

Medical Genetics and Cancer

Genetic Analysis of Human Diseases · The Genetic Basis of Cancer

Human genetic disease can be approached from two complementary directions. The first is diagnostic and mapping-oriented: given a family or population with a disease, how do geneticists work out its mode of inheritance and locate the responsible gene? The second concerns a single, extraordinarily important class of genetic disease in its own right - cancer - which arises when mutations accumulate in the genes that normally govern cell division, DNA repair and cell death.

1. Genetic Analysis of Human Diseases

1.1 Pedigree analysis

A **pedigree** is a diagrammatic family history that records affected and unaffected individuals across generations using standard symbols (squares for males, circles for females, filled symbols for affected individuals, a dot inside an open symbol for an unaffected carrier). Because controlled breeding experiments are impossible in humans, pedigree analysis is the primary tool for working out how a trait is inherited, by looking for characteristic patterns in which generations are affected, the sexes of affected individuals, and how the trait is transmitted from parent to offspring.

Pedigree Analysis: Recognising Modes of Inheritance

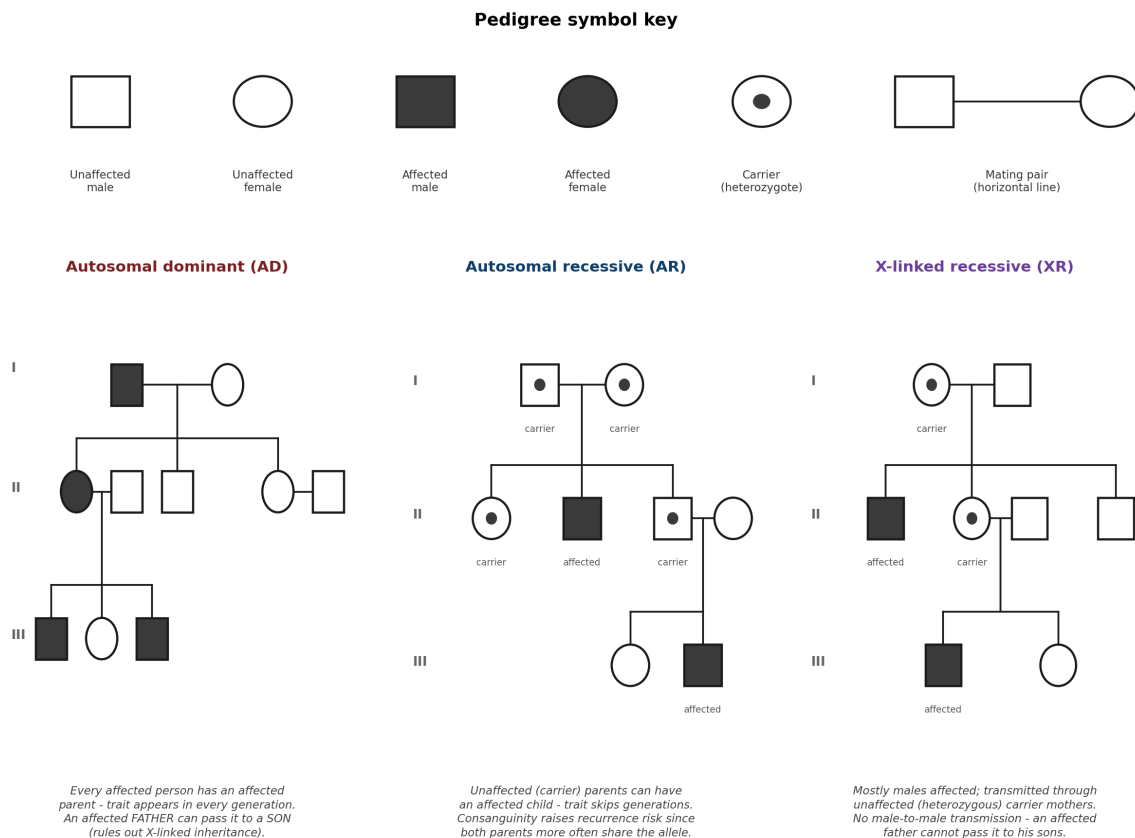


Fig. 1 - Pedigree symbols and the classic three-generation patterns produced by autosomal dominant, autosomal recessive, and X-linked recessive inheritance.

