CHAPTER 5

Genetic Testing Techniques

ALICIA GOMES, MS • BRUCE KORF, MD, PHD

INTRODUCTION

Genetic testing consists of the analysis of DNA (deoxyribonucleic acid), or in some cases RNA (ribonucleic acid) transcribed from DNA, for variations that are associated with disease or risk of disease. DNA is the molecule responsible for transmission of genetic information in the cell. It consists of a double helix comprised of a sugar-phosphate chain, with the two strands bridged by hydrogen bonding between the four nucleoside bases—adenine (A), guanine (G), cytosine (C), and thymine (T). Base pairing is specific: A always binds with T on the opposite strand, and C binds with G. The order of bases determines the coding sequence, which is copied into a complementary RNA molecule (consisting of a single-stranded ribose-phosphate chain with the same four bases, except that uracil [U] substitutes for thymine). This messenger RNA is then exported to the cytoplasm, where RNA associates with ribosomes to direct the synthesis of proteins, with triplets of bases encoding specific amino acids. The genetic information, consisting of 3.15 billion base pairs of DNA in each haploid genome, is tightly compacted into the cell nucleus and divided into 23 chromosomes, including 22 nonsex chromosomes (autosomes), and the X and Y sex chromosomes (Fig. 5.1).

The essence of genetics, and the basis for genetic testing, is genetic variation. This can occur at multiple levels, from abnormalities at the level of the entire chromosome down to changes at the single nucleotide level. In this chapter we will consider the various methods available for detection of this genetic variation, along with pitfalls in interpretation of the results of genetic tests.

SOMATIC VERSUS GERMLINE GENETIC VARIATION

Every individual begins with a single pluripotent zygote that replicates into over 37 trillion differentiated cells. If a genetic variant is present in the zygote, it will be replicated in every cell, including future germ cells. This is classified as a **germline variant**. Germline variants can be without phenotypic effect or can result in phenotypes that may include specific genetic syndromes and disorders. Germline variants are associated with a risk of transmission to future offspring and can be transmitted in either a recessive or a dominant pattern. Germline variants can be identified in any tissue specimen, as the variant is represented in all cells in the body. Peripheral blood is commonly used for germline genetic testing, as it is readily accessible for collection and provides an ample source of genetic material within white blood cells.

Genetic variants may also be acquired by a new mutation in somatic cells. Such postzygotic mutations result in the phenomenon of **mosaicism**. The phenotypic effects of a mosaic variant depend on the developmental stage at which the mutation occurred, the gene involved, and the cell lines affected. Mosaicism can result in a fully expressed phenotype, a mildly expressed phenotype if only a small proportion of cells are involved, or a phenotype confined to a single region of the body. In some cases, a mutation occurs during germ cell development, resulting in **germline mosaicism**; in such cases, there is no phenotype, but genetic transmission to multiple offspring may occur (Fig. 5.2).

Throughout life, additional mutations can be acquired somatically, either spontaneously or because of exposure to mutagenic agents, such as chemicals or radiation. This may have no phenotypic effect, but in some cases mutation of a tumor suppressor gene or oncogene can result in malignancy.

CYTOGENETIC VERSUS MOLECULAR GENETIC TESTING

Chromosomes were discovered in the 1840s in plant cells by Karl Wilhelm von Nageli. The identification of mitosis in animal cells by Walther Flemming in the 1880s paved the way to recognition of the chromosome as a unit of inheritance in the cell. Routine human cytogenetic analysis was developed in the mid-1950s when it was discovered that hypotonic treatment could be used to swell cultured cells, spreading chromosomes at metaphase, the time of greatest condensation, to permit analysis. Subsequent advances included the use of phytohemagglutinin

