Can genomic medicine detect early…

What question do you want to ask?

The question

- Discovery question: Identify variants associated with a disease in a pilot study
- Discovery of variants associated with facial dysmorphia - clinical outcome?
- What happens if you are successful?
- Therapeutic target discovery in low-grade serous ovarian cancer (LGOSC)

Technology approach

- What machine is generating the data?
- Potential impact on format, quality, and management

- Budget?

Methods for analysis - Pipeline

- Input data ? selection of cases, case controls? Numbers? Exons scan? Whole Genome? Panel?
- How to do alignment and calling? (Family Trios?- how to combine data?)
- Annotation, Filtering, Prioritization
- Pathogenic? Clinically actionable? tool list choice / understanding

Data that will be used

- Source – generated? Private? Public?
- File Formats
- Where to store?
- Access and security policy?

Results and clinical outcome

- Describe what is expected
- How will it be reported? Added to databases?
- What if you are successful? What will you do with the discovery/data? Further projects
- Translation into NHS practice?

Clinical calling standards and best practices

- What standards for classifying variants?
- How to ensure:
  - Results are reliable?
  - Reproducible?

Grading

Presentation:
Technology / approach for sequencing – 15%
Choice of analysis methods – 20%
Dataset – 20%
How results will be interpreted –20%
Adherence to standards and best practices – 15%

Response to questions – 5%
Asking questions (participation) - 5%

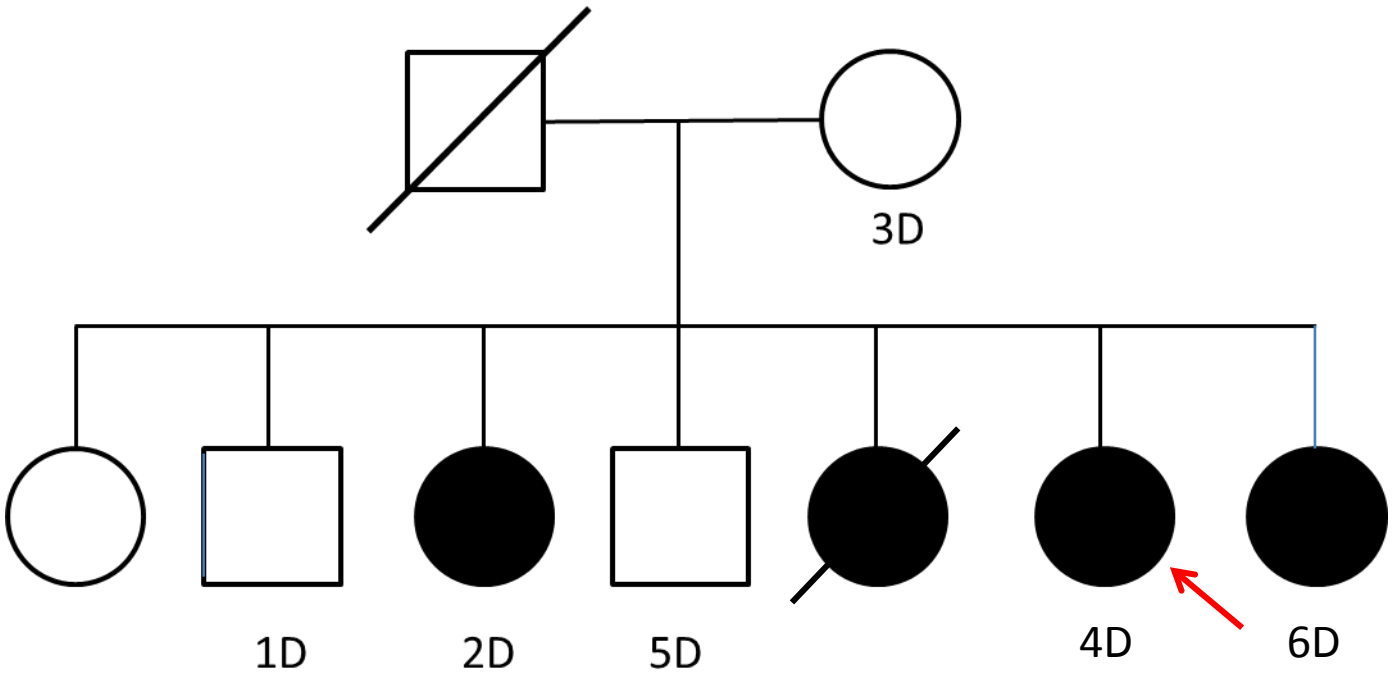
Write-up:
Does the proposal address an important problem and, if successful, will the results have an impact on have a sustainable impact on Genome Medicine in the UK? - 10%
Is the choice of technology appropriate and well reasoned? - 10%
Is the project focused and achievable? - 10%
Are the methods and approach well reasoned and achievable? - 50%
Are the measures of success appropriate? - 10%
Does the project recognise and utilise Genome England resources? - 10%

