

Polymorphic Markers

Genetic linkage, repeat-length polymorphisms, and molecular identity

Molecular genetic diagnosis

Dr Abu Suraih Sakhri

MBBS, MD (Pathology)

www.abusuraihsakhri.com

Learning outcomes

- Explain how linked polymorphisms can track a familial pathogenic allele.
- Distinguish SNPs, microsatellites, and minisatellites.
- Interpret schematic PCR genotypes and familial segregation.
- Describe identity testing, post-transplant chimerism, and the related epigenetic assays.

Why indirect genetic markers are useful

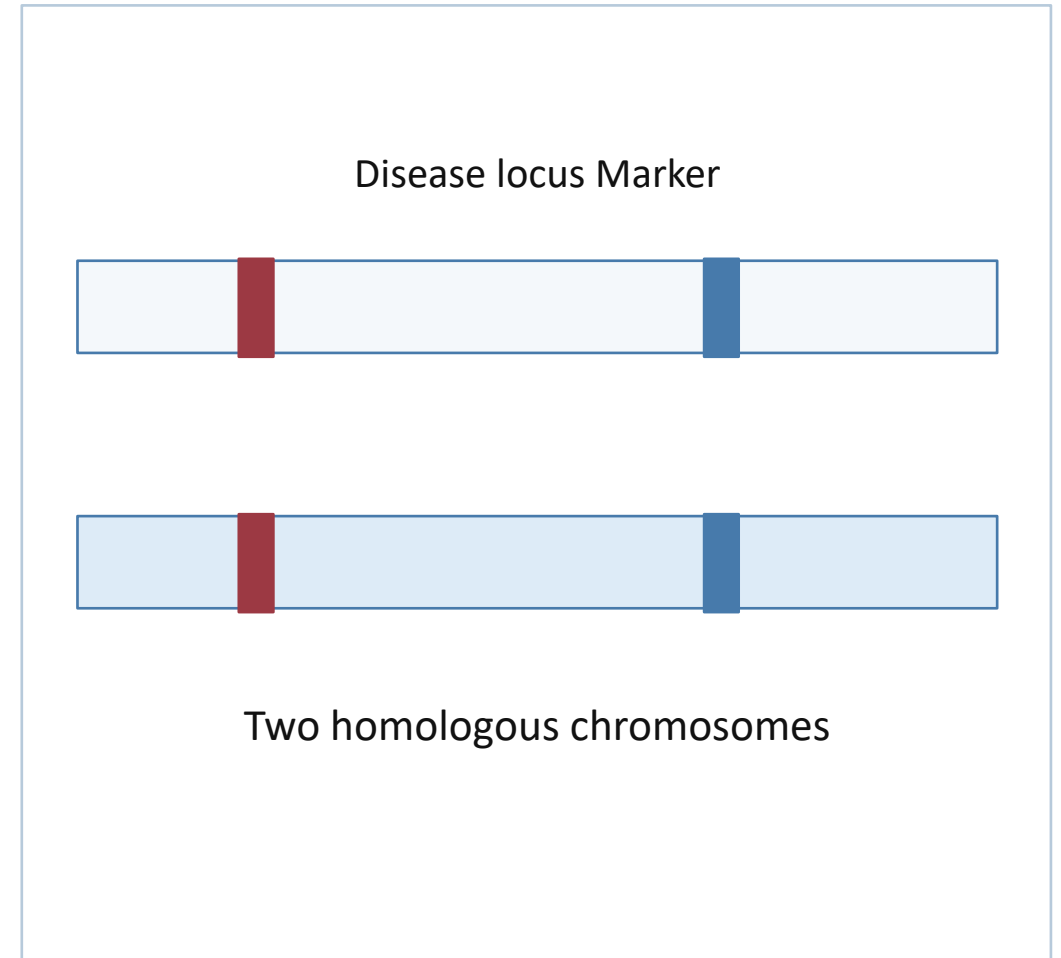
- A disease-specific sequence assay requires the relevant genetic defect to be characterized.
- Some defects are unknown or technically difficult to assay directly.
- Nearby polymorphic loci can serve as surrogate tags for a familial chromosome.
- Tracking a marker provides indirect evidence—not direct identification—of a mutation.

Direct and indirect molecular testing

Direct sequence-based testing	Linkage-marker testing
Interrogates the pathogenic variant itself	Interrogates a neighboring, variable DNA locus
Requires knowledge of the target sequence/defect	Requires an informative familial marker–disease association
Can identify a variant in an individual sample	Usually interprets segregation across relatives
Not contingent on meiotic proximity	Can be uncoupled by meiotic recombination

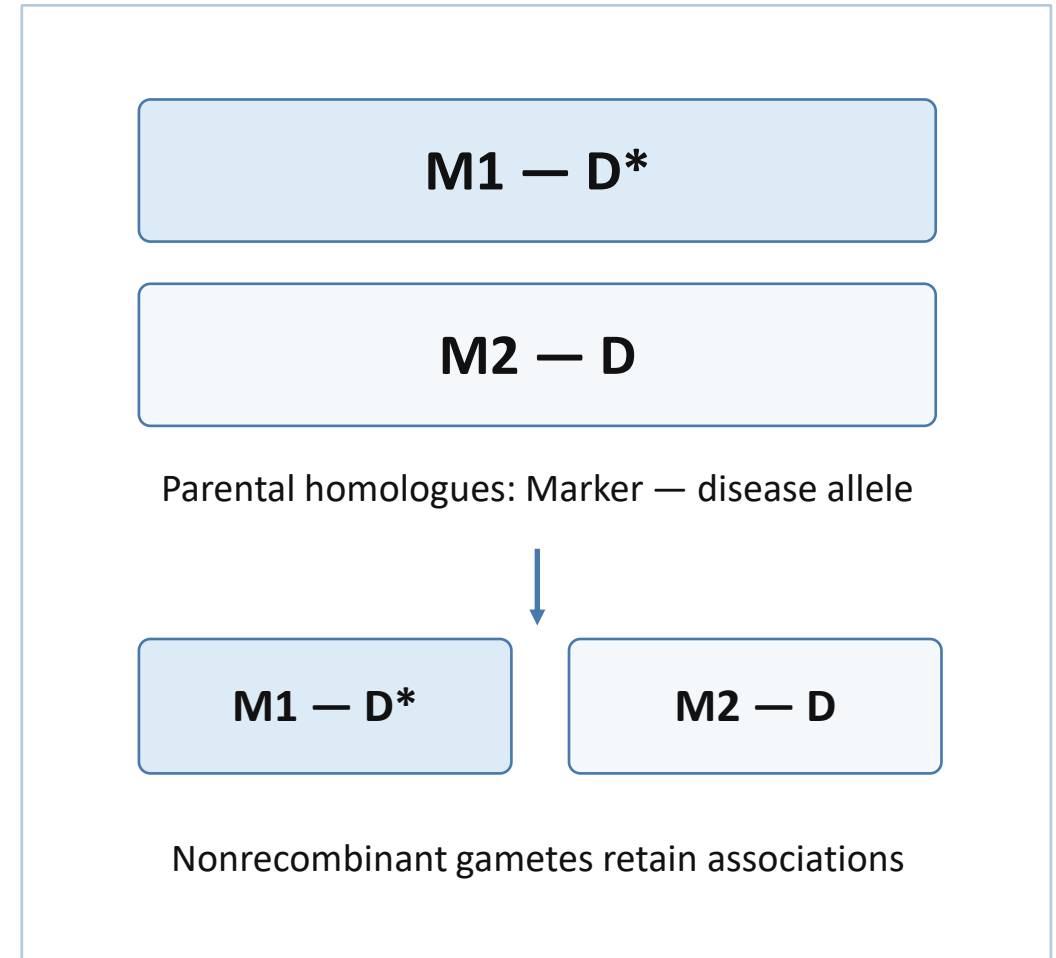
Two loci on the same chromosome

- A locus is a defined genomic position.
- A disease locus and a nearby polymorphic marker may reside on one chromosomal copy.
- The marker allele can identify the chromosome carrying a familial disease allele.



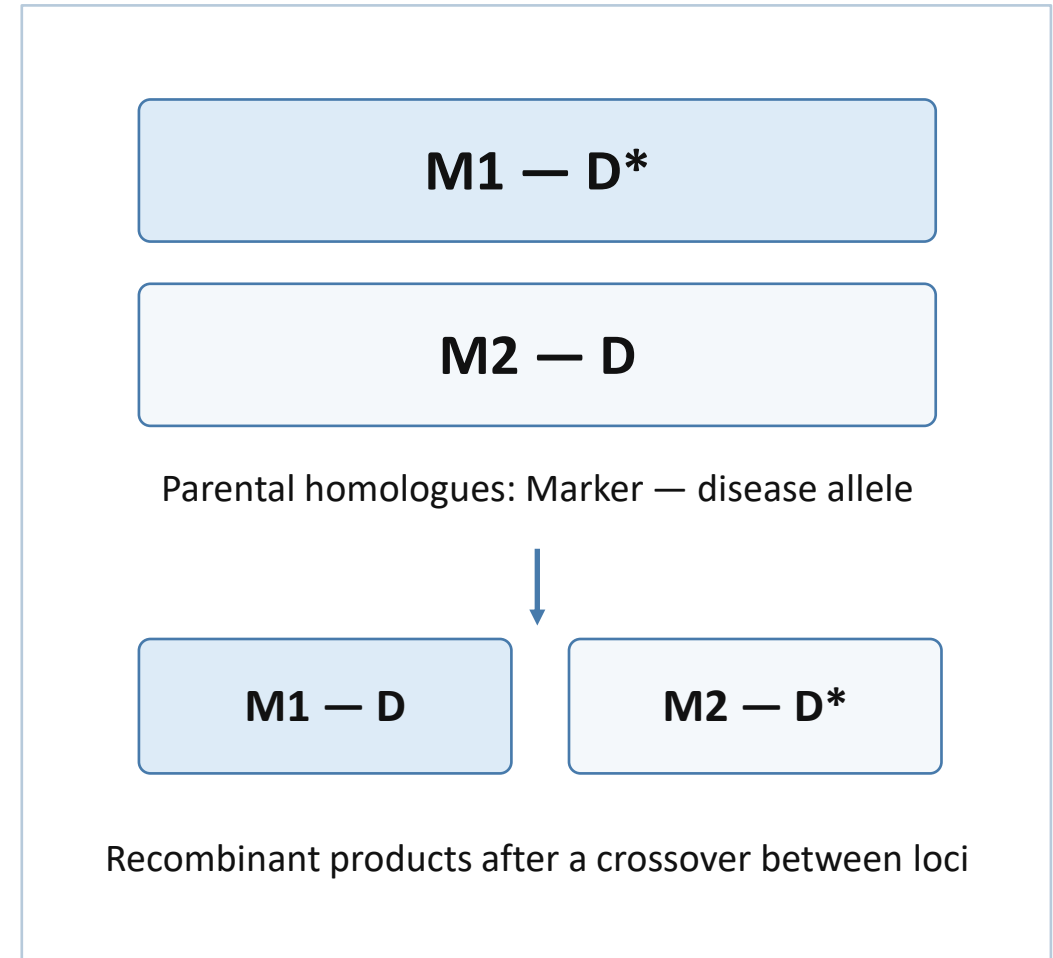
Meiosis creates the linkage opportunity

- A child inherits one parental chromosomal copy at each locus.
- Linked loci tend to be transmitted together in a gamete.
- Observed cosegregation across relatives can reveal the disease-associated marker.



