

Array Technology

Genome-wide DNA copy number and SNP genotyping

Comparative genomic hybridization • Single-nucleotide polymorphism arrays • Loss of heterozygosity

Molecular genetic diagnosis

Dr Abu Suraih Sakhri

MBBS, MD (Pathology)

www.abusuraihshakhri.com

Learning objectives

- Explain why genome-wide arrays complement tests directed at a suspected locus.
- Describe the workflow and copy-number readout of array comparative genomic hybridization (CGH).
- Interpret how SNP arrays combine DNA dosage with allelic genotype information.
- Connect copy-number variation and loss of heterozygosity to clinical diagnostic questions.

The central diagnostic problem

- A pathogenic chromosomal change may be suspected before its genomic location is known.
- A locus-specific probe is effective when the suspected region is already identified.
- A genome-wide platform surveys many intervals without specifying one candidate region in advance.

The testing question determines whether targeted or genome-wide interrogation is needed.

Targeted FISH versus genome-wide arrays

FISH: Locus-directed

- Uses probes directed at specified chromosome regions.
- Can examine aneuploidy, deletions, amplifications or rearrangements at targeted loci.
- Requires a diagnostic hypothesis about the region to investigate.

Microarray: Genome-wide

- Surveys many loci distributed across the genome.
- Detects relative copy-number gains and losses across interrogated intervals.
- SNP platforms add allelic genotype information.

What class of genomic change is being sought?

Deletion

A segment contains less DNA than the reference expectation.

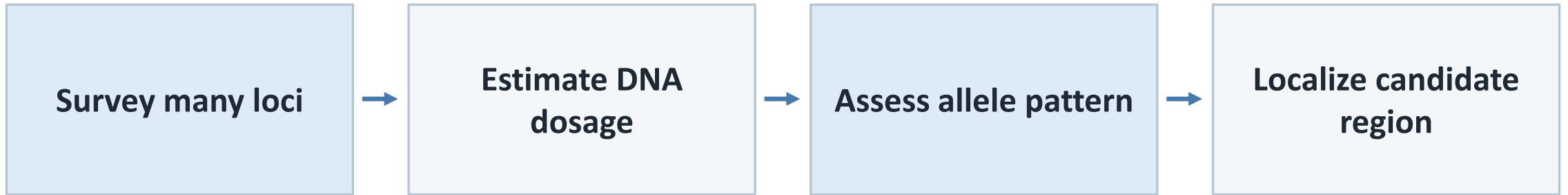
Duplication / gain

A segment contains more DNA than the reference expectation.

Loss of heterozygosity

A region loses the mixed-allele pattern, with or without a decrease in copy number.

A genome-wide measurement answers a different question



- An array provides a genomic map of abnormalities.
- Clinical interpretation still requires a phenotype and an appropriate diagnostic context.

