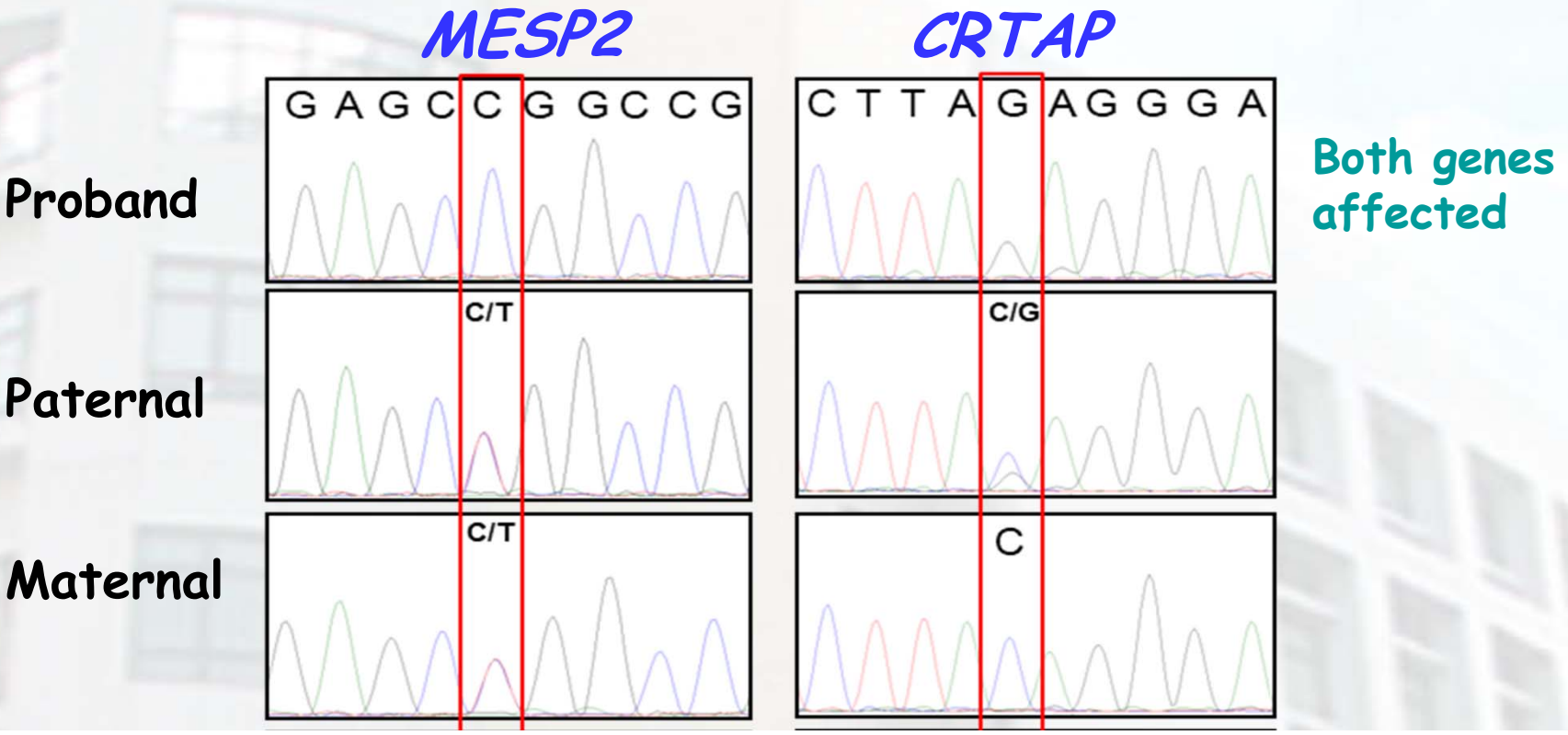
Candidate genes/variants involved in skeletal dysplasia