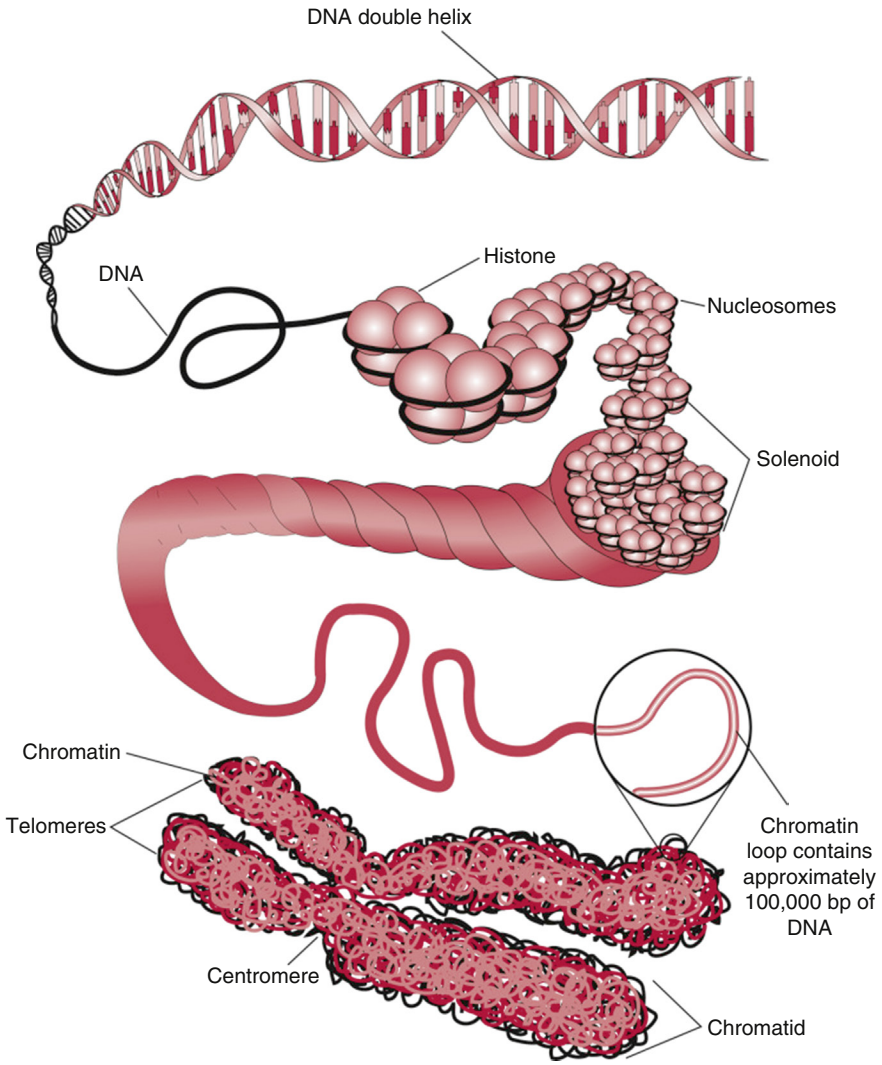


FIG. 5.1 DNA organized within a chromosome. (Courtesy of Jorde LB et al. 2009 Medical Genetics, 4th edn. Elsevier, Edinburgh.)

to stimulate T cell proliferation in peripheral blood cultures and development of staining techniques to elicit characteristic banding patterns that enable precise identification of each chromosome and detection of subtle structural variants. Beginning in the 1980s, fluorescence in situ hybridization (FISH) was developed, enabling detection of specific labeled segments of DNA hybridized to their homologous sequences and consequent detection of deletions or duplications (**copy number variants**). Molecular cytogenetic technologies were further developed in the 2000s with the advent of cytogenomic microarrays, based on hybridization of DNA in a test sample against an array of oligonucleotides on a "chip," permitting high-resolution detection of copy number variants down to a few thousand base pairs.

In contrast with these high-level analyses of structural variants or copy number changes, molecular analyses began to be developed to detect variants down to the level of a single nucleotide. Detection of variants at the DNA level has also undergone evolution, including the development of sequencing technologies applicable to specific genes (e.g., Sanger sequencing) and now toward sequencing of the entire genome using massively parallel ("next-generation") technologies. The scale of detection for molecular analysis ranges from the single nucleotide level to an entire gene or group of

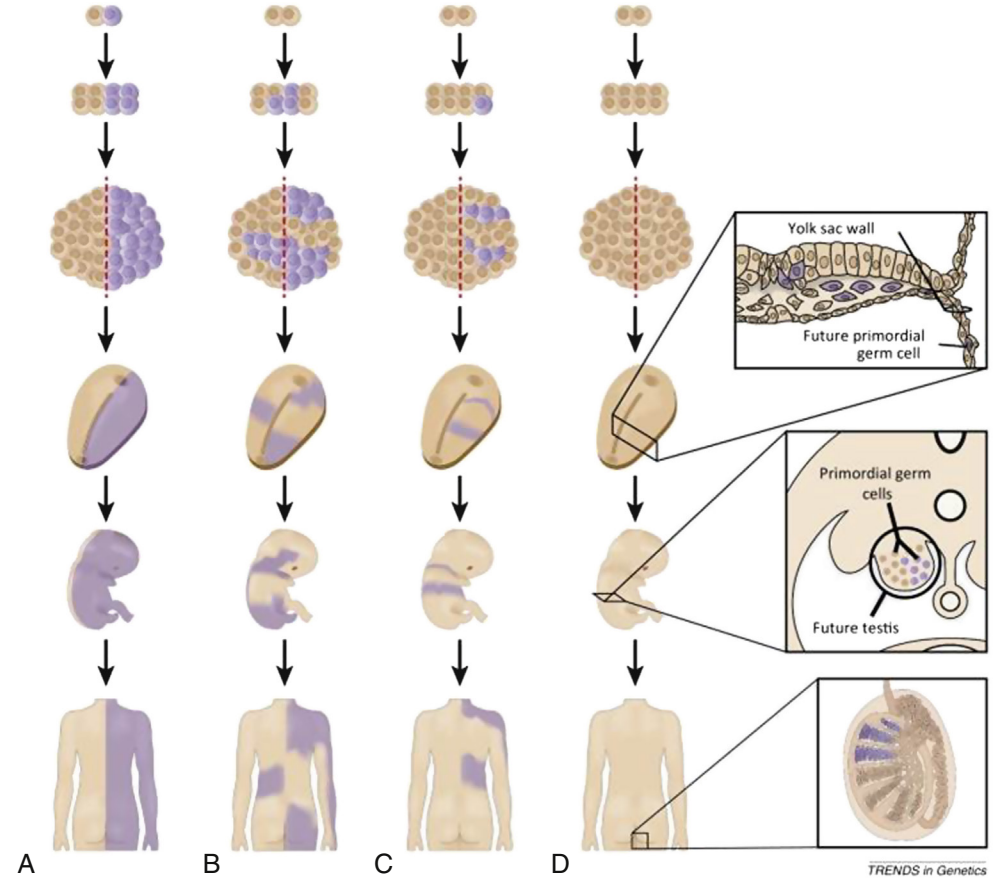


FIG. 5.2 Types of mosaicism. (Courtesy of Campbell I, Shaw C, Stankiewicz P, et al. Somatic mosaicism: implications for disease and transmission genetics. *Trends Genet*. July 2015;31(7):382–392.) Examples A-C describe possible phenotypic severity outcomes based on the timing of the mutation during embryogenesis. Example D describes the unique phenomenon of germline mosaicism when a mutation is isolated to a percentage of the primordial germ cells.

genes, forming now a continuum between cytogenetic/cytogenomic to molecular analysis.