Establishing the causative gene for BVVL in an Indian Family

Janine Kirby

Brown Vialetto Van Laare

- Rare neurological disorder
  - Progressive pontobulbar palsy
  - Sensorineural hearing loss

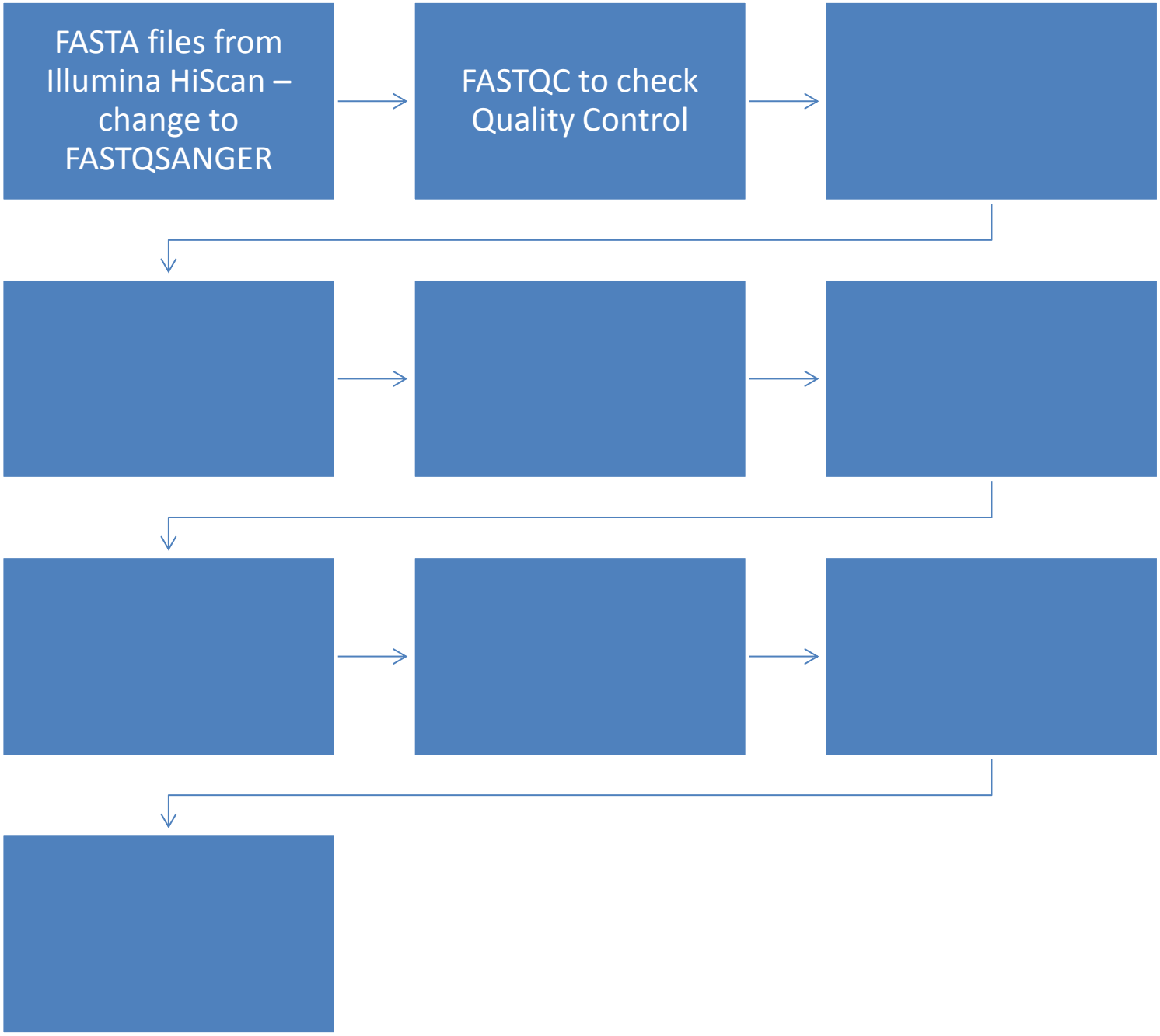


ASSUMPTIONS:

AR inheritance

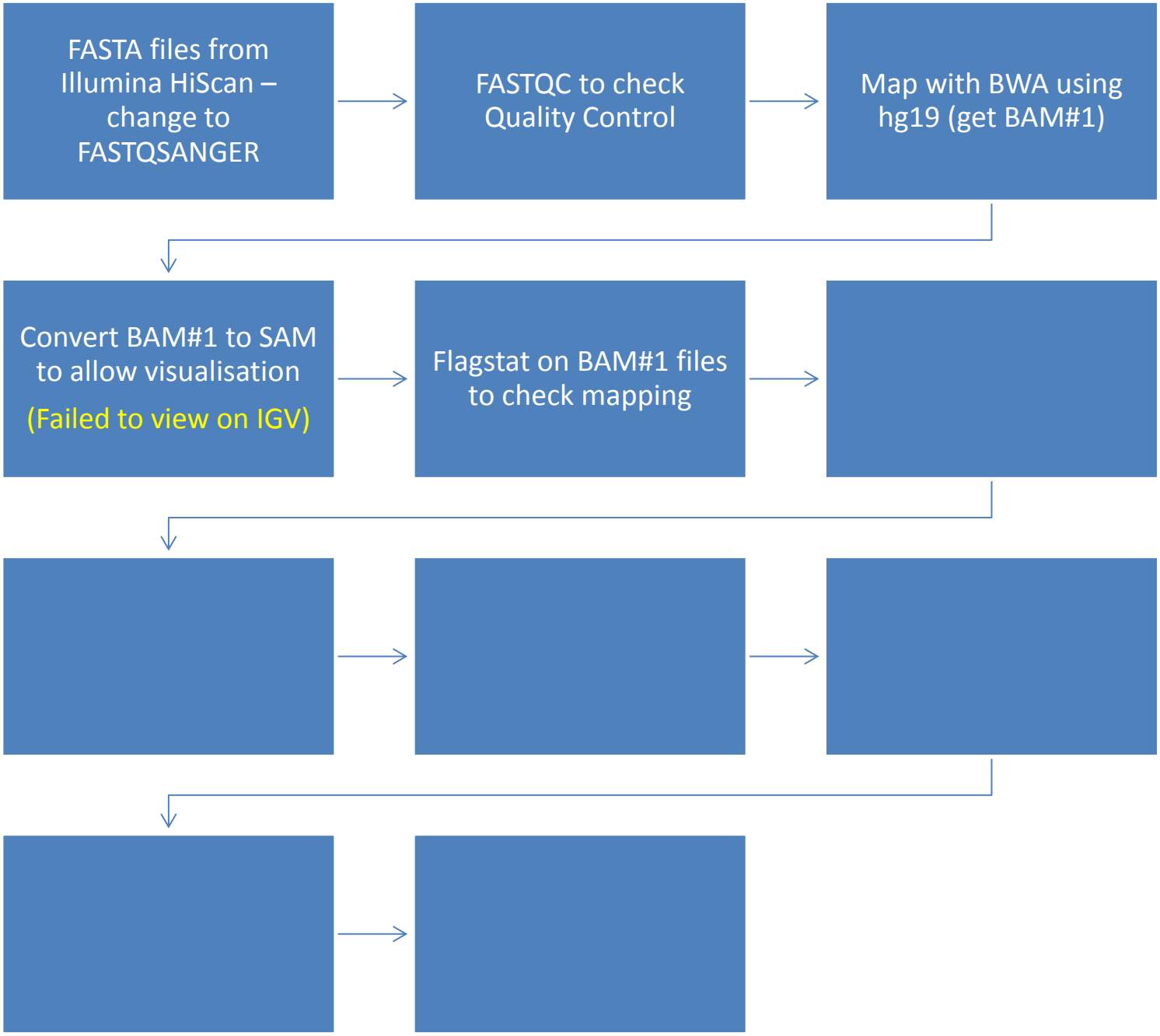
(X-linked?)