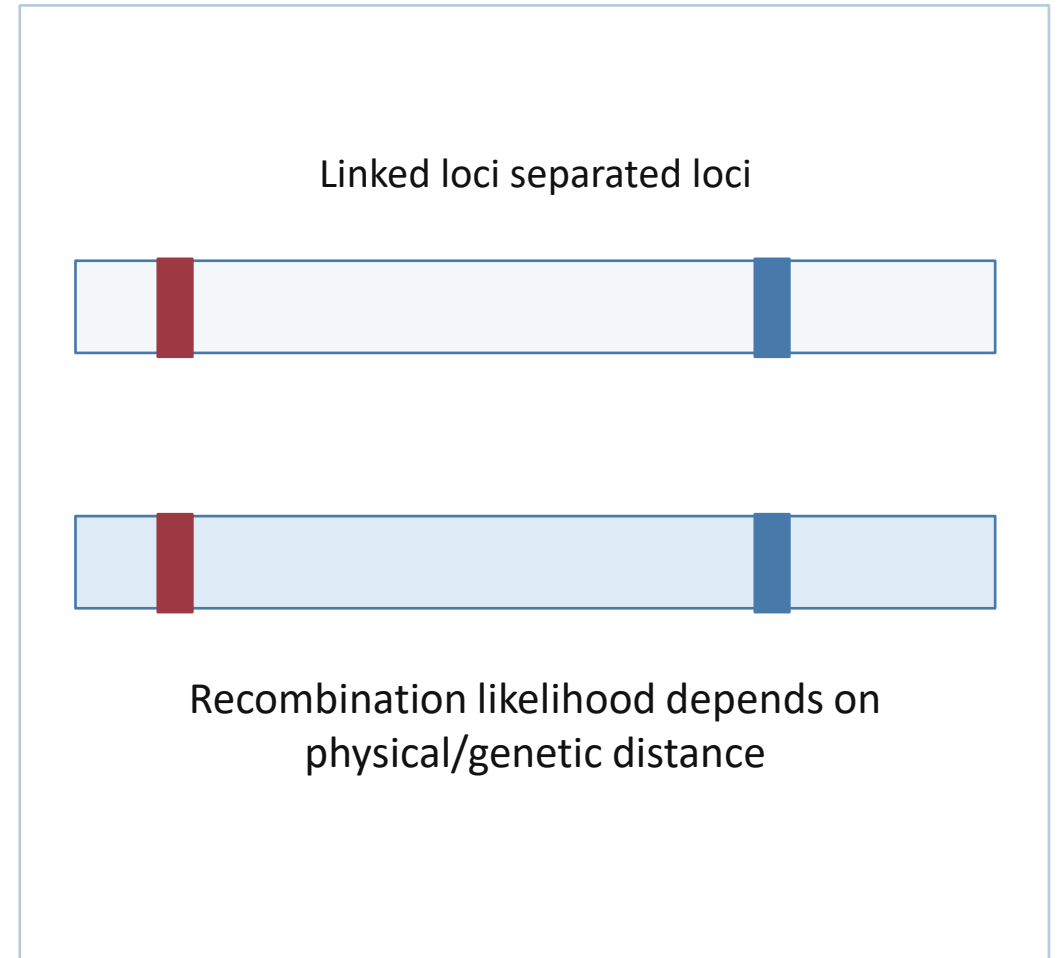
Crossing-over can separate a marker and variant

- Homologous chromosomes may exchange segments in meiosis.
- A crossover between marker and disease locus breaks their original association.
- Linkage is probabilistic; No surrogate marker is perfectly equivalent to the defect.



Distance influences recombination

- Nearby loci are less likely to be separated by a crossover.
- Greater separation generally creates more opportunities for recombination.
- The precise risk depends on genomic location; Proximity is not a universal guarantee.

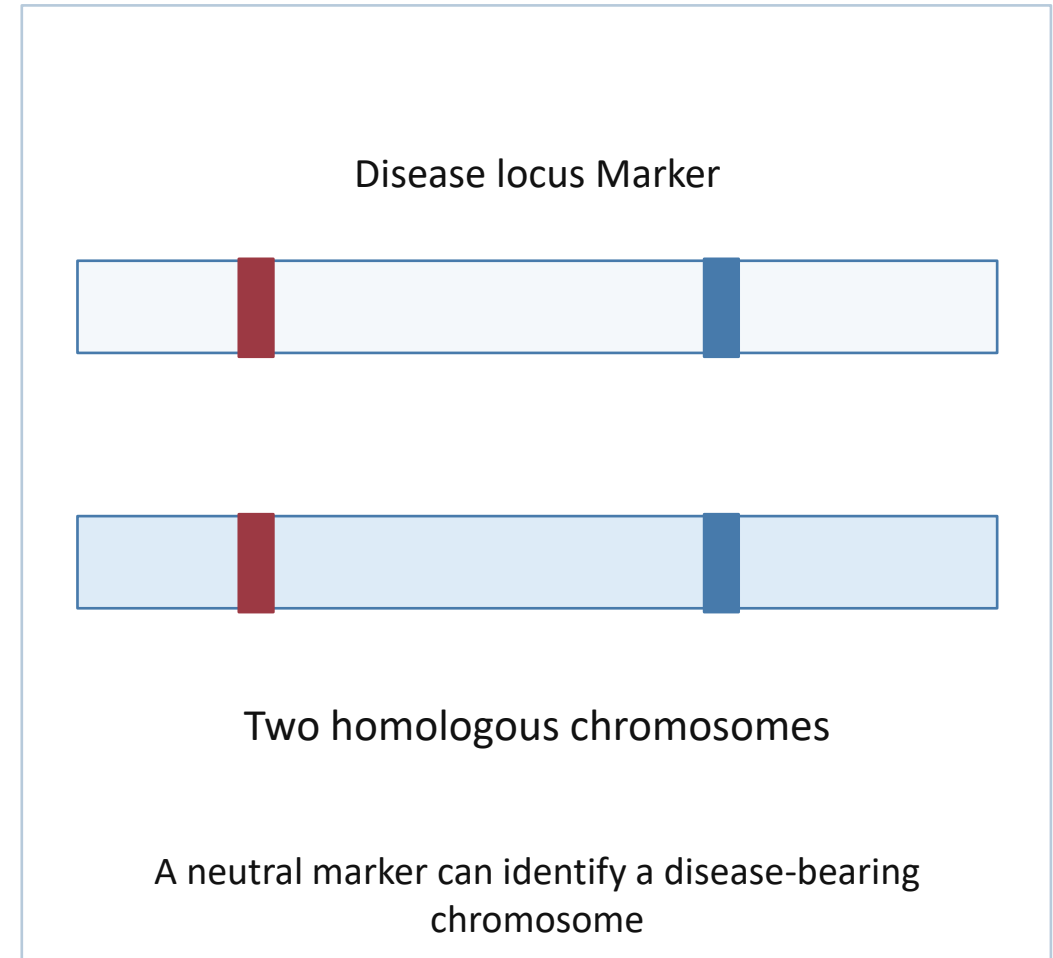


Linkage and linkage disequilibrium

Family-based linkage	Population linkage disequilibrium
Tracks cosegregation in pedigrees	Describes nonrandom association of alleles across a population
Uses informative family members	Can span long haplotype blocks in some regions
Supports inheritance inference	Does not itself demonstrate pathogenicity
Recombination affects both over generations	Population history also affects observed association

A marker tags a chromosome, not a disease

- The marker may be functionally neutral.
- Its value arises from coinheritance with a pathogenic allele.
- The same marker allele in an unrelated family need not identify the same disease chromosome.



Requirements for a useful linkage marker

- The locus must vary among relevant individuals.
- Its alleles must be distinguishable by the chosen assay.
- Marker–disease phase must be determined in the specific pedigree.
- A sufficiently close marker reduces—but does not abolish—recombination uncertainty.

Haplotype: Alleles inherited together

- A haplotype denotes alleles residing on one homologous chromosome.
- Marker genotypes alone do not always reveal which alleles are in cis.
- Family segregation can establish phase and identify the disease-linked haplotype.

Maternal homologue

Marker 1: A — Marker 2: C

Marker 1: B — Marker 2: D

Each row is a haplotype; Unphased genotypes
conceal the pairing

Main classes of polymorphic markers

Two complementary sources of genetic variation

Single nucleotide polymorphisms (SNPs)

Microsatellite repeats

Minisatellite / variable-number tandem repeats (VNTRs)

SNP: Variation at one base position

- A single nucleotide polymorphism changes one nucleotide at a defined genomic site.
- SNPs are commonly biallelic.
- Individuals may be homozygous or heterozygous at an informative SNP site.

... G C A T G ...

... G C G T G ...

One genomic position differs: A versus G

SNP genotypes in a diploid sample

- A two-allele SNP can yield AA, AB, or BB genotype calls.
- Heterozygosity can distinguish maternal and paternal homologues.
- A homozygous SNP provides less allele-level distinction within that individual.