CLASSIFICATION OF GENETIC VARIATION

Genetic variation can extend from the level of the single nucleotide up to the entire chromosome. Table 5.1 provides a classification of the different types of genetic variants.

CYTOGENETIC ANALYSIS

Karyotype

Methodology

Karyotyping involves analysis of the entire chromosome complement through the microscope. Dividing cells are harvested during metaphase, the time of greatest chromosome condensation, by disruption of the spindle using drugs such as colchicine. Chromosomes are visualized by staining, including the use of special stains to elicit banding patterns (Table 5.2). In general GC-rich regions tend to be gene rich and stain darkly with G-banding, bright with R-banding, and dark with Q-banding[1] (Fig. 5.3).

Karyotyping is able to detect polyploidy, aneuploidy, translocations, inversions, rings, and copy number changes in the size range of 4–6 Mb; smaller copy number changes require the use of molecular cytogenetic techniques.[1]

Types of variants detected

- Polyploidy
- Monosomies/trisomies
- Inversions
- Translocations
- Large microdeletions/duplications

TABLE 5.1
Classification of Genetic Variants

Variant Type	Definition	Example
CHROMOSOME LEVEL		
Polyploidy	An entire extra set of chromosomes	Triploidy, having three haploid chromosome sets
Aneuploidy (monosomy/trisomy)	An entire missing (monosomy) or extra chromosome (trisomy)	Trisomy 21, Turner syndrome
Translocation	An exchange of genetic material between two chromosomes. This can be balanced (no missing material) or unbalanced (partial deletions or duplications)	Burkitt's lymphoma due to a translocation of chromosomes 8 and 14: t(8;14)(q24;q32)
Inversion	A segment within a chromosome that has been reversed. It may include (pericentric) or not include (paracentric) the centromere	Inversion within chromosome 16 seen in an aggressive subtype of pediatric acute megakaryoblastic leukemia: inv(16)(p13.3q24.3)
Copy number variant	Deletion or duplication of DNA segments	21q11.2 microdeletion associated with DiGeorge syndrome; deletion of multiple exons in dystrophin gene resulting in Duchenne muscular dystrophy
NUCLEOTIDE LEVEL		
Deletion/duplication	Any combination of loss or gain of one or several nucleotides; a deletion or duplication that is not an integral multiple of three bases results in a frameshift	One base deletion resulting in frameshift
Nonsense	A base substitution that creates a new stop signal	C to T change in arginine codon, resulting in stop
Missense	A base substitution that alters the amino acid encoded at a specific codon	Sickle cell variant in β-globin
Silent	A base substitution that does not alter the amino acid at a codon	A single nucleotide polymorphism common in the general population
Splicing variant	A base substitution in an exon or intron that alters RNA splicing	Splice donor variant that results in exon skipping

TABLE 5.2
Banding Techniques and Stains

Banding Technique	Stain	Clinical Application
G (Giemsa)-banding	Giemsa	Most commonly used for karyotyping. Used as the reference for banding points. Gene-rich regions stain darkly
R (reverse)-banding	Acridine orange	Gene-rich bands fluoresce brightly
C (centromere)-banding	Giemsa stain with alkali, acid, salt, or heat treatment	Identifies heterochromatin associated with each centromere
Q (quinacrine)-banding	Quinacrine	AT-rich regions fluoresce brightly

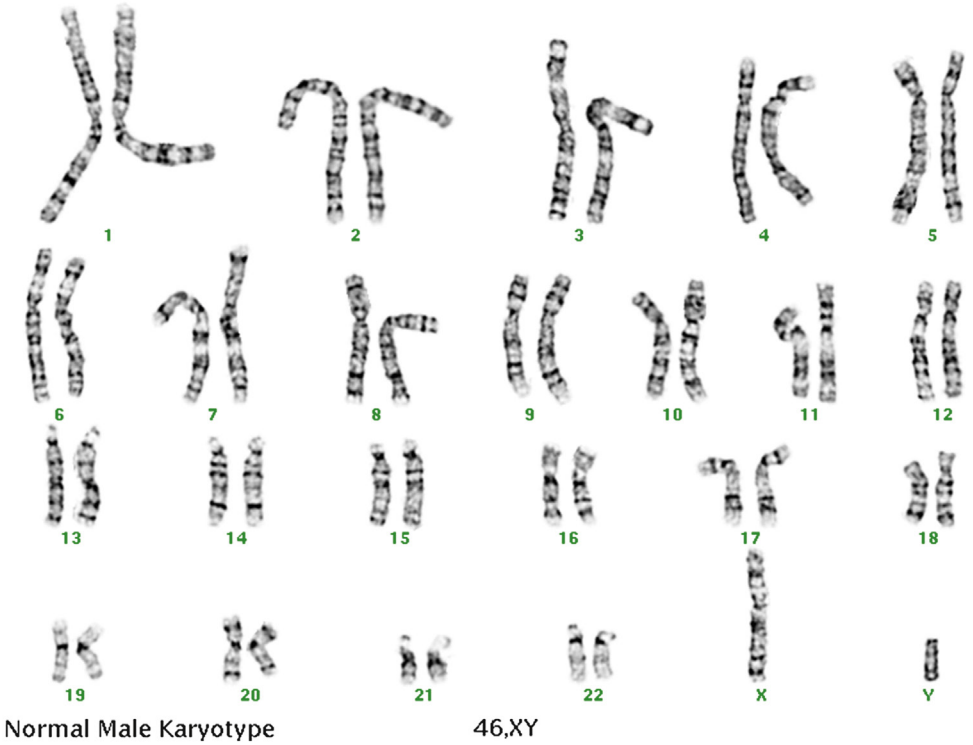


FIG. 5.3 Normal male karyotype. (Courtesy of Dr. Fady Mikhail, MD, PhD, FACMG. University of Alabama at Birmingham Cytogenetics Laboratory.)