Unaffecteds are truly unaffected


FASTA files from Illumina HiScan – change to FASTQSANGER
FASTQC to check Quality Control

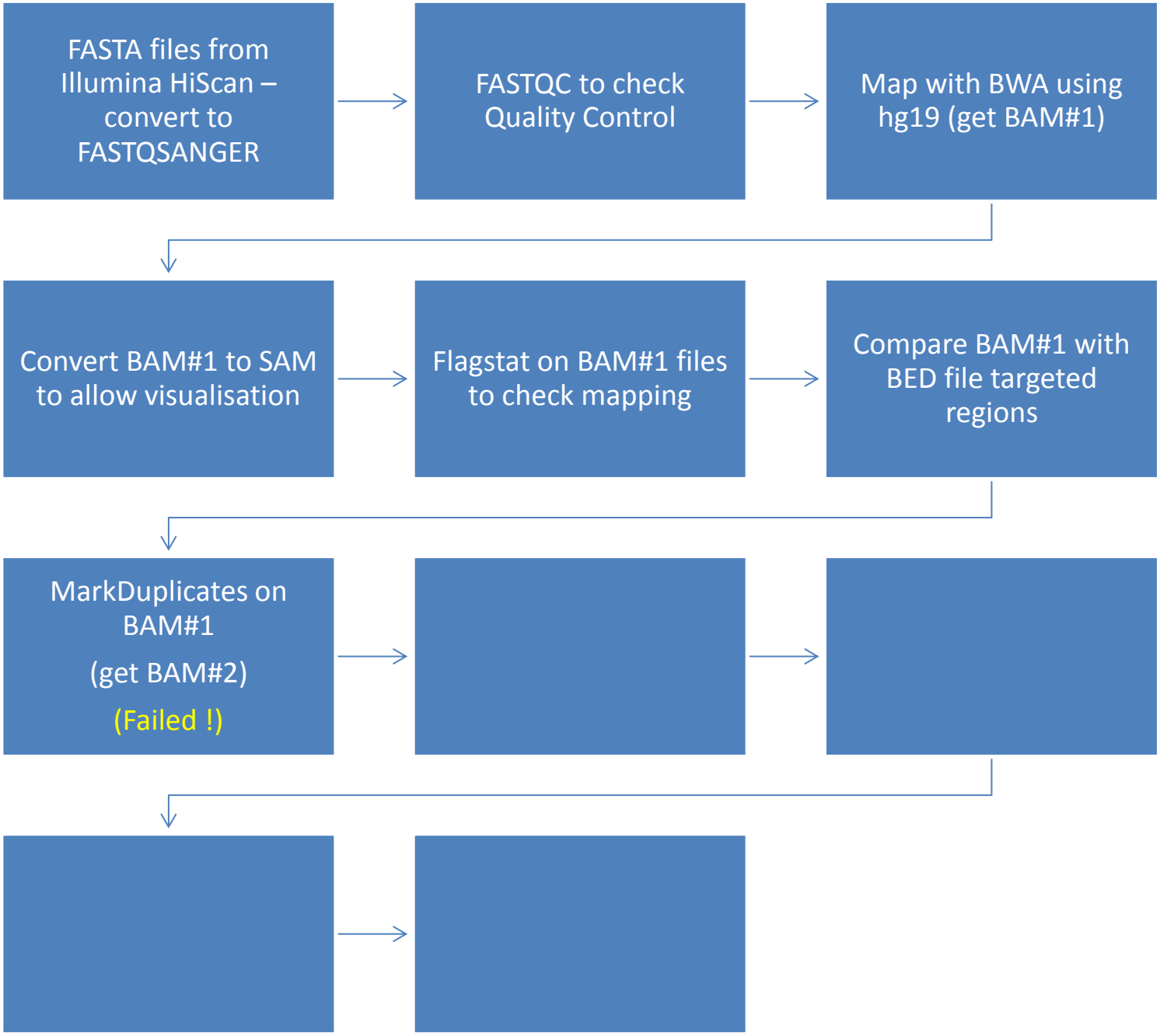
FASTQC Results

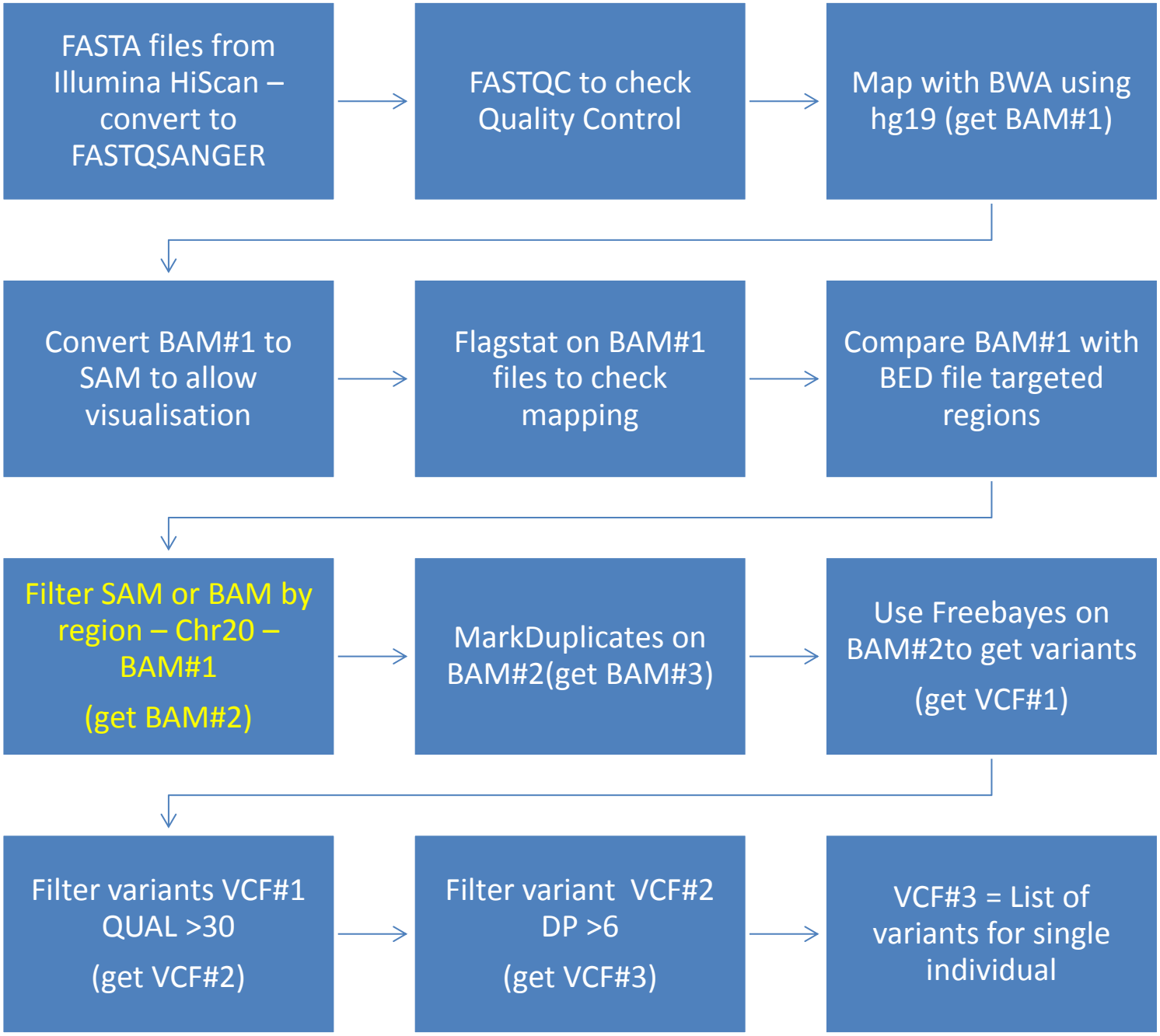
		FASTQC	
		Seq length	No Seqs
1D	R1	93	14112175
	R2	93	14112175
2D	R1	93	8031424
	R2	93	8031424
3D	R1	93	13281840
	R2	93	13281840
4D	R1	93	14794396
	R2	93	14794396
5D	R1	93	15354790
	R2	93	15354790
6D	R1	93	15949664
	R2	93	15949664