Genomic coordinates are interrogated by ordered probes

Chromosomal mapping

A chromosome can be surveyed as successive probe-defined intervals



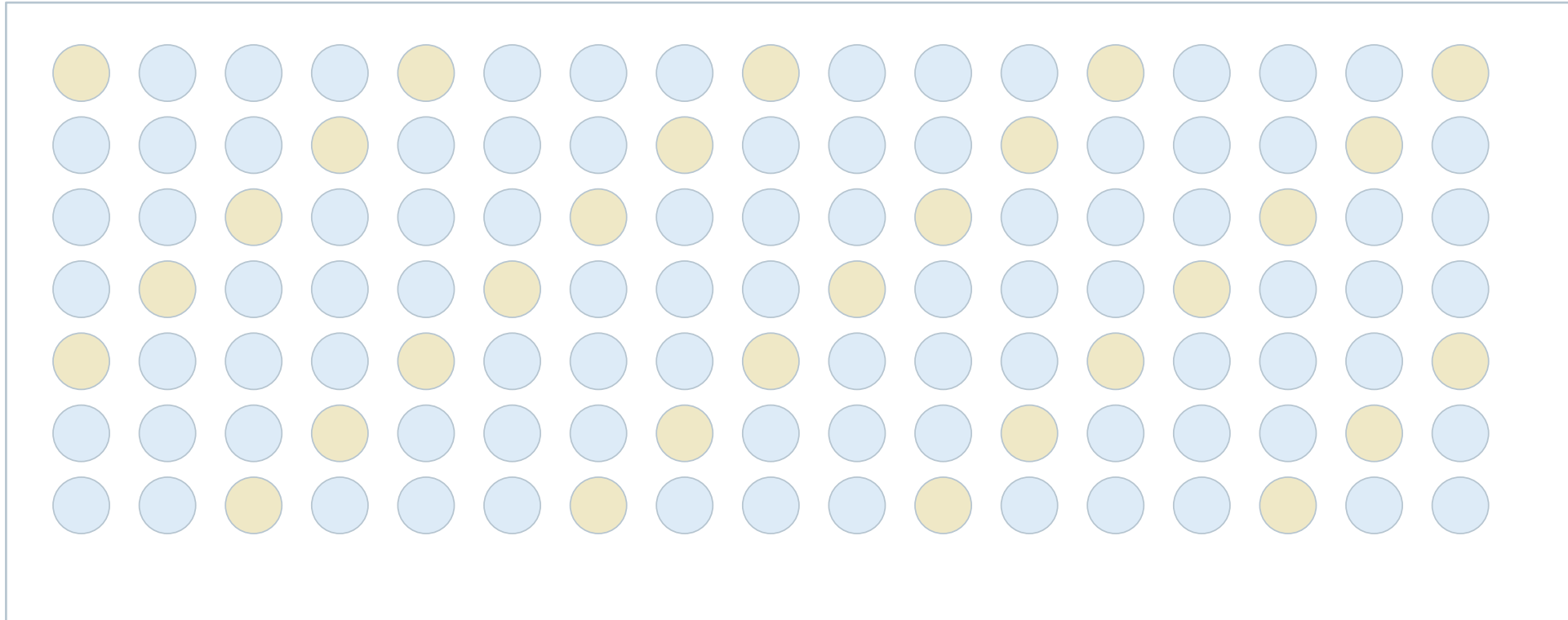
Genomic coordinates →

Contiguous probes provide positional information along the genome

Schematic — not patient data

- Each probe corresponds to a known interval or polymorphic position.
- A chromosome can be interpreted as successive measurements along its length.

An array is a grid of molecular probes



Immobilized probes form a parallel genomic assay

Schematic arrangement of probe spots; Spacing, colors and counts are illustrative

A probe spot is not itself a clinical diagnosis

- It measures hybridization from DNA complementary to a defined genomic target.
- An abnormal intensity or allele pattern must be related back to its genomic coordinate.
- Neighboring probes strengthen interpretation of a regional alteration.

The fundamental unit is a probe-linked observation, not a disease label.

Two principal microarray strategies

Array CGH

Compares relative DNA abundance in test and reference material.

SNP genotyping array

Determines genomic copy-number patterns and genotype/allele information.

Array CGH: The analytical question

Copy-number analysis

- Is the amount of DNA in the test sample comparable to normal reference DNA?
- Which genomic intervals are relatively depleted or enriched?
- How small can an alteration be while remaining supported by the probe design?

Test DNA and reference DNA

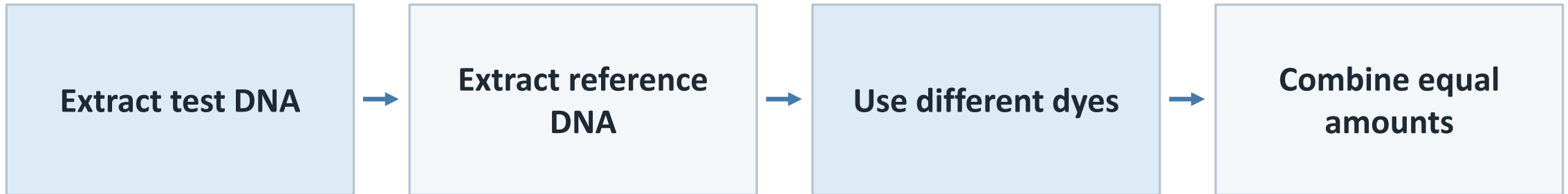
Test specimen

- Genomic DNA from the individual being evaluated.
- Contains the copy-number state under investigation.
- Labeled with one fluorescent channel.

Reference specimen

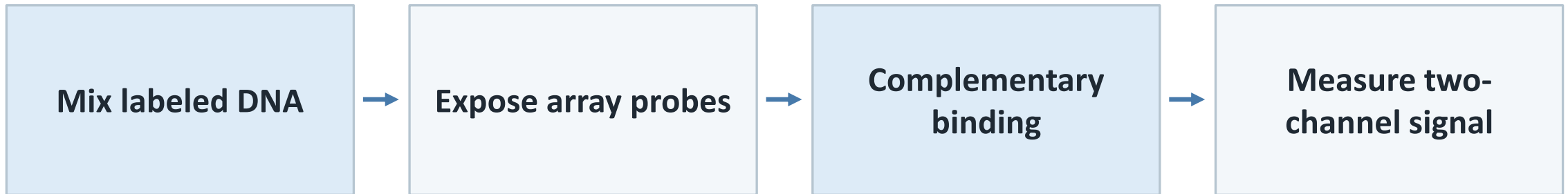
- Normal comparator DNA used as the baseline.
- Contributes the second labeled signal.
- Equal DNA amounts are mixed before hybridization.