Benefits

Within the germline, there are syndromes caused by large structural rearrangements that are known to have an increased predisposition to tumors. Individuals with Down syndrome, associated with trisomy 21, have a 10- to 20-fold risk for acute lymphocytic leukemia, acute myelocytic leukemia, and acute megakaryocytic leukemia.[2] Individuals with Klinefelter syndrome, a condition in males caused by having an extra copy of the X chromosome (XXY sex chromosomes), are at an increased risk for breast cancer, extragonadal germ cell tumors, non-Hodgkin lymphoma, and lung cancer.[3] Individuals with Turner syndrome, caused by having a single copy of the X chromosome in females, have an increased risk for Wilms' tumor, leukemia, gonadal tumors, neurogenic tumors, and uterine tumors in those taking unopposed estrogen.[4]

Karyotypes can be used in somatic cells for treatment and prognosis. Many tumors and cancers acquire chromosomal changes as they progress. A karyotype can be performed on dividing cancer cells to help in classification, as certain chromosomal changes are characteristic of specific cancers.

Limitations

Chromosome analysis will not identify all genetic anomalies. When a deletion or duplication is smaller than 4–6 Mb in size, the variant may be missed by karyotypic analysis. In addition, chromosomal analysis can be challenging in cancer cells, where morphologic preservation may be less than in somatic cells. Accuracy of chromosomal analysis is also heavily dependent on the skill of the technologist. Lastly, mosaicism may be an important factor, as low level mosaicism may be missed. The number of cells routinely analyzed is dependent on the tissue type, whether the specimen has been cultured, and the purpose of the analysis.

Fluorescence In Situ Hybridization

Methodology

Fluorescence In Situ Hybridization (FISH) analysis is performed by denaturing the double-stranded DNA in the fixed chromosomes on a microscope slide. Once denatured, two fluorescently labeled DNA probes are used in combination to analyze each location in question. The first probe serves as a control and hybridizes with DNA on the target chromosome but outside of

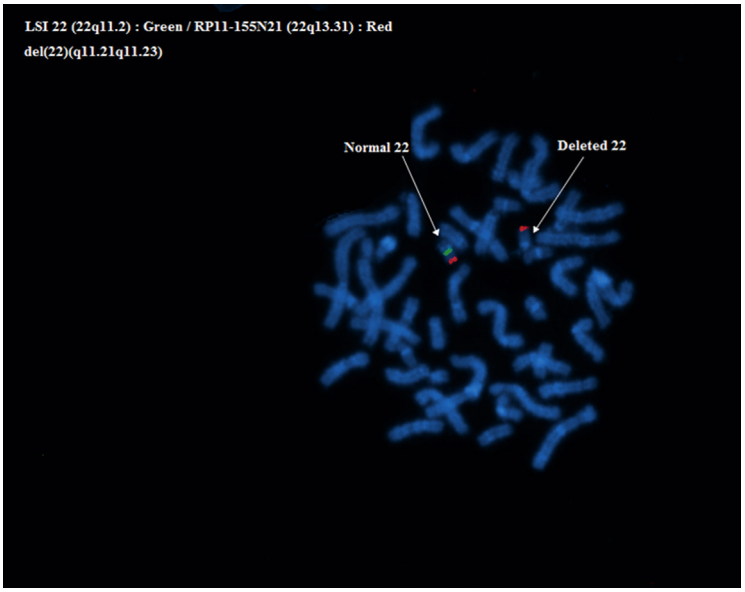


FIG. 5.4 Fluorescence in situ hybridization (FISH) probe notating a deletion on the long arm of chromosome 22. (Courtesy of Dr. Fady Mikhail, MD, PhD, FACMG. University of Alabama at Birmingham Cytogenetics Laboratory.)

the targeted region. The second probe hybridizes to a target location on the individual's DNA sequence. When the sequence is present, the probe will hybridize and fluoresce with a different color than the control probe.

When a deletion is present, the second probe will not hybridize and no fluorescence will be seen. A duplication will result in two fluorescent spots with the test probe (Fig. 5.4).

Types of variants detected

- Monosomy/trisomy
- Balanced/unbalanced translocations
- Microdeletions/duplications

Benefits

FISH analysis can be helpful in germline analysis of large deletions and duplications. This testing process has a quick turnaround of typically 2–3 business days. This enables an immediate clinical confirmation in time-sensitive medical situations. FISH is also used in diagnosis and follow-up of cancer. FISH probes have been designed for most common deletions, duplications, and translocations that have been identified in many cancer types where the mutation in question will provide insights into how well the tumor or cancer in question will respond to a certain therapeutic agent. Some laboratories offer probe panels that include several FISH probes for the most common regions analyzed and offer a simultaneous analysis of the genetic material. FISH analysis can also be very helpful in the identification of mosaicism, as several cells can be analyzed simultaneously in one analysis.

Limitations

FISH analysis requires an adequately sized tissue sample for analysis. It is also necessary to predetermine the target for analysis; if there is a copy number change outside the target region, it will not be detected.

Microarray

Methodology

DNA microarrays can be used to analyze the expression or copy number of multiple genes throughout multiple regions of the genome simultaneously. The following are a list of types of DNA microarrays available and their applications[5]; however, we will only focus on two approaches in this chapter (Table 5.3).