AA

AB

BB

Homozygous Heterozygous Homozygous

Three illustrative diploid genotypes

SNPs occur across the genome

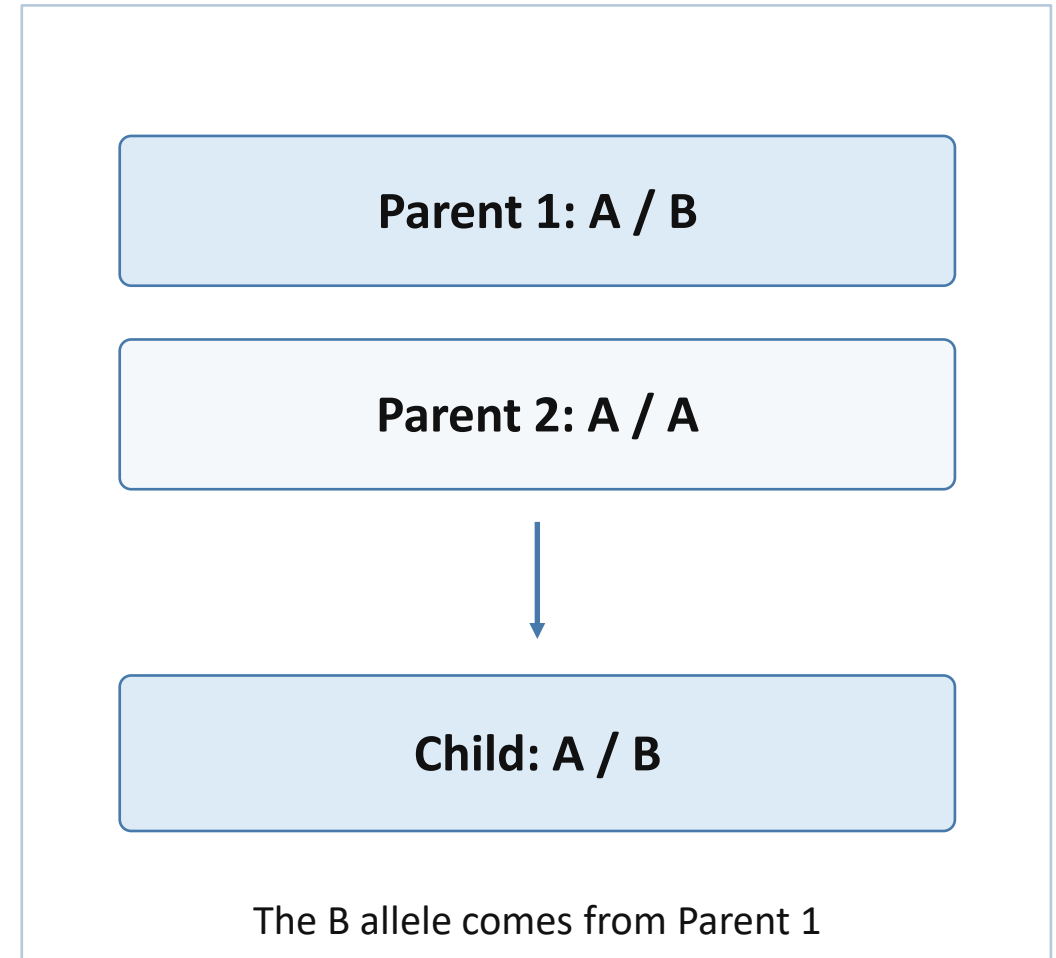
- SNP sites can occur in coding exons, introns, and intergenic DNA.
- Some alter protein sequence or gene regulation.
- Other SNPs appear phenotypically neutral.
- A neutral SNP may still be a highly useful linkage marker.

Biologic effect differs from marker utility

SNP example	Pathologic interpretation
Coding SNP with functional effect	Could contribute directly to altered protein function
Regulatory SNP	May influence expression of a nearby gene
Neutral SNP linked to a variant	May track a disease-associated chromosome without causing disease
Common unlinked SNP	Usually not useful for family-specific tracking of the defect

Tracking one informative SNP

- Assume a parent is A/B at a marker while the other parent is A/A.
- A child with A/B has inherited a B allele from the first parent.
- Whether B indicates disease risk depends on the established family-specific phase.

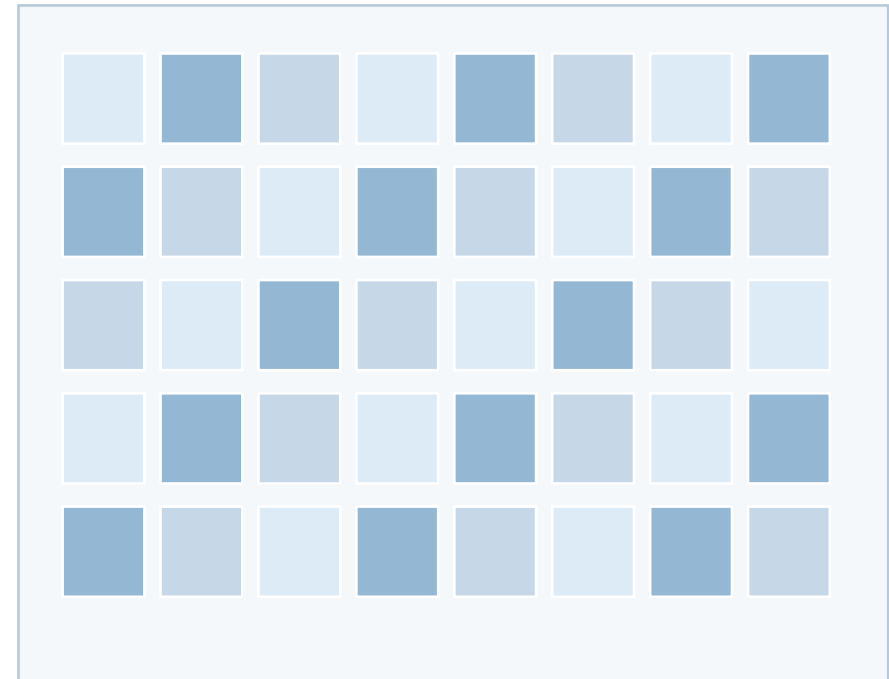


Genotype and haplotype are not synonyms

- Genotype lists the alleles observed at a particular marker.
- Haplotype specifies the alleles carried together on one chromosome.
- Two loci each heterozygous can have different possible cis/trans arrangements.
- Pedigree information resolves which chromosomal combination segregates with disease.

SNP arrays: Markers at many genomic loci

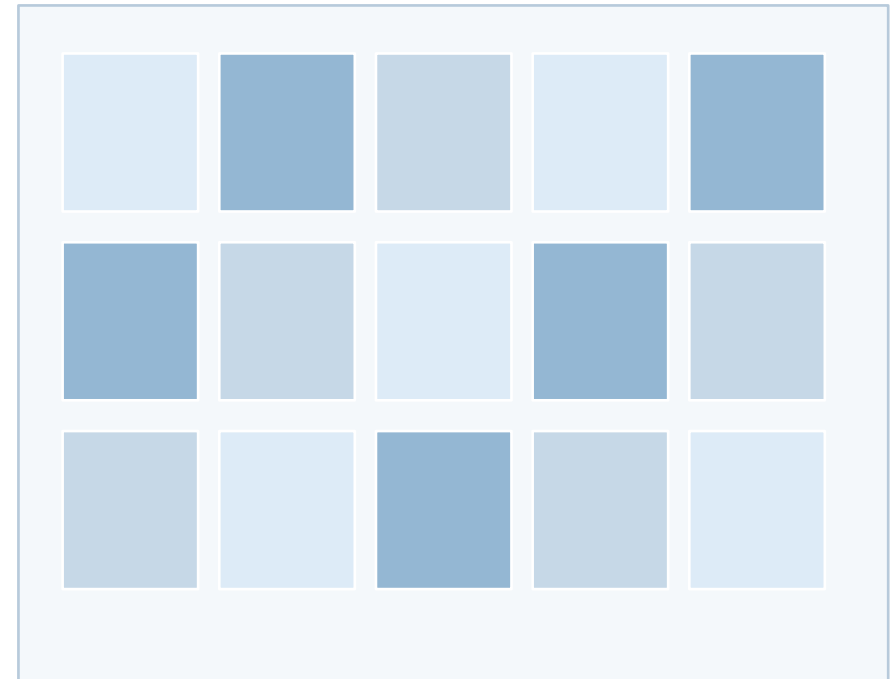
- Arrays identify alleles at numerous informative SNP sites.
- Their intensity and allele-ratio patterns also support assessment of copy-number change.
- Genotype information extends beyond what comparative genomic hybridization alone provides.



Genome-wide allele-specific probe positions

SNP arrays and loss of heterozygosity

- Loss of heterozygous genotype calls can accompany deletion.
- A region of homozygosity may occur without a change in overall copy number.
- The utility of such patterns when evaluating uniparental disomy.



Allele balance plus copy number

Clinical context for SNP-array interpretation

- SNP arrays are used when a chromosomal abnormality remains suspected despite a normal karyotype.
- Examples include congenital abnormalities, dysmorphic features, and developmental delay.
- Copy-number and allele-specific patterns should be interpreted together.
- An SNP-array pattern is not equivalent to a sequence-confirmed causal mutation.

Repeat-length polymorphisms

Variation in the number of tandem DNA repeats

A short sequence motif is repeated consecutively.

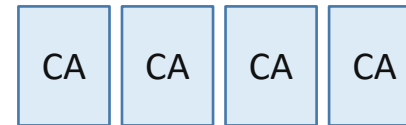
Different chromosomes may carry different repeat counts.

Repeat-count variation creates alleles distinguishable by DNA-fragment length.

A tandem repeat creates length alleles

- The repeated sequence forms a contiguous block of DNA.
- Alleles with different repeat counts have different lengths.
- The flanking sequences allow PCR amplification of the polymorphic locus.

Allele 1



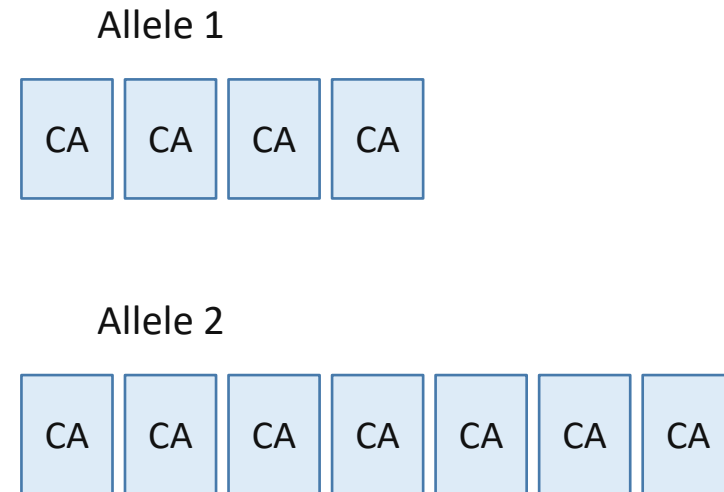
Allele 2



Allele length differs because repeat count differs

Microsatellites: Short repeat motifs

- Microsatellite repeat units are typically 1–6 base pairs.
- The repetitive region is usually under 1 kb.
- Variation in the number of repeat units supplies marker diversity.