1.2 Recognising the mode of inheritance

Six inheritance patterns account for most single-gene (Mendelian) human diseases, each leaving a distinctive signature in a pedigree.

Pattern	Hallmark pedigree features
Autosomal dominant (AD)	Affected in every generation; every affected child has an affected parent; an affected father can transmit to a son; roughly equal numbers of affected males and females.
Autosomal recessive (AR)	Often skips generations; unaffected (carrier) parents can have an affected child; recurrence risk is raised by consanguinity; equal sex ratio among the affected.
X-linked dominant (XD)	Resembles AD but affected fathers transmit to ALL daughters and NO sons; typically more affected females than males (they have two X chromosomes).
X-linked recessive (XR)	Mostly males affected; transmitted through unaffected (heterozygous) carrier mothers; NO male-to-male transmission (a father cannot pass an X-linked gene to his son).
Y-linked	Only males affected; strict father-to-son transmission down every generation; the rarest pattern, but the easiest to recognise.

Pattern	Hallmark pedigree features
Mitochondrial	Transmitted only through the mother (all her children can be affected, but only her daughters transmit it further); shows variable severity due to heteroplasmy.

Key points

- The single most useful discriminating clue is often male-to-male transmission: present, it rules out every X-linked pattern; absent, together with mostly-male involvement, it strongly suggests X-linked recessive inheritance.
- A trait "skipping" a generation is the hallmark of a recessive pattern (autosomal or X-linked); a trait present in every generation points to a dominant pattern.

1.3 Mapping disease genes: linkage analysis and GWAS

Once a mode of inheritance is established, the next question is where the responsible gene actually lies in the genome. Two complementary strategies are used, suited to different kinds of disease.

- **Linkage analysis** uses multigeneration families segregating a single-gene disorder. It tracks how often a DNA marker (historically a microsatellite, now often a SNP) is co-inherited with the disease across meioses. The **recombination frequency** between marker and disease locus gives a genetic distance in **centimorgans (cM)**, and results are combined across families using the **LOD score** (log of the odds); a LOD score of 3 or more is the conventional threshold for declaring significant linkage, after which the region can be narrowed and the gene identified by **positional cloning**.
- **Genome-wide association studies (GWAS)** instead use large, unrelated case-control population cohorts. Several hundred thousand to millions of SNPs spread across the genome are genotyped, and each is tested for statistical association with the trait, exploiting **linkage disequilibrium** between a genotyped SNP and a nearby, untyped causal variant. Results are typically shown as a **Manhattan plot**, with peaks passing a strict genome-wide significance threshold marking associated loci.