The two most frequently utilized DNA microarray technologies for genotyping are the comparative genomic hybridization (CGH) microarray and the single nucleotide polymorphism (SNP) array. The CGH microarray uses a small plate of glass (chip), typically less than an inch in size. The chip contains a grid consisting of thousands of probes to specific segments within the human genome. The test sample DNA is digested to create fragments and a fluorescent dye is added to the specimen.

TABLE 5.3
Microarrays and Applications

Microarray Name	Application
Gene expression profiling	Observes expression levels of thousands of genes simultaneously. Useful when analyzing the treatment and advancement of diseases and choice of therapeutics
Comparative genome hybridization (CGH)	Compares the genetic content of two similar specimens; used to identify copy number changes in a test sample as compared with a reference sample
Single nucleotide polymorphism (SNP)	Measures hybridization with oligonucleotide probes that recognize SNPs across the genome; useful in detecting copy number changes or loss of heterozygosity
Exon junction array	Uses fewer probes per gene to assess the expression of alternative splicing of a gene in a sample
Fusion gene	Detects fusion transcripts resulting from translocations commonly found in cancer specimens for treatment and prognosis
Multistranded DNA	Used in the detection of novel drugs that may inhibit the expression of particular genes
Bacterial artificial chromosomes (BAC)	Oligonucleotide probes consisting of BACs used to target specific regions of the genome for copy number analysis

The sample is combined with a reference DNA sample that has been digested into fragments and labeled with a different fluorescent tag. The combined specimen is separated into single-stranded DNA and washed over the chip of probes so that the specimens can compete with one another to hybridize with the available DNA probes. The results are analyzed by determining the ratio of test DNA to reference DNA at each available probe based on the visible fluorescence.

SNP microarray technology also takes place on a small plate of glass (chip). The chip contains tens of thousands of regions on which unique DNA probes sit to target specific areas of the genome. Each probe contains a small sequence (~25 base pairs) that surrounds an established SNP within the genome. A probe is created to represent each genotype. For example, if there is a SNP that causes the C allele to be replaced by a T:

…..ATGCTGT**C**TGTTAC……

…..ATGCTGT**T**TGTTAC……

A probe will be located on the chip with the complementing G allele to hybridize to individuals with the C allele in their DNA sequence, and a separate probe will be located on the chip with the complementing A allele to hybridize to individuals with the T allele in their sequence.

To perform the analysis, DNA is extracted from an individual. The DNA for each locus is denatured, amplified, and digested. Biotin is also added to enable the individual's DNA sample to adhere to a fluorescent marker later in the process. The individual's DNA is washed over the probe-laden chip for >12 h to provide ample opportunity for hybridization between the test DNA and the probes. The chip is then rinsed to remove the excess DNA that did not hybridize with the regions captured by the probes. A second wash with a fluorescent marker is then added to adhere to the individual DNA that has hybridized to the probes. Any excess stain that does not bind to a DNA sample is removed. A laser then scans the reaction to measure the amount of fluorescence that is created at each SNP region. As a probe is available for each possible genotype for each SNP location, individuals who are heterozygous or homozygous for a specific location can be identified. Further analysis for copy number variation is performed by normalizing the fluorescence produced by the individual's sample to a reference sample. Copy number loss or gain, uniparental disomy, and large regions of homozygosity can be visualized using this technology (Fig. 5.5).

Types of variants detected

- Monosomy/trisomy
- Unbalanced translocations
- Microdeletions/duplications
- Large deletions/duplications

Benefits

Many genetic testing laboratories use a microarray technology to identify copy number changes for the genes analyzed in their panels. This can be very helpful for patients presenting with a phenotype that overlaps with several possible genetic etiologies. An array allows for the simultaneous analysis of multiple genes for identification of copy number variants.