Gene	Chr:Pos	Variant	Inherited model*	Genotype Umbilical DNA&	Genotype Parental DNA	Consequence	Population AF (TWB)	SIFT$	PolyPhen$
MESP2	15:90320496	T > C	AR	hom	♂: het ♀: het	Missense Variant	T: 0.98 C: 0.02 (0.034)	deleterious (0.02)	probably damaging (0.96)
CRTAP	3:3318 3884	C > G	AR	hom	♂: het ♀: wt (?)	Splice region variant# (intron)	C: 0.9999 G: <0.001 (0.001)		

- **Mutation of MESP2 causes spondylocostal dysostosis (SCDO) in an autosomal recessive (AR) manner.** (Whittock et al., 2004; Takahashi et al., 2007; Cornier et al.,2008; Makino et al., 2013)

- **Mutation of CRTAP causes Osteogenesis Imperfecta (OI) in an AR manner.** (Amor et al., 2011; Valli et al.,2012; Barbirato et al., 2015; Costantini et al., 2017)

Validation of the mutations

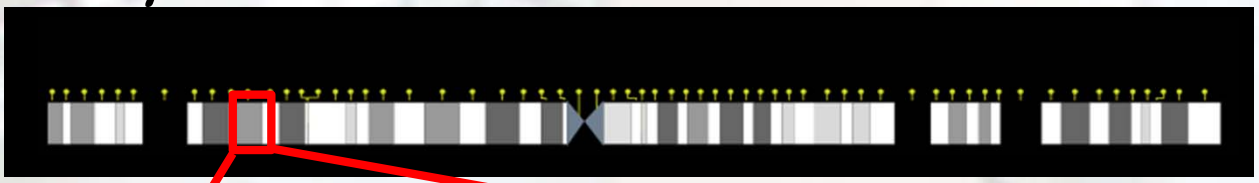


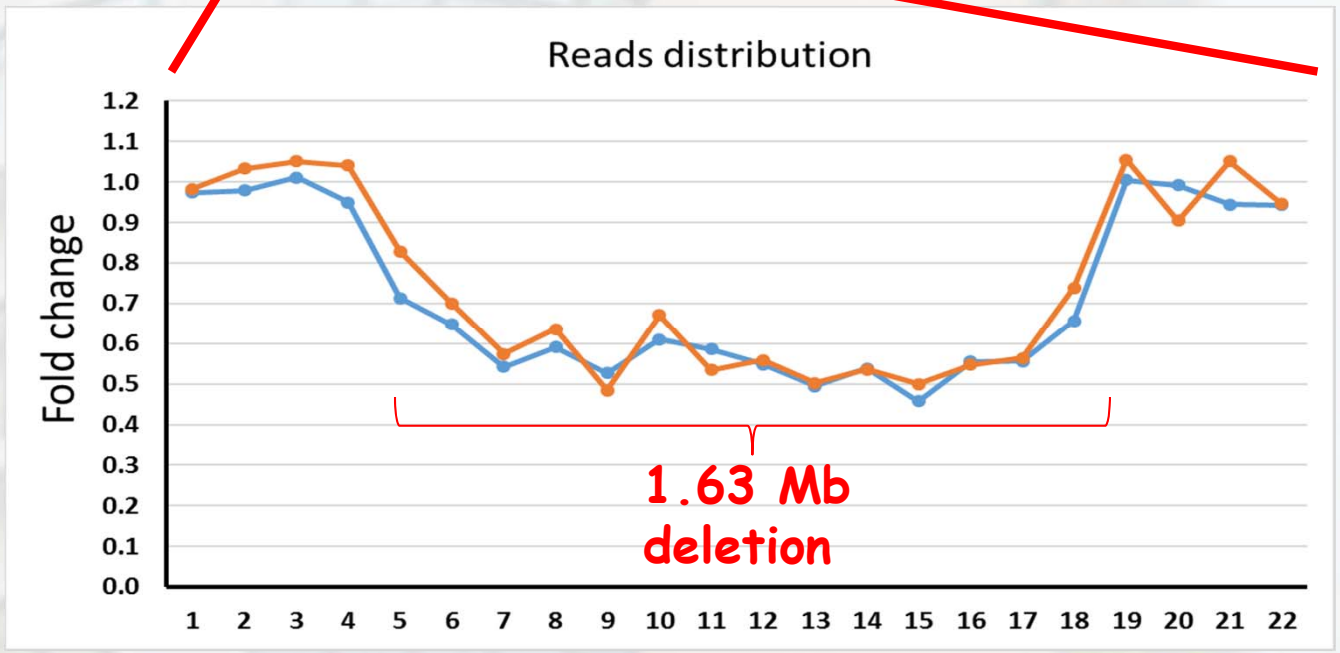
What happens to the maternal allele of CRTAP locus?