Array CGH begins with differential labeling



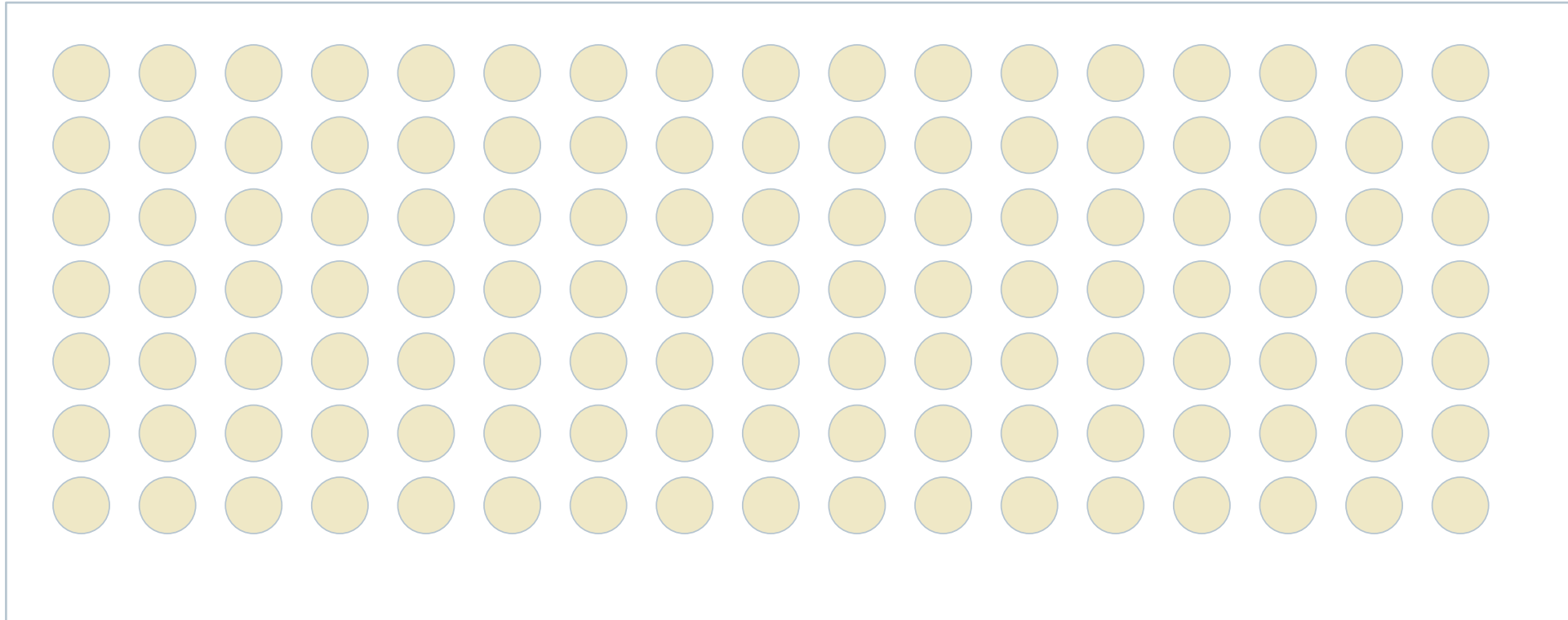
- Two distinguishable fluorescent dyes.
- Dye identity must be known before a gain or loss is inferred.

Hybridization maps the DNA to its matching probes



- Each probe relates signal to a specific genomic sequence.
- Relative binding of test and reference material is compared at every location.