FASTA files from Illumina HiScan – change to FASTQSANGER
FASTQC to check Quality Control
Map with BWA using hg19 (get BAM#1)
Convert BAM#1 to SAM to allow visualisation
(Failed to view on IGV)
Flagstat on BAM#1 files to check mapping

Flagstat QC

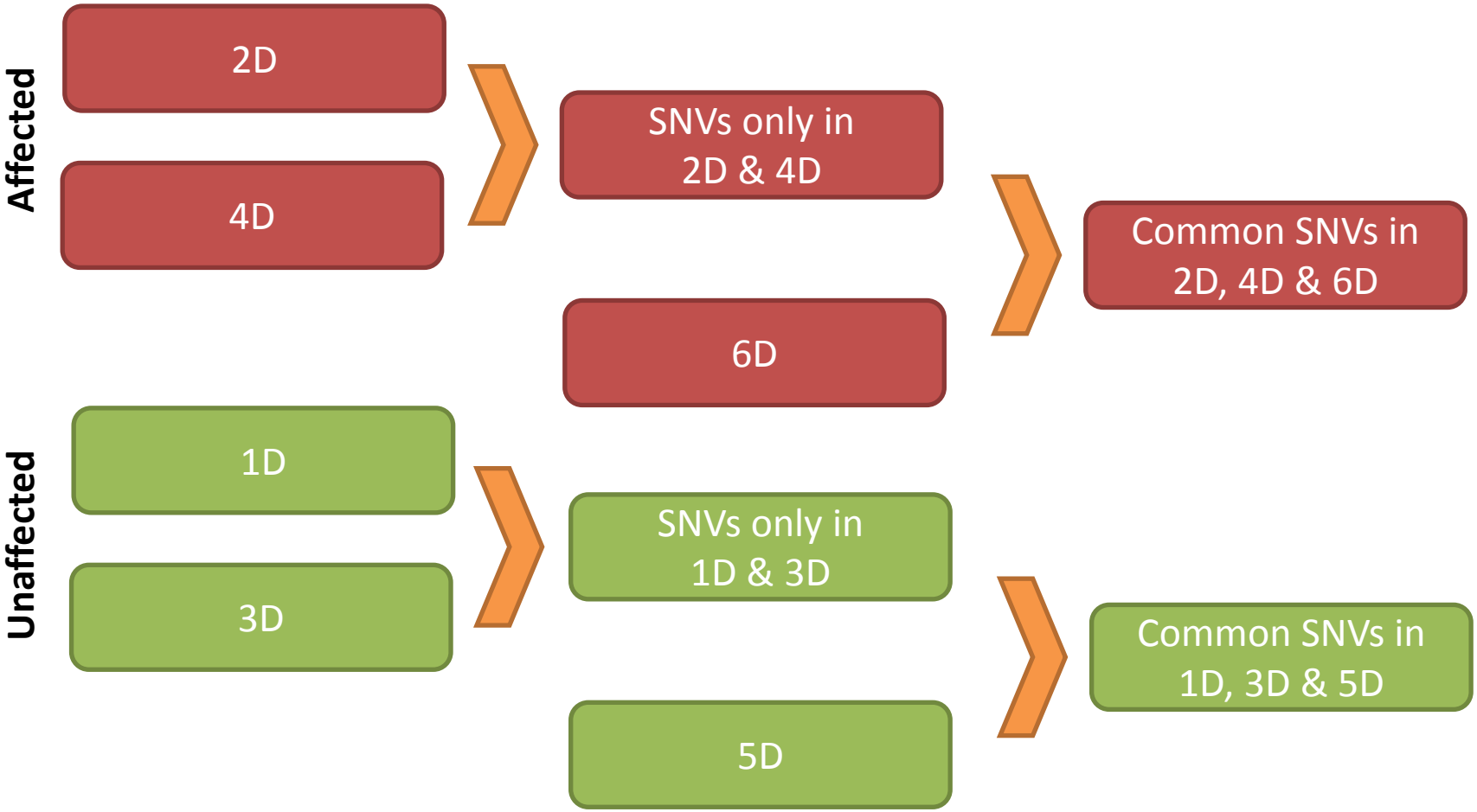
	Initial QC	
	Mapped	Paired
1D	86.78%	82.95%
2D	90.56%	87.33%
3D	88.04%	84.21%
4D	83.59%	78.76%
5D	85.52%	81.13%
6D	84.15%	79.48%