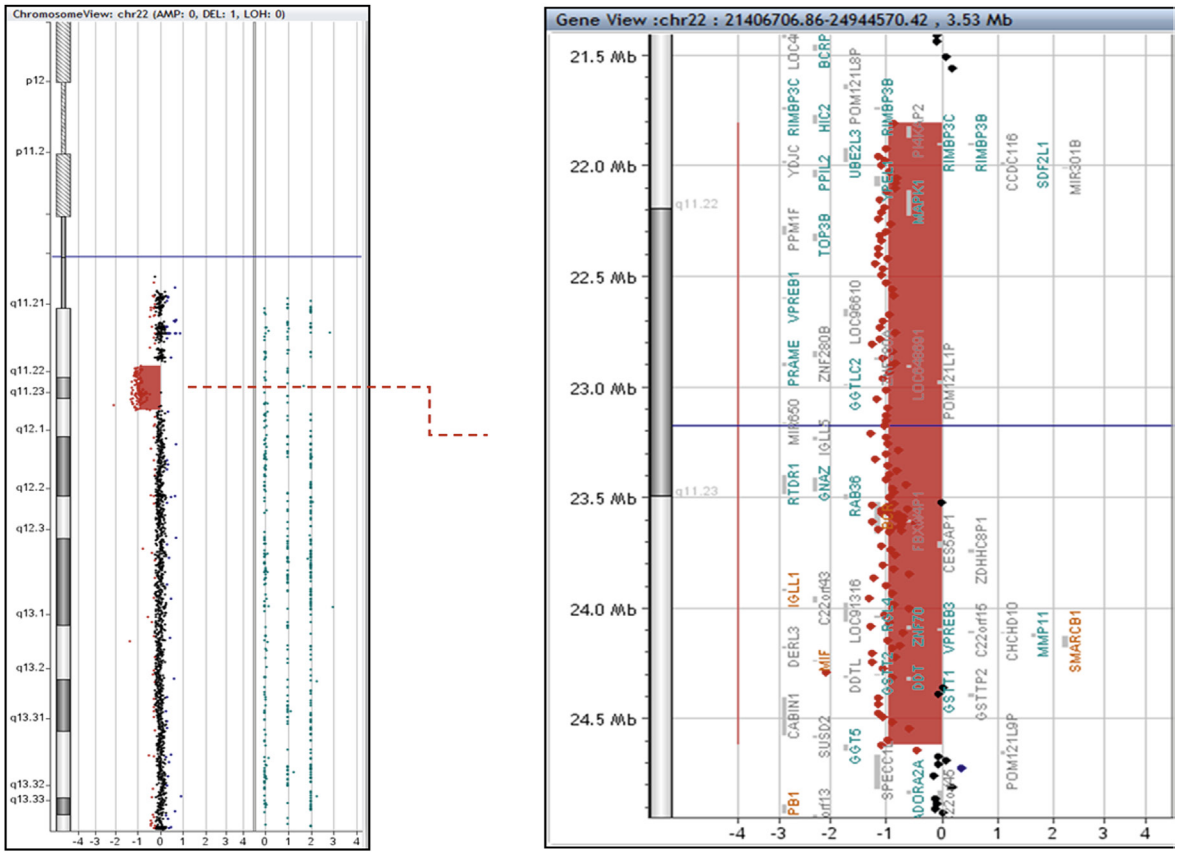


FIG. 5.5 Single nucleotide polymorphism array results showing a deletion along long arm of chromosome 22. TypeIII distal 22q11.2 microdeletion (~2.8 Mb in size) that spans the *SMARCB1* gene (MIM: 611867). (Courtesy of Dr. Fady Mikhail, MD, PhD, FACMG. University of Alabama at Birmingham Cytogenetics Laboratory.)

Microarray-based platforms are highly sensitive in detecting copy number variations ranging from a single exon to aneuploidy/polyploidy. This can be very helpful in classifying a tumor and in guiding treatment and prognosis.

Limitations

Microarray technology can be expensive when compared with other technologic approaches that could be used to identify similar alterations. Microarray technology is also limited in the mutational spectrum that can be identified. While SNP arrays can be used to identify various forms of copy number change, it is unable to identify copy number variants smaller than ~80kb. Balanced translocations are also not identified, which is a limitation in cancer diagnosis. Mosaicism detection is also a limitation of microarrays because a mutation present in fewer than 20%–30% of the cells may be missed by this approach.[6] This can be important when dealing with a starting material containing a mixed etiology or if an individual only harbors a mutation in a small percentage of their cell lines.

Multiplex Ligation-Dependent Probe Amplification

Methodology

Multiplex Ligation-dependent Probe Amplification (MLPA) is a polymerase chain reaction (PCR)-based technique that analyzes several loci simultaneously for small copy number changes, usually involving one or several exons. MLPA begins with a denaturation of double-stranded DNA into single strands and hybridization to MLPA probes located at distinct intervals

throughout the exons of the gene in question. Each probe consists of two separate oligonucleotides (one at the 5′ and one at the 3′ end) of the interval to be tested. A ligation reaction follows when the adjacent probes are able to hybridize with their target sequences. In addition to the gene-specific oligonucleotides, each probe also includes a universal primer sequence that enables a simultaneous PCR reaction to take place. All of the ligated probes in the assay amplify the ligated probes and then each PCR product is separated using capillary electrophoresis. The products are then analyzed by measuring the fluorescence produced by the amount of PCR product that is made after the amplification compared with control peaks that assess the quality of the amplification step and standardizes the amount of DNA present at each locus (Fig. 5.6).

Types of variants detected

- Single/multiexon deletions
- Single/multiexon duplications

Benefits

MLPA is analyzed and interpreted by the presence of PCR product measured in the patient sample in comparison with control peaks. MLPA analysis can detect homozygous or hemizygous deletions and duplications. The probes used for MLPA analysis can detect copy number changes involving as few as 60 nucleotides, enabling identification of single exon deletions. MLPA technology is useful in cancer genetics for both somatic and germline testing. Many germline tumor and cancer predisposition syndromes present with a spectrum of mutations, including copy number variants. It is important to review the mutational spectrum known for each condition in consideration and compare this with the testing options available for each laboratory's testing strategy.

Limitations

MLPA is only capable of identifying copy number variations. Furthermore, identifying a deletion or duplication will not define the exact break points of the mutation; MLPA can only certify that a deletion or duplication is present. If a variant has break points that extend past the last probe deleted in an assay but end before the next probe analyzed, it cannot be concluded where the deletion starts/stops within this region. MLPA is also gene/region-specific. For example, if an MLPA analysis concludes that there is a duplication of all the probes in the assay, it cannot be determined if the duplication is really due to a chromosomal anomaly or a duplication involving multiple genes. Additional technology must be used to define these details. Lastly, MLPA has limited ability to identify a mosaic deletion/duplication if the variant is present in fewer than 30% of the cells analyzed.[7]

Sanger Sequencing

Methodology

Sanger sequencing is a targeted sequencing technique that uses oligonucleotide primers to seek out specific DNA regions. Sanger sequencing begins with denaturation of the double-stranded DNA. The single-stranded DNA is then annealed to oligonucleotide primers and elongated using a mixture of deoxynucleotide triphosphates (dNTPs), which provide the needed arginine (A), cytosine (C), tyrosine (T), and guanine (G) nucleotides to build the new double-stranded structure. In addition, a small quantity of chain-terminating dideoxynucleotide triphosphates (ddNTPs) for each nucleotide is included. The sequence will continue to extend with dNTPs until a ddNTP attaches. As the dNTPs and ddNTPs have an equal chance of attaching to the sequence, each sequence will terminate at varying lengths.