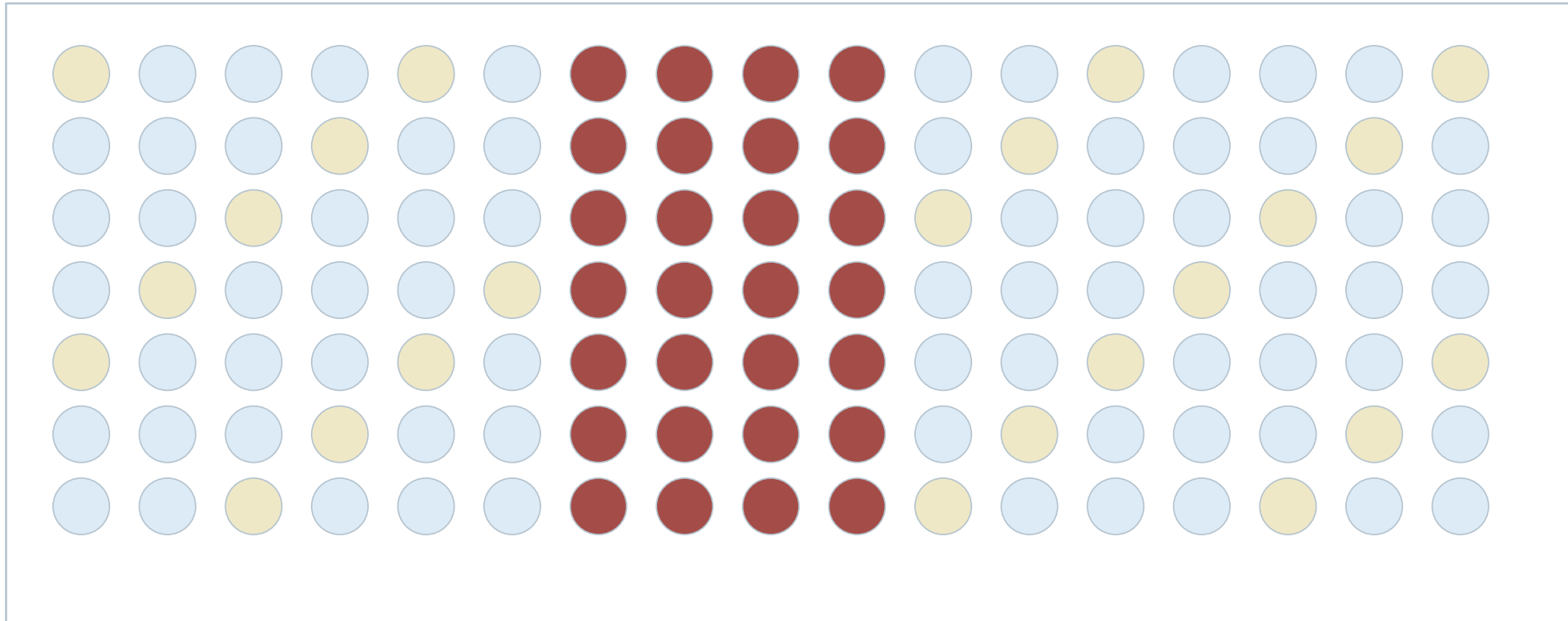
Normal relative dosage yields balanced signals



Balanced red and green emission can appear yellow

Schematic arrangement of probe spots; Spacing, colors and counts are illustrative

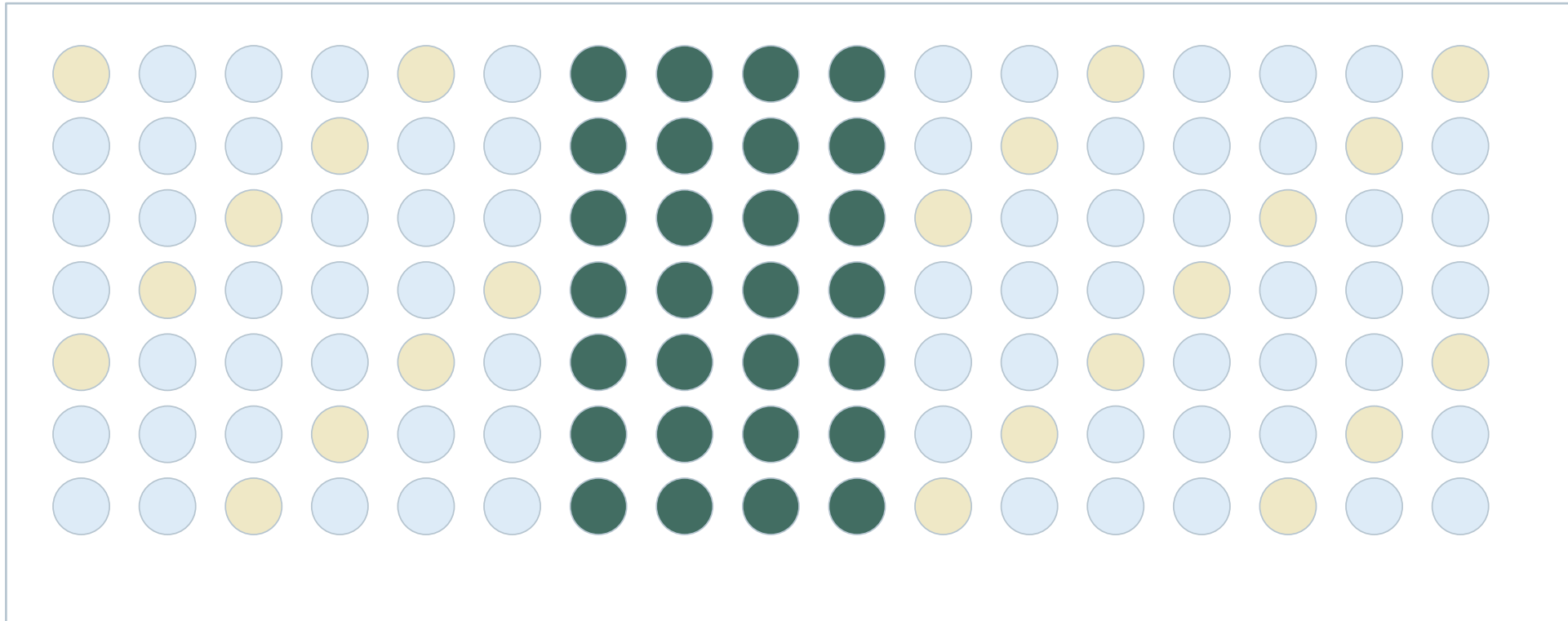
Relative DNA loss skews a subset of spots