Is there a null allele in maternal CRTAP locus?

Reduced read-count suggests CRTAP deletion

Chr3q22.3





Normalized read count from WES data

Maternal/Paternal

Proband/Paternal

Genes in order:

05: CMTM8

13: CRTAP

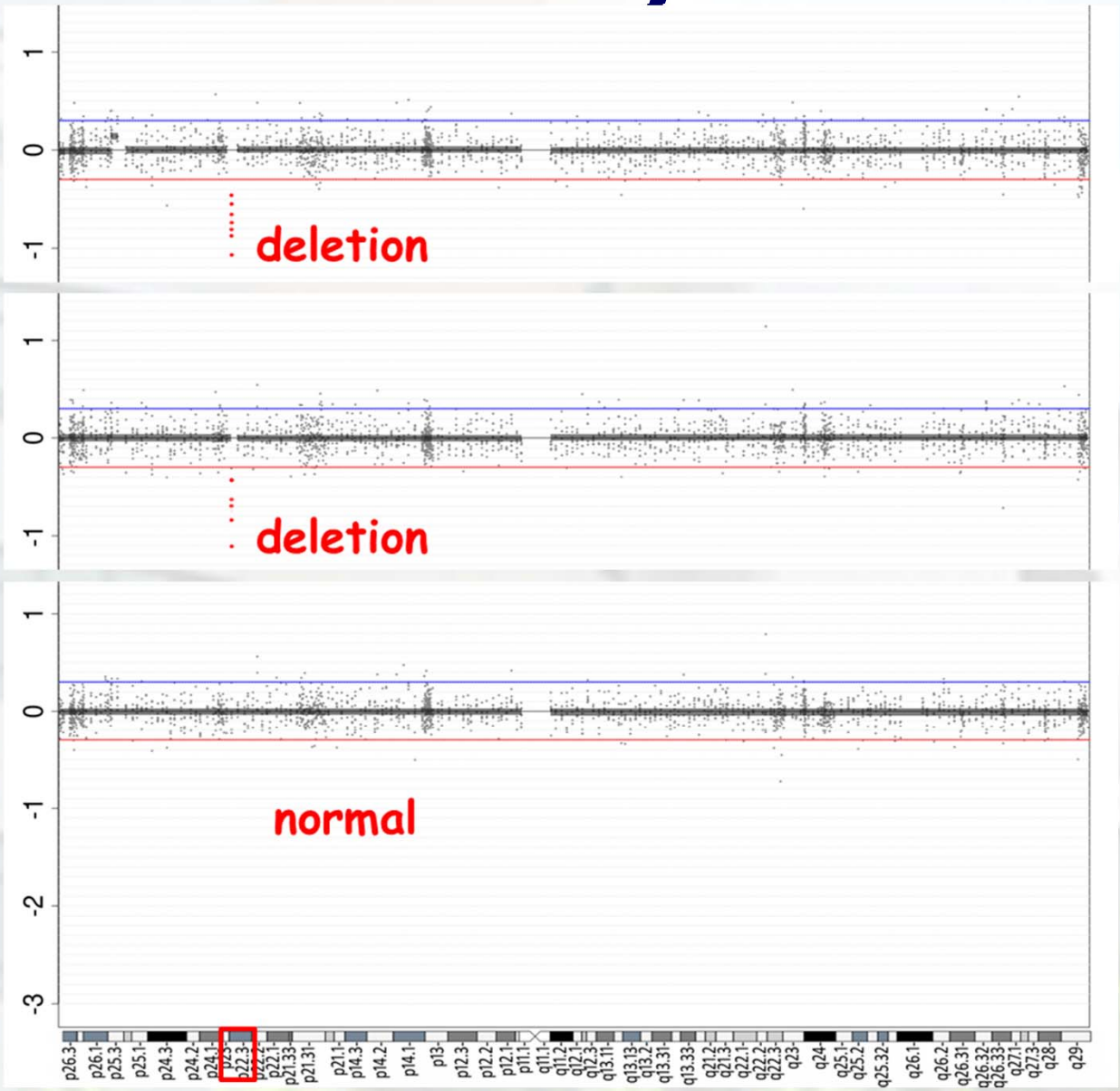
18: PDCD6IP