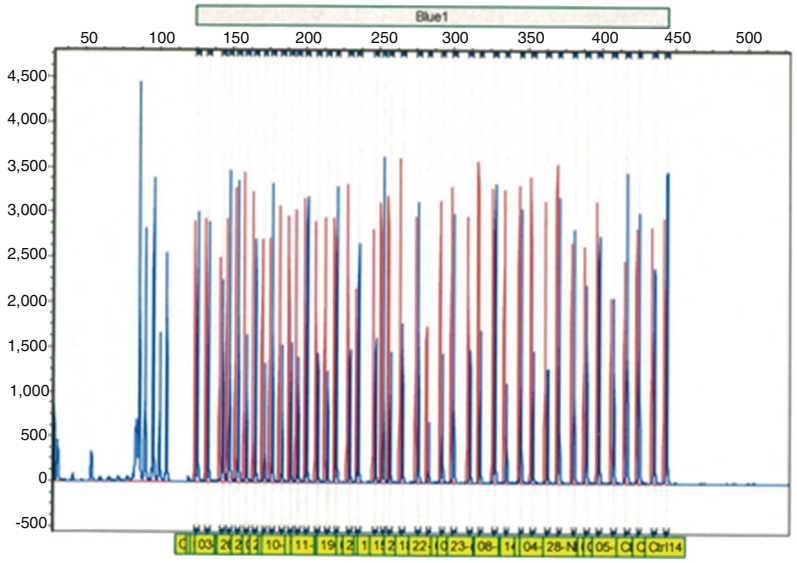
Each ddNTP (ddATP, ddGTP, ddCTP, ddTTP) also includes a fluorescent marker. When a ddNTP is attached to the elongating sequence, the base will fluoresce based on the associated nucleotide. By convention, A is indicated by green fluorescence, T by red, G by black, and C by blue. A laser within the automated machine used to read the sequence detects a fluorescent intensity that is translated into a "peak." When a heterozygous variant occurs within a sequence, loci will be captured by two fluorescent dyes of equal intensity. When a homozygous variant is present, the expected fluorescent color is replaced completely by the new base pair's color (Fig. 5.7).

Types of variants detected

- Silent
- Missense
- Nonsense
- Truncating
- Deletion
- Insertion
- Splicing

Benefits

Sanger sequencing is a robust testing strategy able to determine whether a point mutation or small deletion/duplication is present. It has been widely used for several decades in many settings, including defining the mutational spectrum of a tumor as well as identifying a constitutional variant in diagnostic testing. Primers can be created to cover several regions (amplicons) to cover any size region of interest.



	Probe Name	Bin Size	Height Ratio
1	01-TBX1	177.7	1.219
2	02-TBX1	253.0	1.162
3	03-DGCR8	148.6	1.184
4	04-SNAP29	371.8	0.896
5	05-LZTR1	417.7	1.395
6	06-PPIL2	408.6	1.000
7	07-GNAZ	143.7	0.895
8	08-exon1	336.3	0.341
9	09-exon1	311.7	0.500
10	10-exon2b	195.3	0.459
11	11-exon2b	215.1	0.419
12	12-exon3	160.4	0.474
13	13-exon3	257.9	0.454
14	14-exon4	354.2	0.428
15	15-exon4	265.9	0.490
16	16-exon5	190.6	0.521
17	17-exon5	208.7	0.496
18	18-exon6	283.8	0.390
19	19-exon6	230.3	0.445
20	20-exon7	363.9	0.406
21	21-exon7	184.4	0.497
22	22-exon8	293.0	0.460
23	23-exon8	318.6	0.472
24	24-exon9	172.6	0.491
25	25-exon9	247.7	0.574
26	26-SNRPD3	166.2	0.838
27	27-SEZ6L	276.6	1.056
28	28-NIPSNAP1	390.4	0.835
29	Ctrl01	126.6	1.032
30	Ctrl02	134.1	0.984
31	Ctrl03	154.5	1.021
32	Ctrl04	201.3	1.005
33	Ctrl05	221.4	1.120
34	Ctrl06	236.2	1.233
35	Ctrl07	300.6	0.908
36	Ctrl08	328.4	1.010
37	Ctrl09	346.2	0.923
38	Ctrl10	381.9	1.058
39	Ctrl11	399.0	0.878
40	Ctrl12	425.9	1.061
41	Ctrl13	436.2	0.839
42	Ctrl14	444.6	1.177

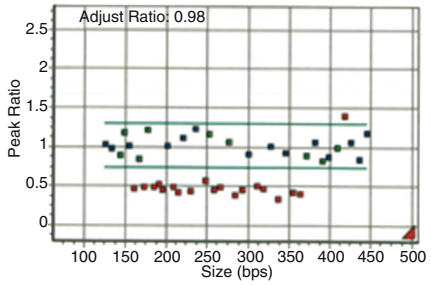


FIG. 5.6 Graphic representation of a Multiplex Ligation-dependent Probe Amplification (MLPA) analysis showing a multi exon deletion. (Courtesy of Dr. Ludwine Messiaen, PhD, FACMG. University of Alabama at Birmingham Molecular Genomics Laboratory.)