Illustrative color skew in affected probe intervals

Schematic arrangement of probe spots; Spacing, colors and counts are illustrative

Relative DNA gain skews a subset of spots



Illustrative color skew in affected probe intervals

Schematic arrangement of probe spots; Spacing, colors and counts are illustrative

Fluorescence color is not a universal biological code

- A red-skewed or green-skewed spot indicates relative imbalance between labeled samples.
- The direction representing gain or loss depends on which dye labels the test DNA.
- Read the assay legend or channel designation before assigning biological meaning.

Color direction follows labeling orientation, not an intrinsic deletion/gain color.

Qualitative CGH dosage states

Compare the labeled test signal to reference at each genomic interval

Balanced	Loss	Gain
Test $\approx$ reference	Test < reference	Test > reference
Comparable DNA abundance	Relative depletion	Relative enrichment
<i>No relative dosage skew</i>	<i>Deletion-type readout</i>	<i>Duplication-type readout</i>

Qualitative illustration — no measured patient signals shown

A copy-number call is localized to a genomic interval

Copy-number loss

A chromosome can be surveyed as successive probe-defined intervals



Genomic coordinates →

Region with reduced relative test DNA

Schematic — not patient data

- The contiguous region with a shifted signal defines a candidate loss.
- Probe coordinates constrain the location of the suspected abnormality.

Genome-wide coverage is built from discrete probes

- Arrays contain many defined genomic targets arranged to interrogate the genome.
- Gaps between targets constrain how precisely a change can be localized.
- Uniform claims of unlimited resolution are unjustified: Design determines sensitivity to focal events.

Probe density determines potential spatial resolution

Sparse tiling

More widely spaced probes can overlook a narrow interval between interrogated sites.

Dense tiling

Closer probes offer finer interrogation of small regions.

Evidence threshold

Several adjacent probes commonly need to support an interval-level call.

Textbook resolution example: 200 bp versus 1 kb

Probe every 200 bp

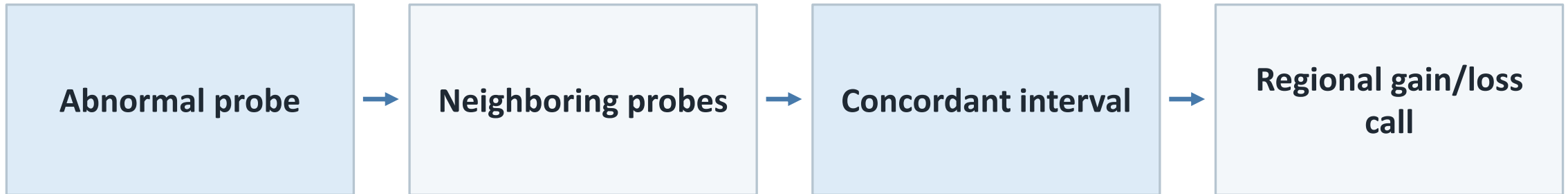
- More sites are sampled over a fixed genomic distance.
- Offers greater potential to detect focal gains or losses.
- Still depends on having a coherent pattern across probes.

Probe every 1 kilobase

- Fewer sites sample that same distance.
- A small event may affect fewer or no probe locations.
- Individual spots should not be overinterpreted.

Illustrates spacing, not a universal clinical detection threshold.

Why adjacent probes matter for a CNV call



- Several probes corresponding to contiguous genomic sequences usually contribute to a detected interval.
- The resulting genomic interval is more interpretable than an isolated spot.

Focal alterations require sufficient probe coverage

Focal gain

A chromosome can be surveyed as successive probe-defined intervals



Genomic coordinates →

Region with increased relative test DNA

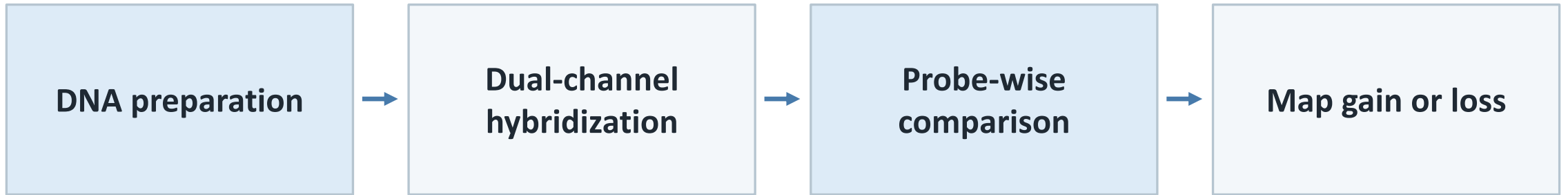
Schematic — not patient data

- An event limited to a short stretch affects only probes inside that interval.
- Denser probe tiling improves the chance that small gains and losses are detected.