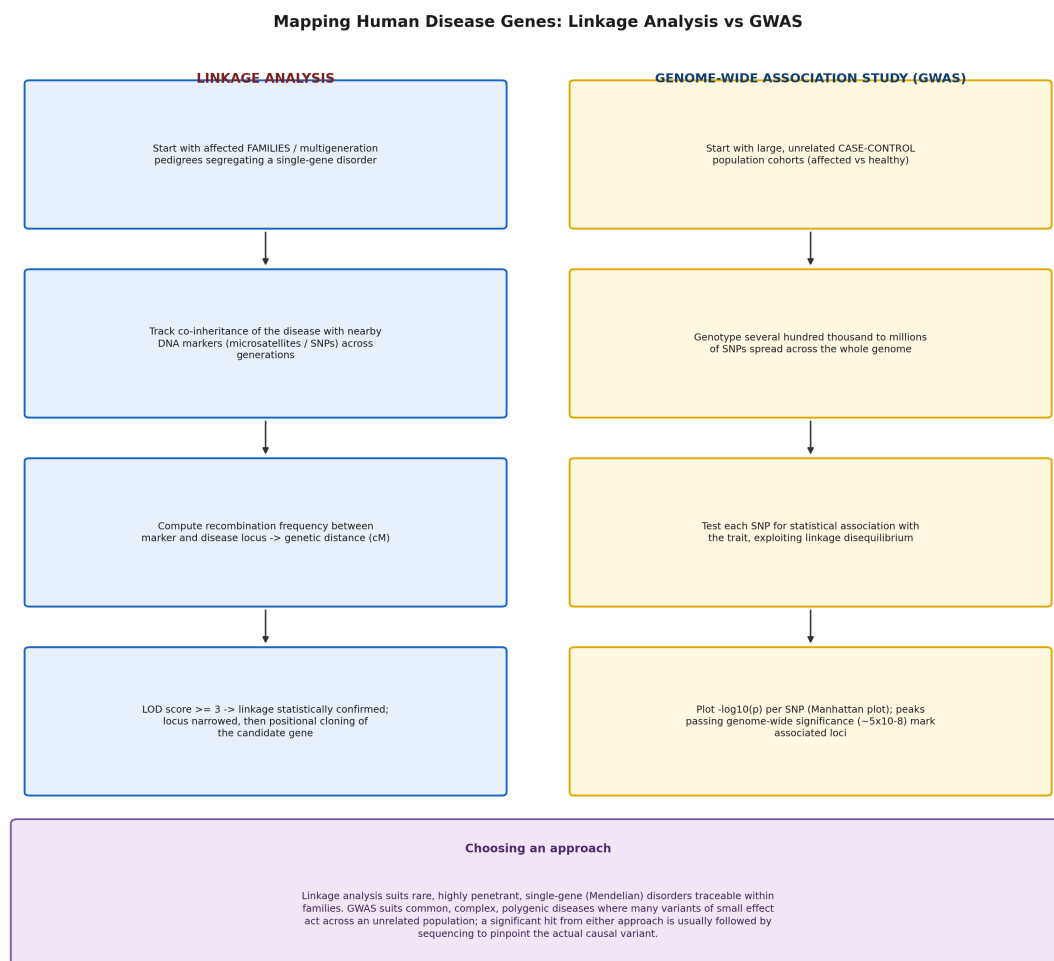


Fig. 2 - Linkage analysis (family-based, LOD score) versus GWAS (population-based, SNP association) as complementary strategies for mapping human disease genes.

Worth noting

- Linkage analysis is well suited to rare, highly penetrant, single-gene disorders that segregate clearly within families (e.g. Huntington disease, cystic fibrosis). GWAS is well suited to common, complex, polygenic diseases (e.g. type 2 diabetes, coronary artery disease) where many variants of small individual effect act together across a population; the two approaches answer different questions and are often followed by direct sequencing to confirm the causal variant.

2. The Genetic Basis of Cancer

Cancer is fundamentally a genetic disease of somatic cells: it arises when mutations accumulate in genes that control cell division, DNA repair and cell survival, allowing a single cell and its clonal descendants to escape the normal restraints on growth. Most cancer-causing mutations are somatic (acquired during a person's lifetime, e.g. from radiation, chemical carcinogens or replication errors), though inherited germline mutations in the same genes can substantially raise lifetime risk. Two broad gene classes are involved.

2.1 Oncogenes: gain-of-function drivers

A **proto-oncogene** is a normal gene that promotes cell division or growth (e.g. *RAS*, *MYC*, *ABL*, *HER2*). It becomes a cancer-causing **oncogene** through a **gain-of-function** alteration affecting just one allele: a point mutation producing a constitutively active protein, gene amplification producing extra copies and excess protein, a chromosomal translocation fusing it to a new, more active promoter or partner gene, or (in some animal cancers) retroviral insertion. Because only one mutant copy is needed, oncogenes behave **dominantly** at the cellular level. The classic textbook example is the **Philadelphia chromosome**, a reciprocal translocation t(9;22) that fuses *BCR* and *ABL1* to create a constitutively active tyrosine kinase driving chronic myeloid leukaemia.