Regional deletion on chr3q22.3 is confirmed by aCGM

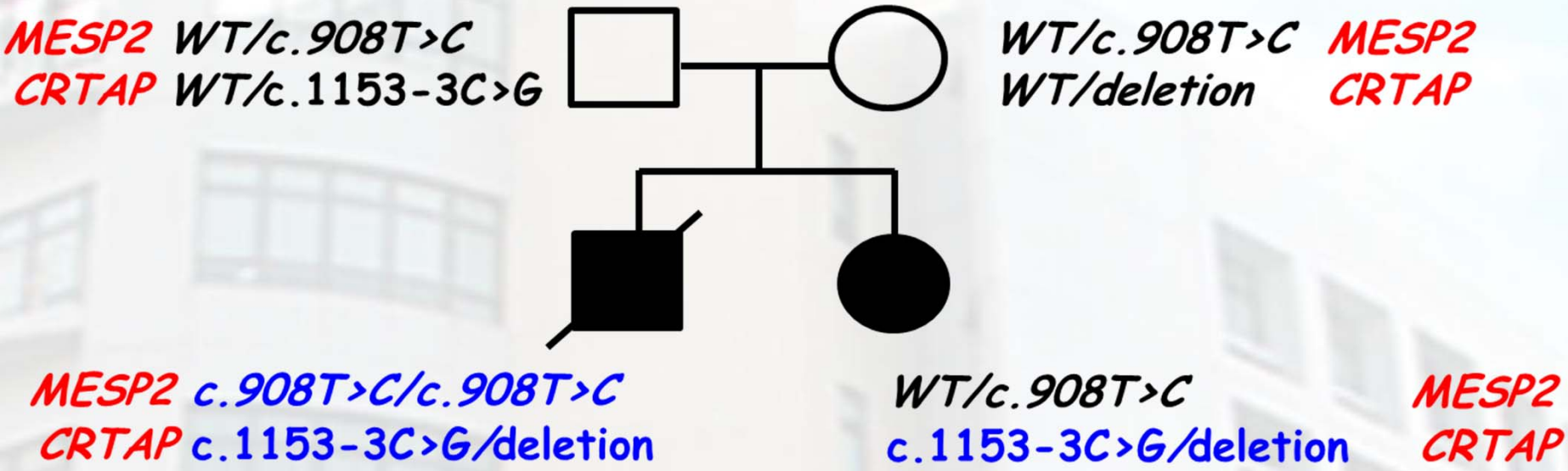
Proband DNA

Maternal DNA

Paternal DNA



Summary of mutations detected in the family with skeletal dysplasia



What are the functional consequences of these mutations?