Analytical signal and biological significance are different

- An altered array signal suggests a genomic copy-number change, not automatically disease causation.
- The interval, direction of dosage change and clinical features all matter.
- The textbook explains array detection; It does not provide universal pathogenicity cutoffs for array calls.

Array CGH workflow from specimen to region



- Genome-wide analysis can reveal an unsuspected locus.
- A candidate interval may then motivate focused interpretation or testing.

Worked pattern: A balanced interval

Conceptual exercise — reference-balanced hybridization

Copy-number track	Expected conclusion
Balanced Comparable total DNA <i>No gain/loss signal</i>	No CNV detected In the interrogated interval <i>Does not assess every variant type</i>

Qualitative illustration — no measured patient signals shown

Worked pattern: A deleted interval

Conceptual exercise — relative test DNA depletion

Copy-number track	Interpretation
Reduced	Copy-number loss
Test DNA lower than reference	Interval-level deletion
<i>Consecutive affected probes</i>	<i>Assess genomic context</i>

Qualitative illustration — no measured patient signals shown

Worked pattern: A gained interval

Conceptual exercise — relative test DNA enrichment

Copy-number track	Interpretation
Elevated	Copy-number gain
Test DNA above reference	Interval-level duplication
<i>Consecutive affected probes</i>	<i>Assess genomic context</i>

Qualitative illustration — no measured patient signals shown

What array CGH does not directly describe

- The CGH readout primarily measures relative DNA dosage, not which SNP alleles are present.
- A diploid region can remain balanced by dosage despite losing heterozygosity.
- This limitation motivates arrays that assay genotype together with copy number.

SNP arrays extend the information content

- Probes recognize many single-nucleotide polymorphism positions across the genome.
- Genotype-dependent signals distinguish alleles as well as informing copy-number calls.
- The two measurements address different biological properties of the same interval.

SNP platforms add a second, allelic dimension.

A SNP is a common single-base polymorphism

- A single-nucleotide polymorphism is a genomic position with allelic variation.
- Two alleles at a locus can be represented abstractly as A and B.
- SNP sites function as both genomic landmarks and polymorphic genetic markers.

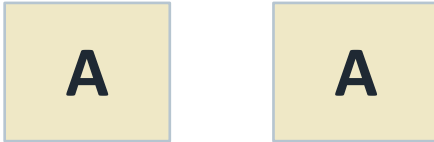
SNPs provide many distributed landmarks

- SNPs as the most common form of DNA polymorphism.
- Its approximate average spacing is one SNP every 1,000 nucleotides across the genome.
- Dense SNP targeting enables genotype information to be mapped to many chromosome regions.

Average SNP spacing is not identical to probe spacing on every array.

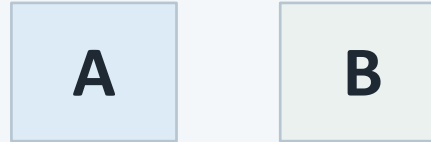
A diploid SNP locus has three possible genotypes

AA



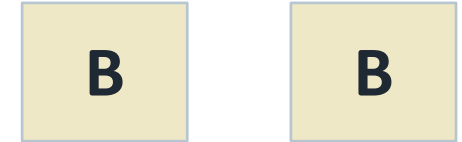
Both alleles A

AB



One A and one B

BB



Both alleles B

AA and BB are homozygous states; AB is heterozygous.

SNP positions anchor allele calls to the genome

SNP landmarks

A chromosome can be surveyed as successive probe-defined intervals



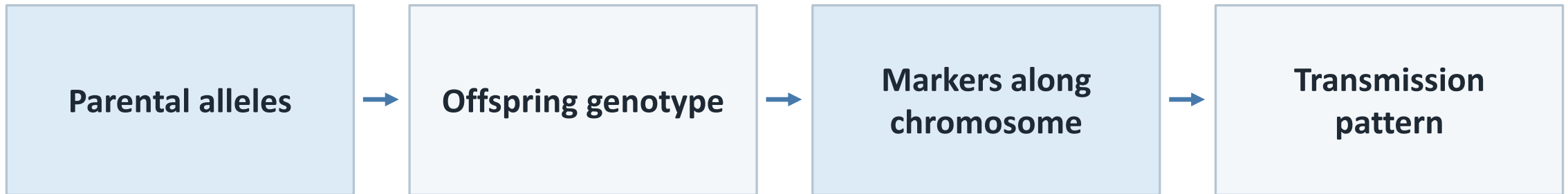
Genomic coordinates →

Contiguous probes provide positional information along the genome

Schematic — not patient data

- A genotype result is interpretable because the SNP has a known genomic location.
- Multiple SNPs across a region allow examination of local allele patterns.