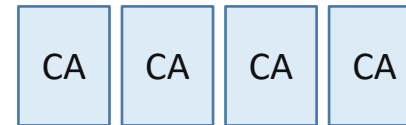


Allele length differs because repeat count differs

CA repeat as an illustrative microsatellite

- CA is a dinucleotide motif.
- The same locus can carry distinct CA-repeat counts.
- A longer CA block produces a larger PCR fragment if the flanking positions are unchanged.

Allele 1



Allele 2



Allele length differs because repeat count differs

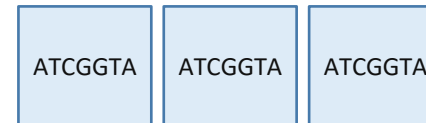
Why microsatellites are informative

- The repeat count may vary substantially between chromosomes.
- High allelic diversity increases the chance of differentiating homologues.
- Inheritance of a specific repeat-length allele can be followed through a family.
- An allele length must be tied to the correct locus and assay.

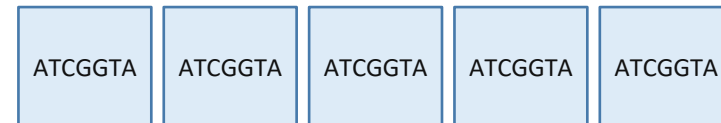
Minisatellites: Larger repeat motifs

- Minisatellite loci typically span about 1–3 kb in the Robbins description.
- Individual repeat motifs are usually 10–100 base pairs.
- Variation in tandem-repeat number also generates distinct fragment-length alleles.

Allele 1



Allele 2



Allele length differs because repeat count differs

Microsatellites versus minisatellites

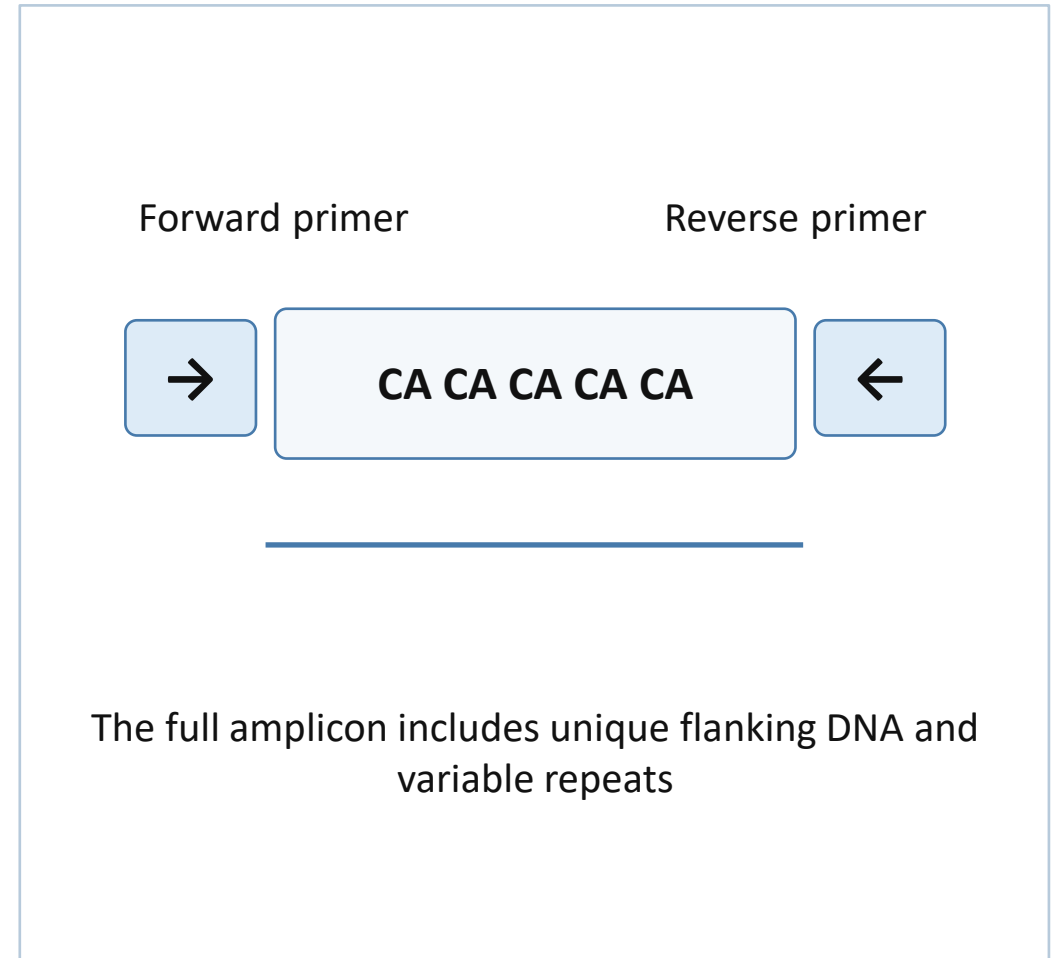
Feature	Microsatellite	Minisatellite
Repeat motif	Typically 1–6 bp	Typically 10–100 bp
Locus length	Usually <1 kb	Usually 1–3 kb
Source of variability	Number of short units	Number of larger units
Teaching application	Family linkage by short repeats	VNTR genotyping and identity assessment

VNTR describes the polymorphic mechanism

- Variable-number tandem repeats contain repeat units present in differing counts.
- The different allele lengths can distinguish two inherited chromosomes.
- Minisatellite VNTR markers as a common identity-testing approach.
- Microsatellite short tandem repeats are also widely used as polymorphic markers.

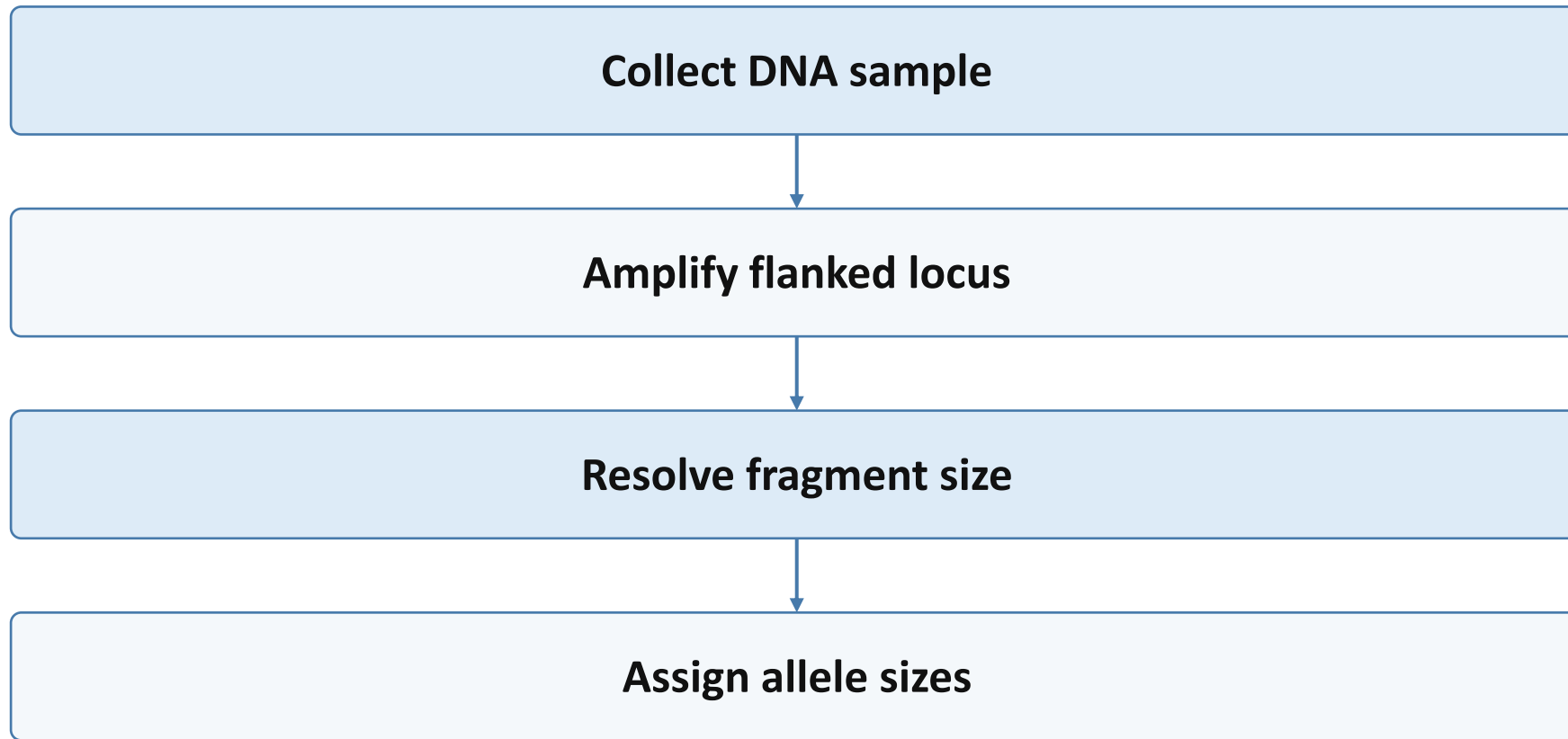
PCR targets the repeat and both flanks

- Forward and reverse primers anneal to unique sequences outside the repeat tract.
- The amplified fragment includes the polymorphic repeat block.
- Alleles differing in repeat count can generate distinct-sized PCR products.