FASTA files from Illumina HiScan – convert to FASTQSANGER
FASTQC to check Quality Control
Map with BWA using hg19 (get BAM#1)
Convert BAM#1 to SAM to allow visualisation
Flagstat on BAM#1 files to check mapping
Compare BAM#1 with BED file targeted regions
MarkDuplicates on BAM#1
(get BAM#2)
(Failed !)


FASTA files from Illumina HiScan – convert to FASTQSANGER
FASTQC to check Quality Control
Map with BWA using hg19 (get BAM#1)
Convert BAM#1 to SAM to allow visualisation
Flagstat on BAM#1 files to check mapping
Compare BAM#1 with BED file targeted regions
Filter SAM or BAM by region – Chr20 – BAM#1
(get BAM#2)
MarkDuplicates on BAM#2(get BAM#3)
Use Freebayes on BAM#2to get variants
(get VCF#1)
Filter variants VCF#1 QUAL >30
(get VCF#2)
Filter variant VCF#2 DP >6
(get VCF#3)
VCF#3 = List of variants for single individual

Identifying SNVs

- VCF#3 files from each individual – use Intercept



Identifying SNVs

- Check for SNVs in
  - SLC52A3 (BVVL1) 20p13
  - SLC52A2 (BVVL2) 8q24
  - SLC52A1 17p13.2
  - UBQLN1 9p21.32
  - **&** not homozygous in unaffecteds

Common SNVs in 2D, 4D & 6D

Common SNVs in 1D, 3D & 5D

SNVs in SLC52A3

SNV	SIFT	Poly Phen2	Mutation Taster	CADD phred	dbSNP	MAF	Aff status
exon3:c.A907G:p.I303V	D	D	P	17.07	rs3746802	0.03	Het
exon3:c.C833T:p.T278M	T	P	P	10.57	rs3746803	0.09	Het
exon3:c.C800T:p.P267L	T	B	P	6.76	rs3746803	0.18	Hom
exon2:c.C222G:p.I72M	T	P	D	18.06	rs35655964	0.03	Het

SIFT: **D**amaging/**T**olerated
PolyPhen: Probably **D**amaging/**P**ossibly damaging/benign
Mutation Taster: Disease causing automatic/disease causing/polymorphism/automatic **P**olymorphism
CADD phred: Higher the score – more likely it is pathogenic

Limitations/Considerations

- InDels not aligned
- Inheritance pattern unclear
- Pedigree provided ≠ Case report summary
- Can’t visualise variants in IGV