2.2 Tumor suppressor genes and the two-hit hypothesis

A **tumor suppressor gene** normally restrains cell division or checks DNA integrity (e.g. *RB1*, *TP53*, *APC*, *BRCA1/2*). Cancer results from **loss-of-function** of such a gene, but because a single working copy is usually sufficient for normal restraint, **both** alleles must be inactivated (by mutation, deletion or epigenetic silencing) - so tumor suppressor loss behaves **recessively** at the cellular level. This is **Knudson's two-hit hypothesis**, formulated from the epidemiology of retinoblastoma: in hereditary cases, one inactivating "hit" is already present in the germline (in every cell), so only a single further somatic hit is needed - explaining the early, often bilateral onset. In sporadic cases, both hits must occur independently in the same somatic cell, a much rarer event - explaining the later, usually unilateral onset.

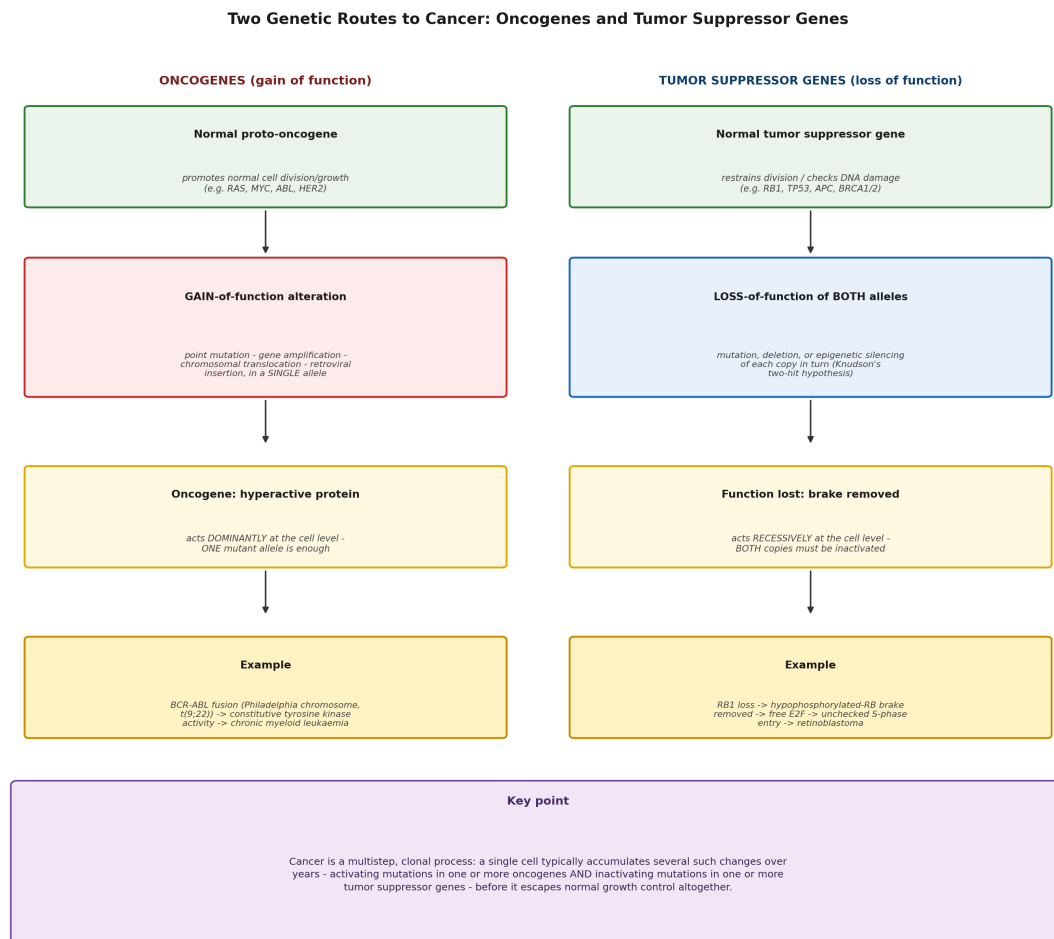


Fig. 3 - Two genetic routes to cancer: gain-of-function activation of an oncogene (dominant) versus loss-of-function inactivation of both copies of a tumor suppressor gene (recessive).