From polymorphic DNA to an allele call

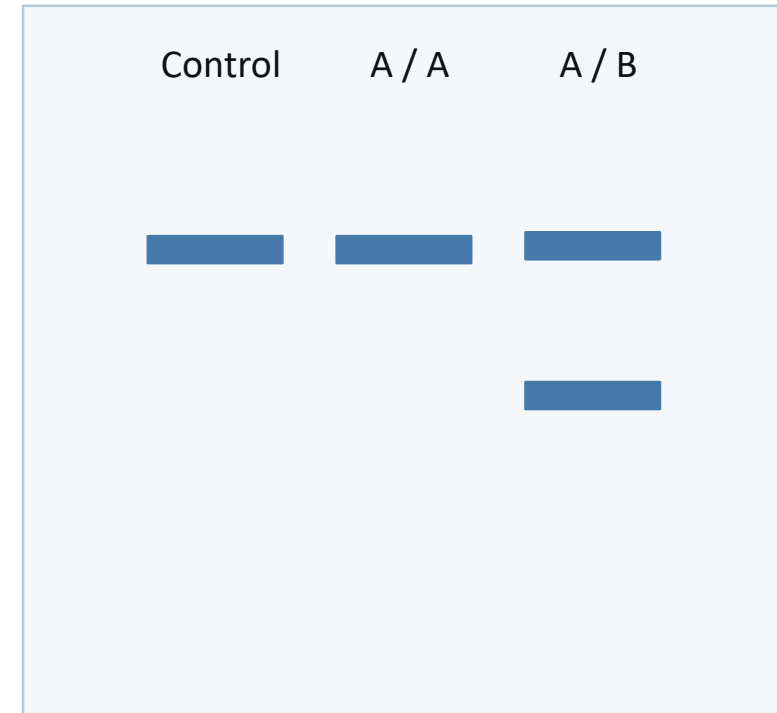
PCR-based repeat-length genotyping



The measurement reflects a marker allele, not necessarily the causal genetic defect.

An allele-size profile is a genotype

- One autosomal repeat locus ordinarily contributes up to two inherited alleles.
- A heterozygote may show two distinguishable fragment sizes.
- A homozygote has matching sizes, which do not identify parental origin by themselves.

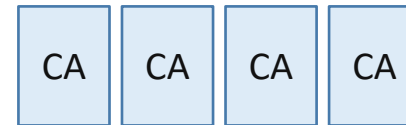


Schematic bands only — not an electrophoresis result

PCR fragment length follows repeat count

- The primer-defined flanks remain the same in this illustration.
- Adding motif copies lengthens the amplicon.
- A marker allele should be interpreted as a measured fragment or repeat count, not as a disease diagnosis.

Allele 1



Allele 2



More repeats → longer amplicon

Interpreting a family repeat-marker result

- Establish the parental allele sets first.
- Determine which marker allele the affected parent transmitted.
- Use informative affected and unaffected relatives to determine phase.
- Consider recombination if marker and phenotype become discordant.

Practice: Is this marker informative?

A parent with disease is 120/136 at a repeat marker; The other parent is 120/120. A child is 120/136. What can be concluded?

A: The child inherited 136 from the affected parent

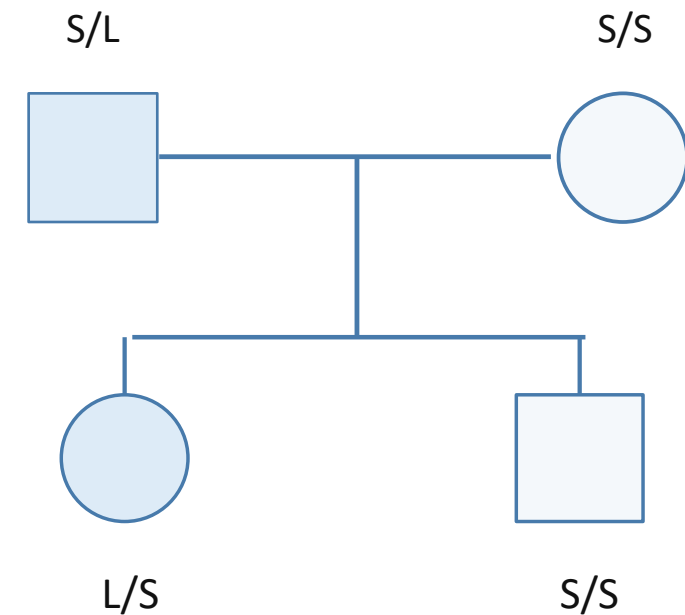
B: The 136 allele is proven pathogenic

C: The child must have the disease

Explanation: The transmitted 136 marker allele is identified; A disease association still requires pedigree-based phase information. Values are illustrative fragment sizes, not clinical reference data.

Illustrative segregation in a family

- In an established family, a particular marker allele can cosegregate with a phenotype.
- Relatives without that allele may inherit the unaffected chromosomal homologue.
- Interpretation is conditional on the family, informative meioses, and recombination.



Illustrative family-specific allele tracking

Phase determines the diagnostic meaning

- An affected parent can carry both a disease-associated and a normal chromosome.
- Only one marker allele is in cis with the pathogenic allele.
- Testing relatives helps identify that linked marker allele.
- Without phase, a marker genotype can be misleading.

Familial PKD1 linkage example

An application

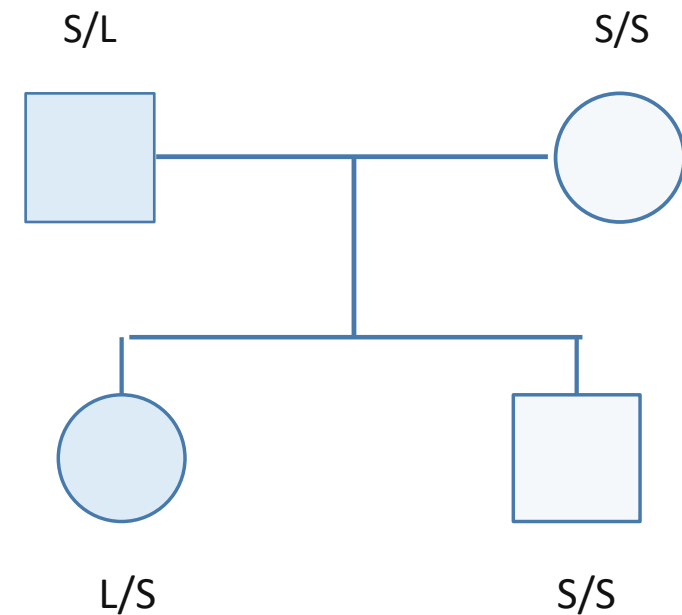
PKD1 was historically difficult to sequence comprehensively.

A linked CA-repeat marker can distinguish parental chromosomal copies.

The disease-associated repeat allele can then be tracked among relatives.

PKD1: Identify the disease-linked marker

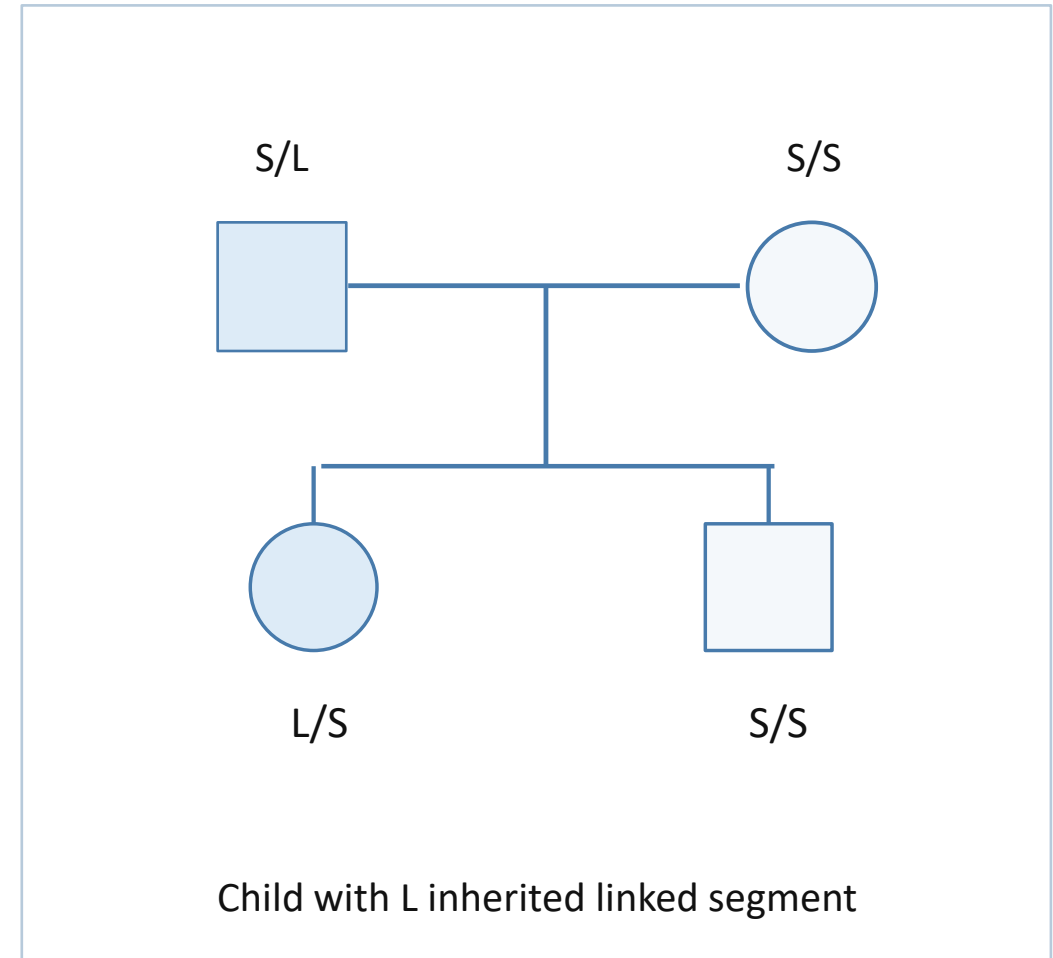
- In a conceptual family, a parent carries marker alleles short (S) and long (L).
- Affected relatives establish that L is linked to the PKD1 disease allele in that pedigree.
- L becomes a surrogate for the disease chromosome—not a causal sequence variant.



Illustrative family-specific allele tracking

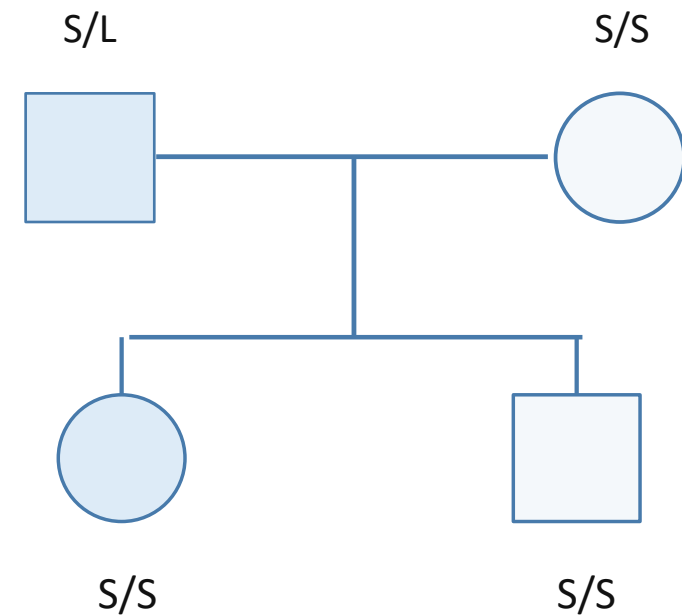
PKD1: Child inherits the linked marker

- If a child inherits L from the informative parent, the child has inherited the tracked chromosomal segment.
- This supports inheritance of the familial disease-linked chromosome.
- Recombination and assay/family limitations must still be considered.