Functional consequence of MESP2 mutation (p.Leu303Pro)

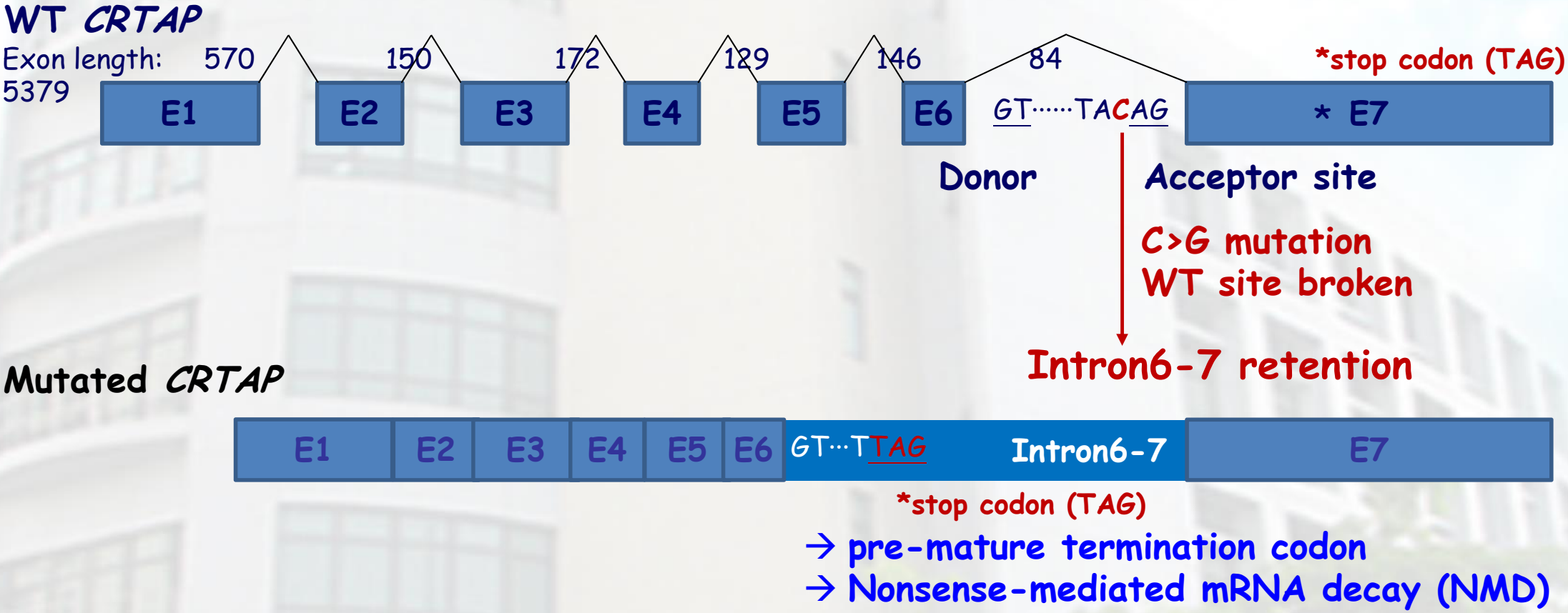
- Highly conserved amino acid by PhastCons (0.841) and by PhyloP (0.935).

Multiple Alignment of different species (UCSC)



- Deleterious (0.02) by SIFT, and probably damaging (0.96) by PolyPhen.
- Probably deleterious by MutationTaster2