2.3 Cell cycle checkpoints: the p53-RB axis

The clearest illustration of tumor suppressor function is the **G1/S checkpoint**, which stops a cell with damaged DNA from replicating it. DNA damage activates the kinases ATM/ATR, which stabilise and activate **p53** - often called the "guardian of the genome". Activated p53 induces **p21**, which inhibits cyclin-CDK complexes; this keeps **RB** in its hypophosphorylated state, in which it binds and sequesters the transcription factor **E2F**, blocking transcription of the genes needed for S phase and arresting the cell cycle to allow repair. If damage is too severe to repair, p53 instead tips the balance of BCL2-family proteins towards **apoptosis**, eliminating the damaged cell altogether. Loss of **TP53** or **RB1** function removes this checkpoint, allowing damaged cells to keep dividing and accumulate further mutations - a key step in tumor progression.

The Two-Hit Hypothesis and the p53-RB Checkpoint

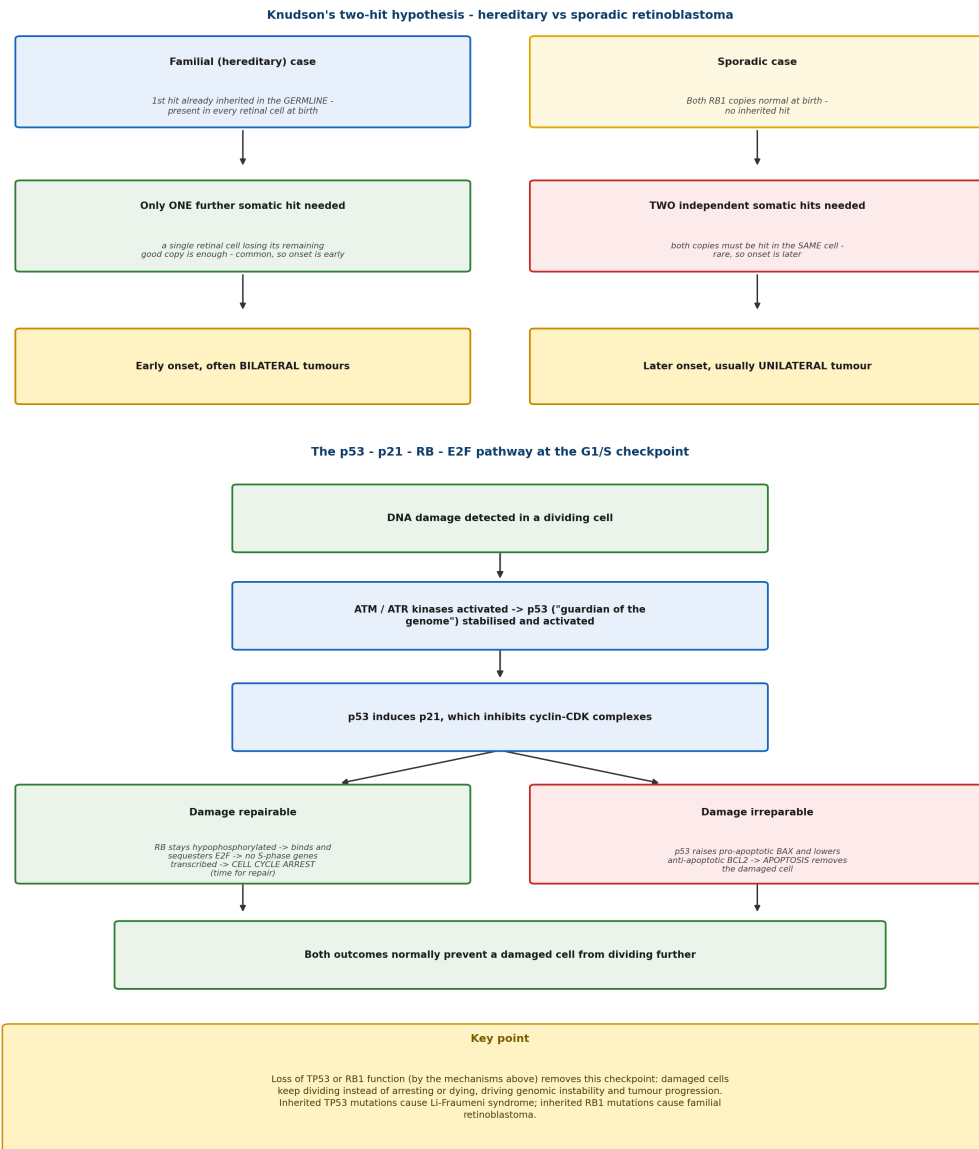


Fig. 4 - Knudson's two-hit hypothesis applied to hereditary versus sporadic retinoblastoma, and the p53-p21-RB-E2F pathway that enforces the G1/S checkpoint.

2.4 Hereditary cancer syndromes

Germline mutations in tumor suppressor genes cause well-characterised hereditary cancer syndromes, in which every cell already carries the "first hit" and only a single further somatic event is needed to initiate a tumor.

Syndrome	Gene	Typical tumors
Familial retinoblastoma	RB1	Early-onset, often bilateral retinoblastoma
Li-Fraumeni syndrome	TP53	Wide range of cancers at unusually early ages (sarcomas, breast, brain, adrenal)
Familial adenomatous polyposis (FAP)	APC	Thousands of colonic polyps; near-certain colorectal cancer if untreated
Hereditary breast/ovarian cancer (HBOC)	BRCA1 / BRCA2	Substantially elevated lifetime breast and ovarian cancer risk

Key points