PKD1: Child inherits the other allele

- Inheritance of S identifies the alternate tracked chromosomal segment in the schematic family.
- That result argues against inheritance of the L-tagged disease chromosome.
- Indirect genotyping cannot categorically exclude a mutation if recombination or phase error occurs.

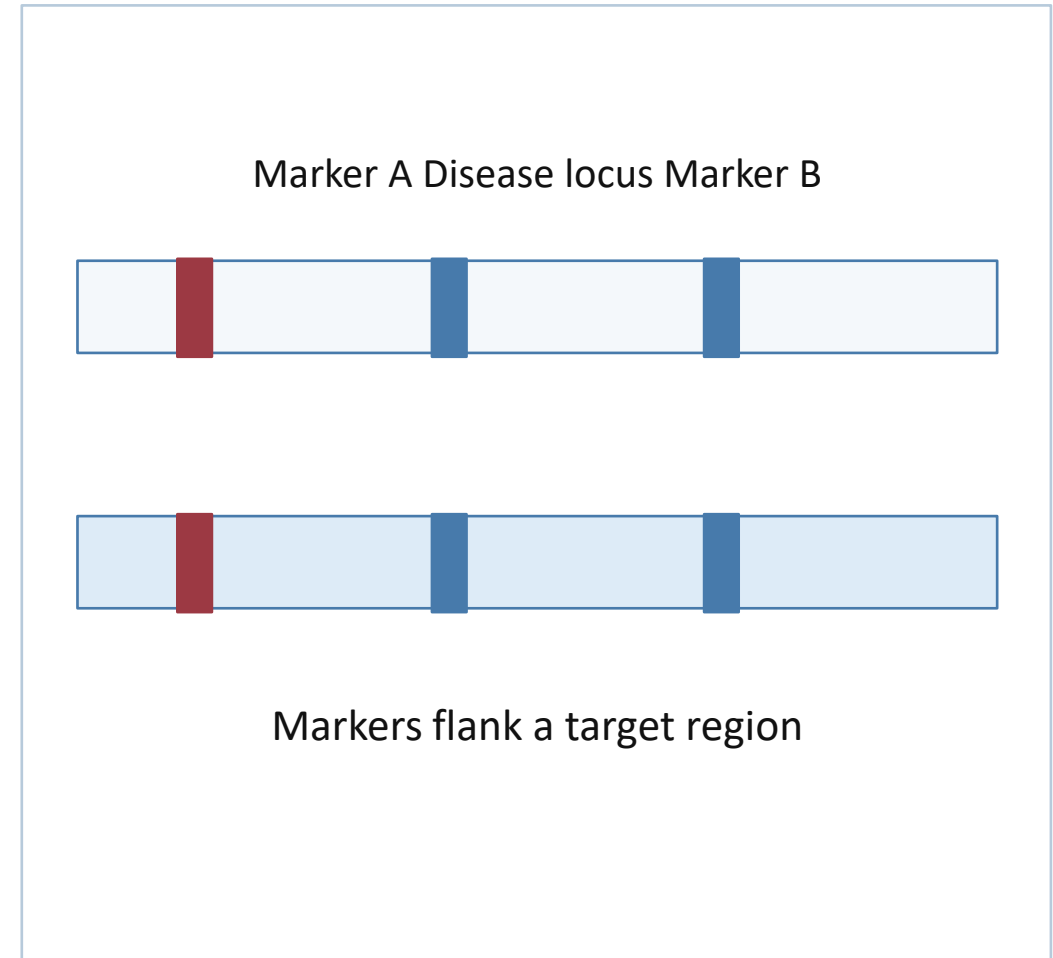


Limits of a single linked marker

- A crossover between the marker and PKD1 could reverse the inferred association.
- Uninformative alleles can prevent discrimination between chromosomes.
- A phenotype may be uncertain or not yet expressed in a younger relative.
- Family-specific interpretation should not be generalized to unrelated pedigrees.

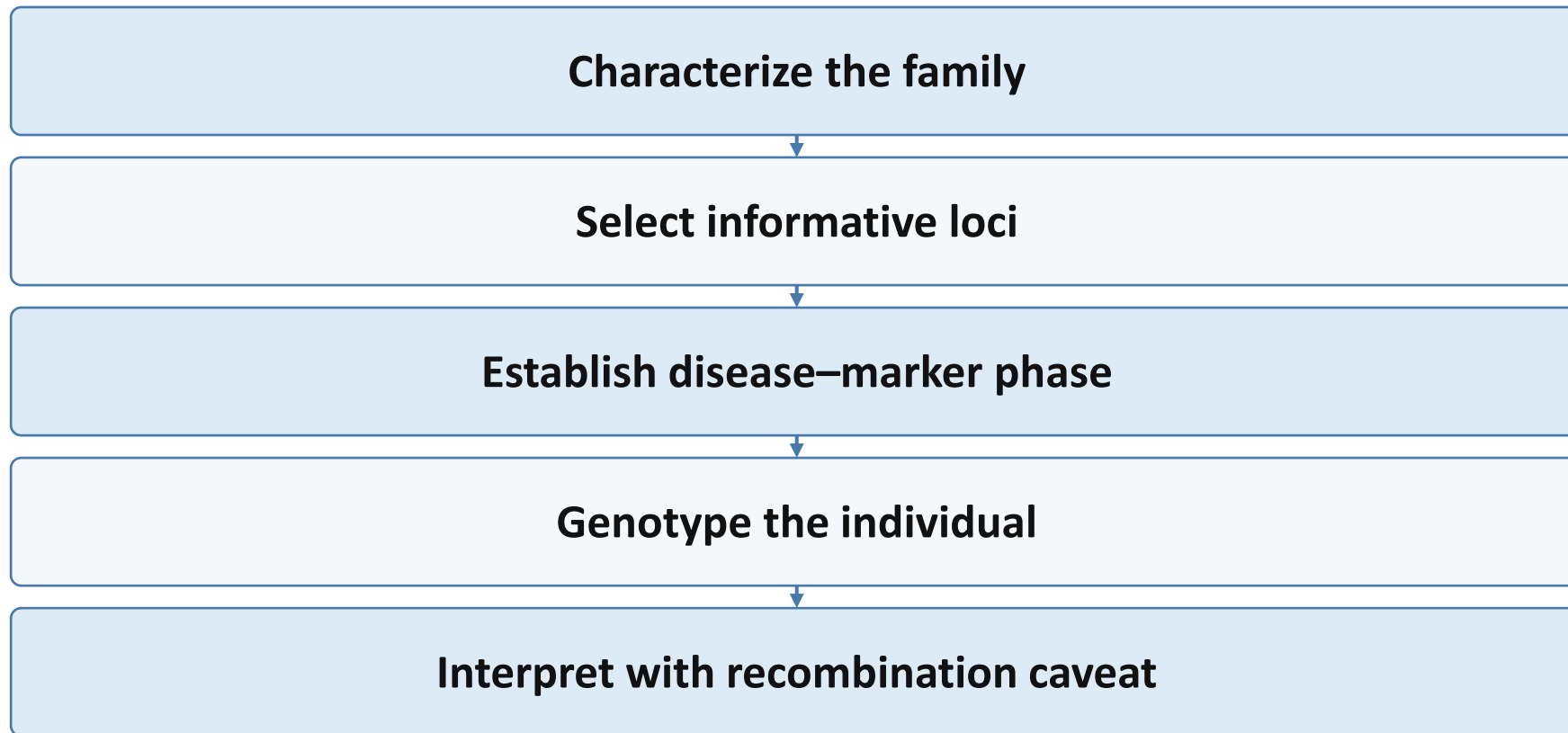
Multiple flanking markers strengthen tracking

- Markers on either side of a locus can identify a chromosomal interval.
- Discordant allele patterns can suggest a crossover within that interval.
- Several informative loci supply more segregation evidence than one isolated marker.



A practical linkage-testing sequence

Illustrative diagnostic-laboratory reasoning



Direct pathogenic-variant testing is preferable when feasible and clinically validated.

Selecting an approach to inherited disease

Clinical situation	Most direct interpretation
Known familial pathogenic allele	Variant-specific molecular diagnosis
Difficult or unresolved familial variant	Informative linked marker / haplotype approach
Uninformative pedigree	Surrogate-marker inference is weak or impossible
Suspected chromosomal imbalance	Copy-number and allele-pattern assessment