SNP alleles can be followed through inheritance



- SNPs as genetic markers that can be followed from parent to child.
- Allelic information is additional to the total DNA amount.

SNP array probes discriminate among alleles



- Different commercial platforms implement SNP analysis differently.
- The shared principle is genotype-dependent signal at known genomic locations.

SNP array output has two complementary dimensions

Each genomic interval can be evaluated for both abundance and allelic identity

Total signal	Allele-specific signal
Copy number	Genotype
DNA dosage across loci	AA / AB / BB states
<i>Gain / neutral / loss</i>	<i>Zygosity information</i>

Qualitative illustration — no measured patient signals shown

Total intensity addresses dosage

- Intensity across probes supports calls of copy-number variation.
- The resulting regional pattern can suggest DNA gain or deletion.
- An intensity-based call does not by itself reveal which alleles are represented.

Allelic ratio addresses zygosity

- Allele-specific signals discriminate the alleles present at a SNP locus.
- A diploid AB genotype represents a mixed-allele (heterozygous) position.
- AA and BB are homozygous positions; Their distribution matters across a region.

Normal diploid regions may contain AA, AB and BB

AA

A

A

Both alleles A

AB

A

B

One A and one B

BB

B

B

Both alleles B

A normal region has a mixture of SNP genotypes across different loci.

A deletion may alter both intensity and genotype pattern

The textbook example combines decreased signal with loss of AB genotype calls

Copy-number evidence

Decreased dosage

Signal below diploid baseline

Supports deletion

Allele evidence

AB calls absent

Mixed-genotype pattern lost

Supports reduced heterozygosity

Qualitative illustration — no measured patient signals shown

Loss of heterozygosity is an allelic observation

- A region that would otherwise contain mixed alleles instead shows reduced heterozygosity.
- A deletion can produce this pattern when only one copy remains.
- Loss of heterozygosity can also occur without a net reduction in DNA copy number.

A diploid interval may still show loss of heterozygosity

Copy-neutral region

A chromosome can be surveyed as successive probe-defined intervals



Genomic coordinates →

Copy number unchanged; Alleles may nevertheless become uniform

Schematic — not patient data

- Copy-number neutrality refers to retention of overall diploid DNA dosage.
- SNP genotype information can show an extended pattern of homozygosity despite that dosage.

Deletion versus copy-neutral homozygosity

The distinguishing feature is whether the copy-number measurement also decreases

Deletion with LOH

Reduced copies

Allelic diversity reduced

Dosage and allele change

Copy-neutral LOH

Diploid dosage

Allelic diversity reduced

Allele change without net loss

Qualitative illustration — no measured patient signals shown

Uniparental disomy links copy number and inheritance

- Uniparental disomy (UPD) refers to both homologous chromosome copies or a chromosomal region originating from one parent.
- Total dosage can remain diploid even when the inheritance pattern is abnormal.
- Regions of homozygosity detected by SNP arrays can therefore raise a UPD question.

An SNP genotype pattern alone does not establish parent of origin.

A deletion and UPD raise different genetic questions

Deletion-type finding

- DNA dosage is reduced in the involved genomic interval.
- An allele may be lost with its copy.
- Array analysis identifies the candidate interval.

Possible UPD-related finding

- DNA dosage may remain diploid.
- A long homozygous region may reflect altered parental inheritance.
- Further correlation is needed to establish the diagnosis.

Why imprinting disorders are a relevant example

- The source names Prader–Willi syndrome and Angelman syndrome as conditions in which UPD may matter.
- An SNP array can identify a compatible homozygosity pattern while copy number remains diploid.
- This is a screening or interpretive clue, not a direct measurement of imprinting status.

Leukemia example: Concurrent copy loss and LOH

- An SNP-array study of the short arm of chromosome 12 in pediatric leukemia.
- Reduced overall intensity is paired with disappearance of the heterozygous AB genotype pattern.
- Together these observations support a heterozygous (single-copy) deletion in that interval.

Interpret the copy-number and genotype observations together.

Joint interpretation of SNP-array tracks

Dosage readout	Genotype readout	Interpretive direction
Balanced	Expected AA / AB / BB mixture	No evident dosage or zygosity abnormality in interval
Reduced	Loss of mixed AB calls	Deletion with loss of heterozygosity
Diploid	Extended homozygosity	Copy-neutral region of homozygosity; Evaluate possible causes
Increased	Allele pattern must also be reviewed	Copy-number gain

Clinical array testing addresses unresolved chromosome questions

- The source describes routine SNP-array use in pediatric patients when an abnormality remains suspected despite a normal karyotype.
- Typical referral features include congenital anomalies, dysmorphic features, developmental delay and autism.
- A genome-wide scan may detect copy-number changes not apparent from conventional chromosome assessment.