Predicted effects of *CRTAP* mutations

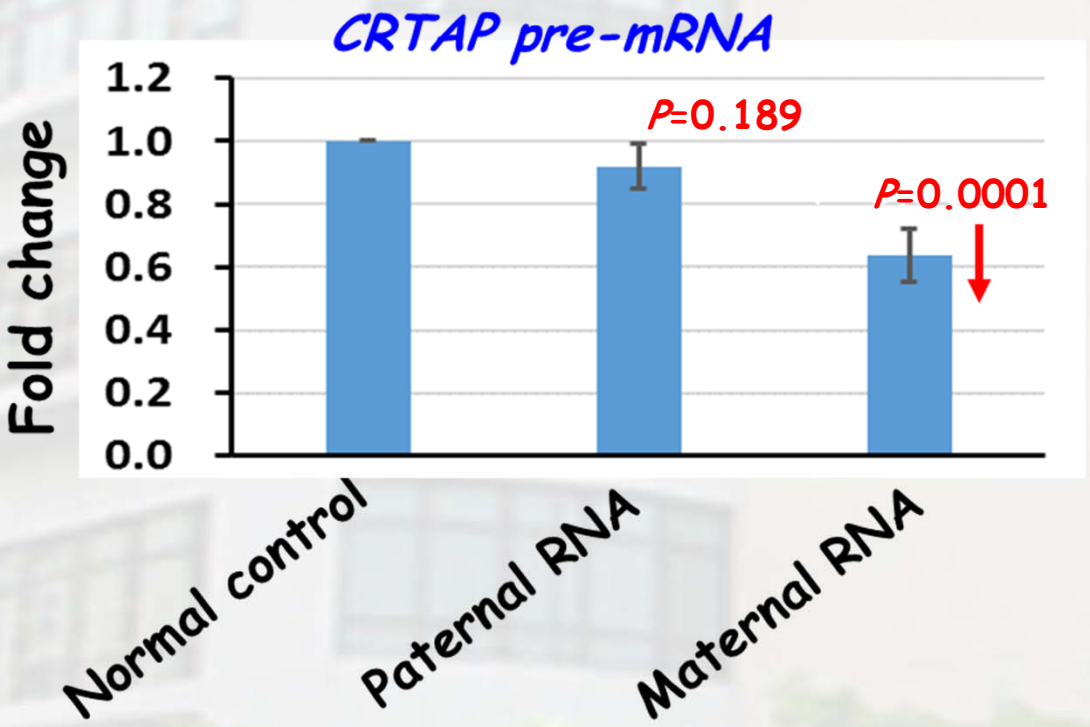


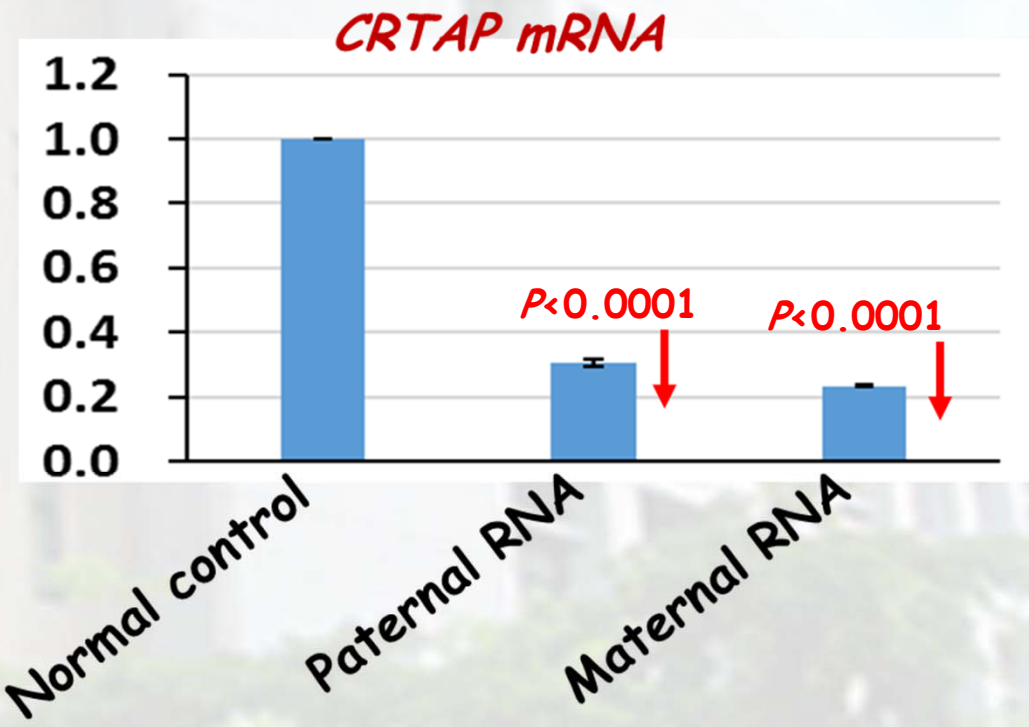
Functional consequences of *CRTAP* mutations



qRT-PCR

Primer for mature mRNA





Conclusion

- WES-based genetic screening workflow successfully identify mutations in a family with Skeletal dysplasia.
- The mutation identified in the MESP2 gene has population frequency of 0.02-0.03. The fact points out the needs for "carrier screening" to avoid the transmission of "not so rare" mutations to the offspring.
- To take one step further, an AI-assisted, comprehensive workflow is on the way to provide next generation prenatal genetic screening.