Other applications of polymorphic markers

Identity and relative contributions of different genomes

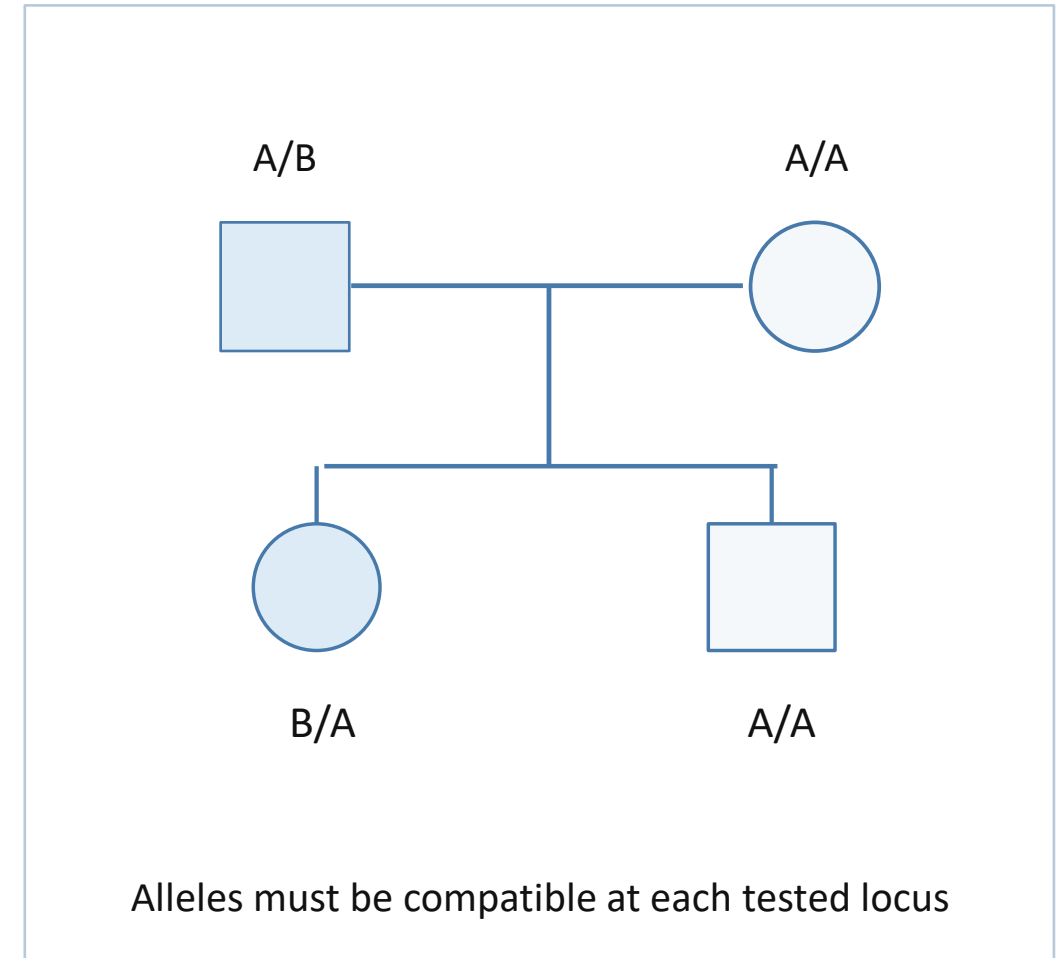
Genetic identity and relatedness

Paternity and forensic investigations

Allogeneic hematopoietic transplantation

Paternity: Inheritance at polymorphic loci

- Each offspring receives one allele from each parent at an autosomal locus.
- Informative marker comparisons can assess whether alleged relationships are compatible.
- Strong conclusions require multiple independent informative loci and validated statistical analysis.

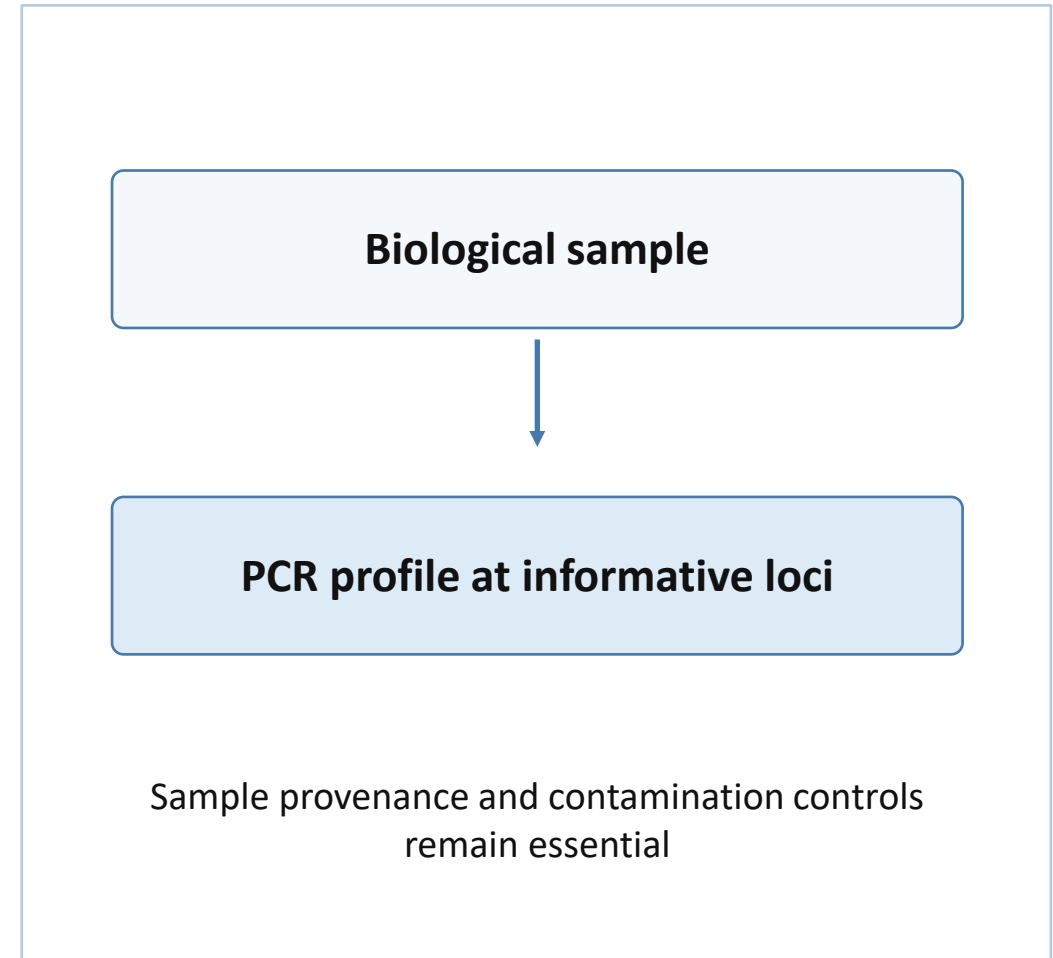


Identity analysis needs multiple markers

- A single repeat-size match is not unique to a person.
- Variation across many loci improves individual discrimination.
- Shared inheritance permits testing of biological relationships.
- The analytical interpretation must account for sample quality and relevant uncertainty.

Forensic DNA: Amplify informative fragments

- PCR enables analysis from limited or degraded biological material in suitable circumstances.
- Repeat-marker genotypes can support identification or exclusion.
- Results depend on sample integrity, contamination control, and validated interpretation.

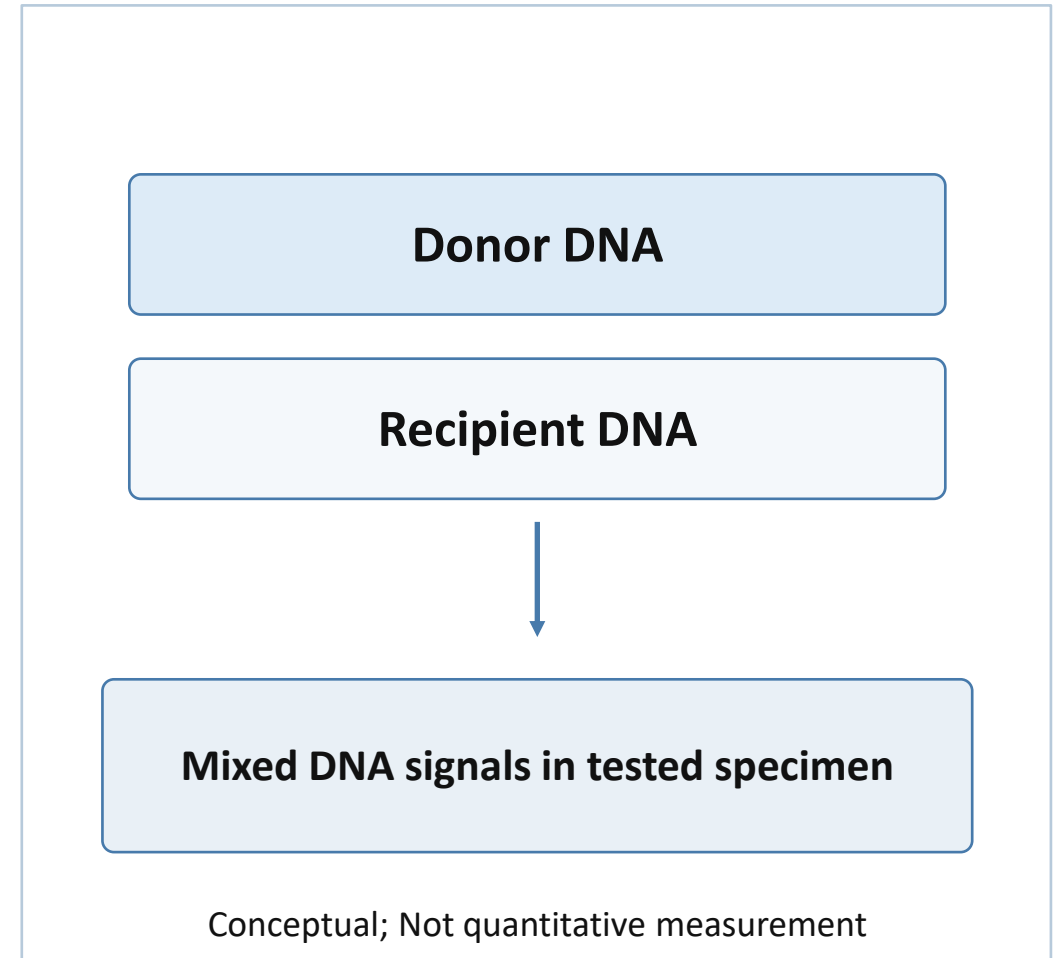


Do not equate a DNA match with an event

- A DNA profile pertains to the origin of biological material.
- Analytical results do not independently establish when, how, or why material was deposited.
- Marker evidence must be interpreted in the appropriate forensic context.
- This as a major application, not a replacement for the full investigation.