- Oncogenes are activated (gain of function, dominant, one hit); tumor suppressor genes are inactivated (loss of function, recessive, two hits) - opposite mechanisms converging on the same outcome of unchecked proliferation.
- The p53-RB checkpoint pathway is the cell's principal brake on damaged cells entering S phase; its two central genes, *TP53* and *RB1*, are among the most frequently altered genes across all human cancers.
- Hereditary cancer syndromes are a direct clinical application of the two-hit hypothesis: inheriting one "hit" in every cell converts a rare two-event process into a near-inevitable one.

3. Quick Revision Summary

Concept	Essence
Pedigree analysis	Family-history diagram used to infer a trait's mode of inheritance
Autosomal dominant	Affected every generation; affected parent needed; father can pass to son
Autosomal recessive	Skips generations; unaffected carrier parents can have affected child
X-linked recessive	Mostly males; carrier mothers transmit; no male-to-male transmission
Linkage analysis / LOD score	Family-based mapping; LOD ≥ 3 confirms linkage; leads to positional cloning
GWAS	Population case-control SNP association; suits common, polygenic disease
Proto-oncogene -> oncogene	Gain-of-function change (mutation/amplification/ translocation); acts dominantly
Tumor suppressor gene loss	Loss-of-function of BOTH alleles; acts recessively at the cell level
Two-hit hypothesis	Knudson: hereditary cancer needs 1 somatic hit; sporadic needs 2 -> earlier/bilateral onset
p53	"Guardian of the genome"; triggers p21-mediated arrest or apoptosis after DNA damage
RB / E2F	Hypophosphorylated RB sequesters E2F, blocking S-phase gene transcription
Hereditary cancer syndromes	Germline TSG mutation (RB1, TP53, APC, BRCA1/2) supplies the first hit in every cell

Notes prepared for study purposes. Consult current primary literature for details required in formal citation or research use.