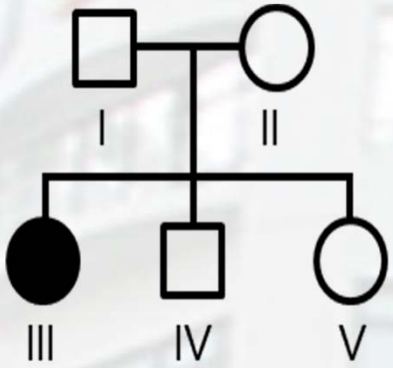
CRTAP (Chr3:33183884)

Type	Coverage	Variant	Genotype	DP	AD (REF, ALT)	AF
gDNA	*~100x*	C > G	Hom	17	0,17	1.00
cfDNA	*~500x*	C > G	Het	59	51,8	0.14

Identification of pathogenic variants for Autism

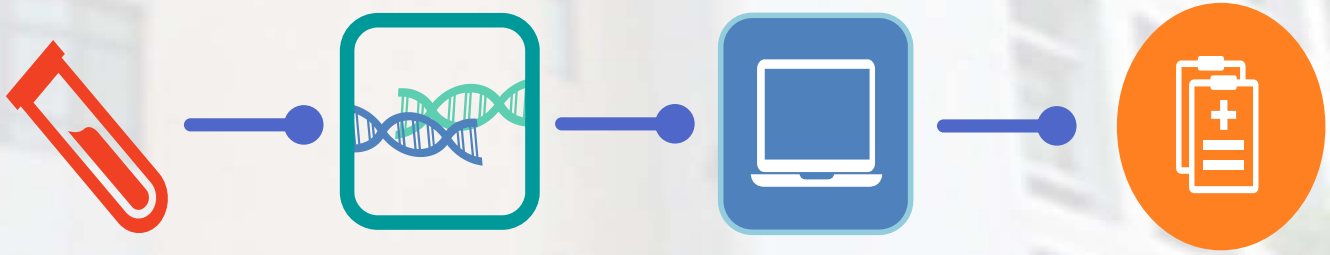
Complex disorders

Complex disease: Autism



Autism
Intellectual disability

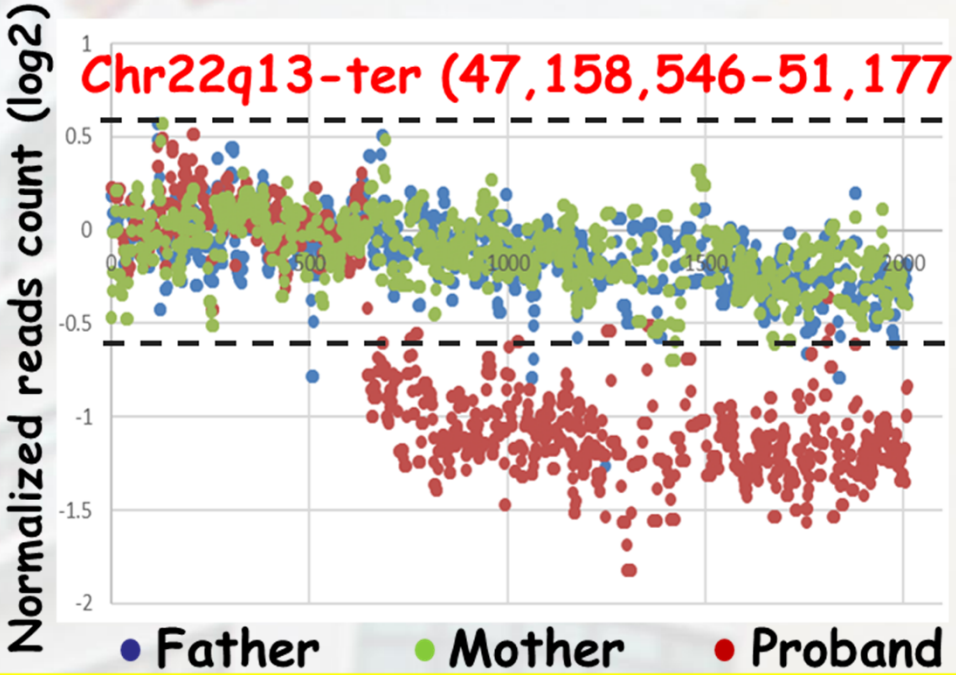
Autism spectrum disorder (ASD) is a complex, neuropsychiatric disease. More than 100 genes and genomic regions have been associated with ASD. Large copy number variants (CNVs) and de novo mutations (DNMs) in genes account for only 10%–30% of simplex autism cases.
(Turner et al., *Cell*, 2017; Krumm et al, *Nat Genet*, 2015; Wurst, *PLoS Genet*, 2016)