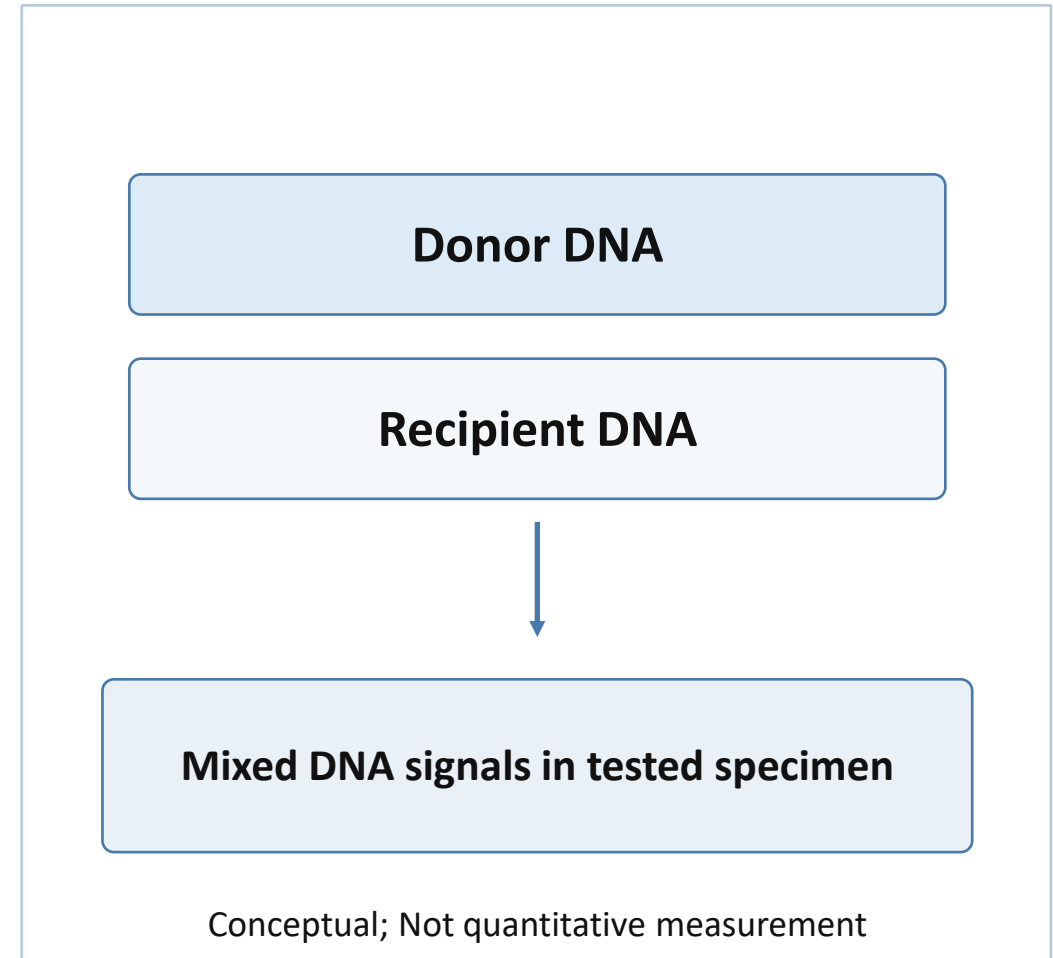
Why transplantation requires donor markers

- Allogeneic hematopoietic cell transplantation introduces donor-derived hematopoietic cells.
- Donor and recipient may have distinguishable genetic marker alleles.
- Marker assays can assess their relative DNA contributions after transplant.



Chimerism: Two genetic DNA sources

- Post-transplant marrow or blood may contain donor-derived and recipient-derived DNA.
- Informative polymorphic markers distinguish these sources.
- The assay can detect and quantify the relative contributions using validated laboratory methods.



Reading a chimerism result correctly

- A change in donor/recipient marker signal reflects changing sample composition.
- The clinical meaning depends on specimen, timing, cell lineage, and disease context.
- A marker assay measures genetic contributions, not directly graft function or malignancy.
- The monitoring application without supplying a universal cutoff.

Associated topic: Epigenetic alterations

Regulation without a change in primary DNA sequence

DNA sequence polymorphisms identify variable genetic loci.

Epigenetic modifications alter DNA or chromatin chemistry.

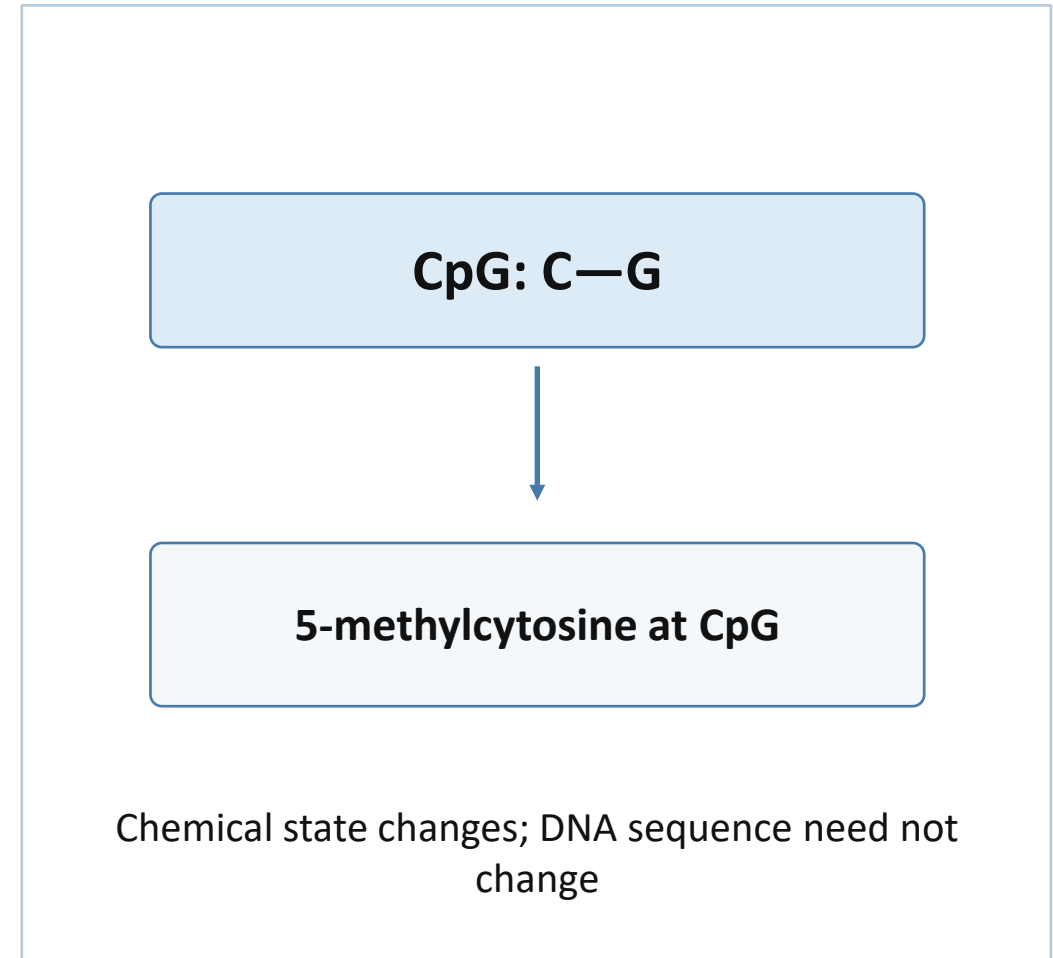
Different laboratory methods are required to detect methylation status.

Sequence polymorphism vs epigenetic state

Polymorphic DNA marker	Epigenetic alteration
Underlying nucleotide sequence or repeat length differs	DNA sequence can remain unchanged
Allele is inherited as genomic sequence	Chemical modification can alter regulatory state
SNP/repeat genotyping assays	Methylation-sensitive or bisulfite-based assays
Linkage, identity, chimerism	Imprinting disorders and gene-silencing assessment

CpG methylation is an epigenetic signal

- DNA cytosines in CpG dinucleotides can be methylated.
- Chemical modification changes regulatory properties without changing the base sequence.
- Promoter methylation assays can therefore detect information invisible to ordinary sequence genotyping.



Chromatin modifications are also epigenetic

- Histone proteins may undergo methylation or acetylation.
- These modifications influence chromatin and gene regulation.
- They are conceptually distinct from sequence polymorphisms.
- A repeat-marker assay cannot determine histone modification state.

Clinical examples of methylation testing

- Fragile X syndrome: FMR1 hypermethylation is associated with gene silencing.
- Prader–Willi syndrome: Methylation analysis is essential to the diagnostic workup.
- Angelman syndrome: Methylation analysis is likewise essential.
- The significance of each assay depends on locus-specific imprinting and disease mechanism.

Bisulfite separates methylated and unmethylated C

- Sodium bisulfite converts unmethylated cytosine to uracil.
- Methylated cytosine is protected from this conversion.
- Following amplification/sequencing, the two states can be distinguished by their sequence readout.

Unmethylated C → U

Methylated C → C

Following PCR/sequencing, U is read as T

Alternative methylation assay mechanisms

- Methyl-CpG binding proteins can enrich DNA carrying methylation.
- Methylation-sensitive restriction enzymes distinguish DNA by cleavage behavior.
- Some sequencing platforms directly sense modified nucleotides using polymerase kinetics or nanopore signals.
- These methods assay epigenetic state rather than repeat-number polymorphism.

Match the clinical question to the assay

Question	Relevant assay principle
Which repeat allele was inherited?	Flanking-primer PCR and fragment sizing
Which SNP alleles are present?	SNP genotyping or allele-specific arrays
What DNA fraction is donor-derived?	Informative polymorphic marker quantification
Is a CpG region methylated?	Bisulfite or other methylation-sensitive assay

Application: Interpreting a linked marker

A family-specific repeat allele cosegregates with disease. An unaffected child lacks that allele. What is the most defensible conclusion?

A: Disease is excluded with absolute certainty

B: The child did not inherit the tracked disease-linked segment

C: The repeat itself must cause the disorder

Explanation: The marker tracks a chromosomal segment in that pedigree. Recombination, incomplete phase data, and test limitations prevent absolute inference.

Application: The transplant laboratory

After an allogeneic transplant, recipient-specific DNA markers remain measurable. What has this assay shown?

A: Presence of recipient-derived DNA in the tested specimen

B: Proof of leukemia relapse

C: Proof of failed engraftment

Explanation: The marker assay demonstrates donor/recipient genetic contributions. Interpretation as engraftment, residual disease, or relapse requires additional clinical and laboratory information.

Summary: The logic of polymorphic markers

- Polymorphisms make inherited chromosomes distinguishable.
- Linkage permits indirect tracking of difficult-to-assay familial disease alleles.
- SNPs and repeat-length polymorphisms provide complementary genetic landmarks.
- PCR-based marker assays extend to kinship, forensic identification, and transplant chimerism.

Reference and source boundaries

- Primary text: Robbins, Cotran & Kumar, Pathologic Basis of Disease, 11th ed. (2026), Chapter 5, pp. 162–164.
- Core section: “Polymorphic Markers”; adjacent section: SNP genotyping arrays.
- Associated section in this teaching unit: “Epigenetic Alterations.”
- Examples labeled “illustrative” are original teaching schematics, not clinical specimen data or reproduced textbook figures.