Why normal karyotyping does not end the evaluation

Karyotype is normal

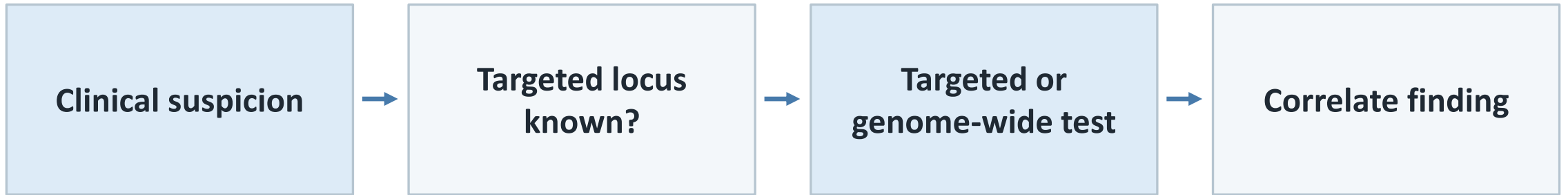
- No visually recognized abnormality at the resolution of the initial analysis.
- This result does not rule out smaller structural gains or losses.

Chromosomal suspicion persists

- The phenotype may suggest a genomic imbalance.
- Higher-resolution genomic interrogation can be indicated.

Testing decisions remain phenotype- and context-dependent.

A conceptual diagnostic strategy



- FISH is useful when a particular locus is under investigation.
- Arrays are useful when survey-level assessment for gains/losses is required.

Application case: Unexplained developmental delay

A child has developmental delay and dysmorphic features. Conventional karyotyping reveals no abnormality.

Case A

Fictional educational case; Not a clinical record

- What question can a genome-wide microarray address?
- What class of aberration is the array primarily searching for?
- Would a negative result establish that no genetic disorder is present?

Case A: Array testing as a genomic survey

Reasoning

1

The unsolved question is whether an otherwise unrecognized genomic interval has altered copy number.

2

A genome-wide array evaluates many possible deletion or duplication intervals without a preselected locus.

3

The analysis must be integrated with phenotype; Absence of an array-detected imbalance does not exhaust every genetic possibility.

Application case: Diploid dosage, uniform genotypes

An SNP array shows a chromosomal region with apparently diploid DNA dosage but an extended absence of heterozygous SNP calls.

Case B

Fictional educational case; Not a clinical record

- Why would a CGH-only interpretation miss part of this observation?
- What does the allele pattern add to the dosage result?
- What further biological question does this raise?

Case B: Interpret copy number and zygosity jointly

Reasoning

- 1** The total amount of DNA in the region is not measurably reduced.
- 2** The allele pattern suggests regional homozygosity or loss of heterozygosity.
- 3** This pattern can warrant assessment for possible uniparental disomy in an appropriate clinical context.

Choosing among molecular genetic techniques

Clinical question	Method concept	Why it fits
One specified locus?	Targeted FISH	Directly visualizes the suspected chromosomal region
Unknown deletion/duplication?	Array CGH	Maps relative DNA dosage across the genome
CNV plus allele pattern?	SNP array	Measures copy-number and genotype-related information
Sequence-level variant?	Sequencing assay	A different information class from array dosage

Boundaries of array interpretation

- Array design and probe coverage constrain the resolution of copy-number analysis.
- CGH signal primarily reports relative DNA dosage, while SNP probes also support allelic interpretation.
- An SNP-array homozygosity pattern is not synonymous with a proven parent-of-origin mechanism.
- The textbook provides these principles, not universal laboratory thresholds for pathogenicity or all platform-specific limitations.

Three high-yield conceptual distinctions

Locus-specific vs genome-wide

Targeted localization differs from broad genomic discovery.

DNA dosage vs genotype

Copy-number and allelic status are separate measurements.

Detected pattern vs diagnosis

Genomic interpretation requires clinical and laboratory context.

Knowledge check: Interpret the readout

Consider three descriptions from an array report: (A) relative DNA depletion; (B) diploid dosage with extended homozygosity; (C) test signal enriched over reference.

Retrieval practice

Fictional educational case; Not a clinical record

- Which description supports a copy-number deletion?
- Which can indicate copy-neutral loss of heterozygosity?
- Which supports a copy-number gain?

Knowledge check: Reasoning and answers

Reasoning

- 1** A: Reduced test DNA relative to reference is consistent with a copy-number loss.
- 2** B: Normal dosage plus extended homozygosity points to copy-neutral loss of heterozygosity.
- 3** C: Increased test DNA relative to reference supports a copy-number gain.

Interpreting an unrevealing array result

- An unrevealing array result means no supported target abnormality was identified under that assay's coverage and interpretive conditions.
- A copy-number assay cannot answer every question about allelic inheritance or nucleotide sequence.
- The unresolved clinical question should determine whether a different molecular assay is appropriate.
- Interpret negative findings in the context of the phenotype and the information the method can measure.