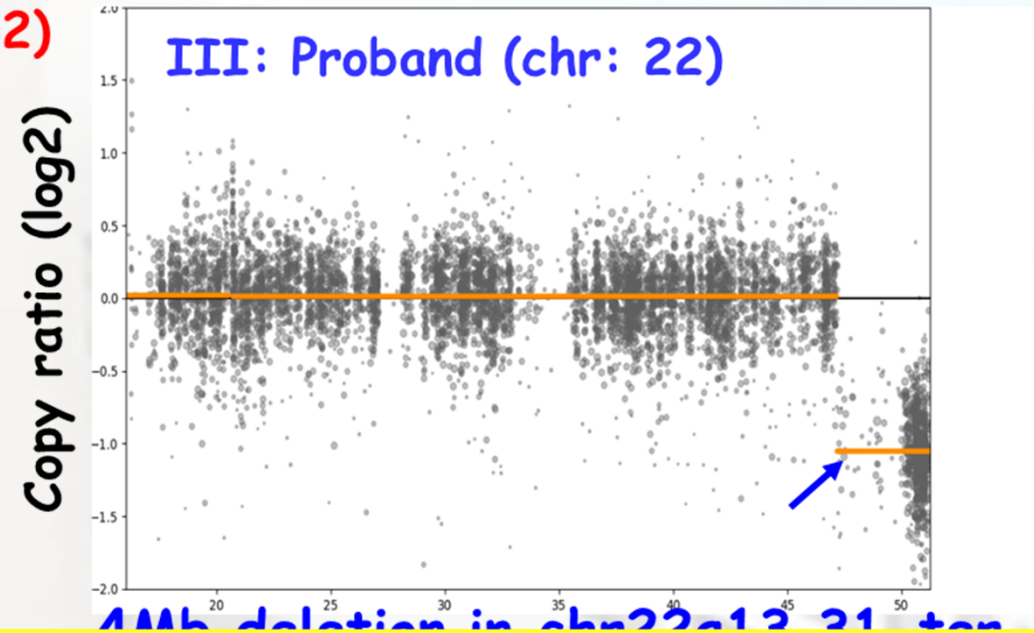


Trio: I, II, III WES Variant Calling Discovery

→ There's no phenotype-relevant SNV found in the proband.

Identification of copy number variation (CNV) in the proband





WES platform works very well → → PGT-M

Phelan-McDermid syndrome (22q13.3 Deletion Syndrome, AD)

Common features: neonatal hypotonia, global developmental delay, normal to accelerated growth, absent to severely delayed speech, autistic behavior, and minor dysmorphic features (OMIM #606232).

Conclusion

We provide a total solution for the prenatal care with high quality

PGT-M

WES

Pathogenic variant discovery

PGT-A

Survival Index (SI) prediction



Healthy Baby

Better Life--Preventive Medicine



Thank you for your attention