Limitations

Although one could use individual Sanger sequencing reactions to cover any desired region, this testing approach can be costly when compared with other multiplex testing systems. Therefore, most currently available Sanger sequencing tests are gene-specific or analyze a small subset of genes. Sanger sequencing is able to identify mosaic mutations including as low as 20% of the cells, but Sanger sequencing is not precisely quantifiable. For example, one cannot conclude if a mutation is present in 25% versus 40% of cells based on peak sizes; additional testing strategies must be used for quantification.

Next-Generation Sequencing

Methodology

"Next-generation" sequencing (NGS) is a high-throughput technology that enables simultaneous sequencing of multiple DNA segments in a sample. This analysis is accomplished by parallel sequencing of small fragments

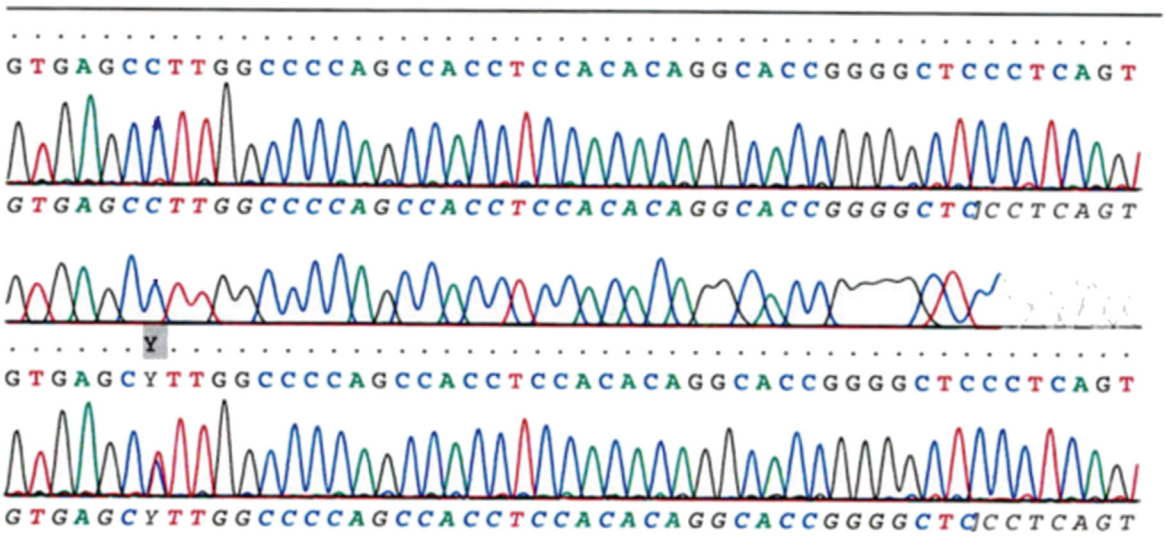


FIG. 5.7 Electropherogram from Sanger sequencing of a nucleotide change from C to T (mutation noted with a Y) compared to sequencing of normal contral samples. This mutation is a heterozygous mutation as both alleles harbor a different nucleotide. (Courtesy of Dr. Ludwine Messiaen, PhD, FACMG. University of Alabama at Birmingham Molecular Genomics Laboratory.)

and aligning these to a reference sequence. Different platforms use different sequencing approaches, with current major platforms summarized in Table 5.4. NGS can be used to sequence the entire complement of DNA in a sample (whole genome sequencing—WGS) or specific segments can be isolated for sequencing. These can include all of the exons (whole exome sequencing—WES), or specific regions of interest, creating "panels" of genes to be sequenced.[8]

Types of variants detected

- Silent
- Missense
- Nonsense
- Truncating
- Deletion
- Insertion
- Splicing

Single gene versus panel versus whole exome/genome testing

One major benefit of NGS is the ability to test tens to hundreds to thousands of genes simultaneously. This is especially helpful in testing individuals who present with an ambiguous clinical presentation in which multiple genes could be the underlying cause. Panel-based testing eliminates the time and financial strain required to sequence several individual genes until the right gene is identified. Panel-based testing is also especially helpful when any of several genes may be known to cause a specific disorder.

NGS testing strategies also produce multiple, individually produced readings of the target area compared with Sanger sequencing, which only provides one aggregated read. This is most beneficial when evaluating for mosaicism. NGS testing becomes a quantitative approach of determining mosaicism by calculating the percentage of individual reads produced with the mutation versus those without the mutation (Fig. 5.8).

NGS can also expand beyond panel testing options, into whole exome or genome sequencing. Whole exome/genome sequencing can be useful when it is suspected that the genetic cause for an individual clinical presentation is not well established and likely unavailable in an individual gene or panel option. Whole exome/genome sequencing can also become helpful when someone has exhausted the panel-based options currently available for a condition and no mutation has been found. Lastly, when an individual presents with a phenotype that overlaps with multiple conditions, a whole exome/genome testing approach can be helpful in examining several genes simultaneously. The choice of whole genome versus exome sequencing depends in part on costs (WGS is significantly more expensive than WES currently) and degree of coverage required. WGS offers the possibility

TABLE 5.4
Sequencing: Platform, Target Amplification, Sequencing Platform, and Sequencing Chemistry

Platform	Target Amplification Technique	Sequencing Platform	Sequencing Chemistry
454/Roche	On-bead emulsion PCR	Beads are placed in individual wells for sequencing	Uses pyrosequencing by detecting pyrophosphate as it is released when each base pair is added
Illumina (MiSeq/ HiSeq/NextSeq)	Array-based "Bridge-PCR"	Each targeted DNA region is sequenced within a cluster randomly placed on the surface of a chip	The fluorescence of each base pair is captured as it is added to each amplicon
SOLiD	On-bead emulsion PCR	Beads are randomly located on the surface of a chip and each small oligonucleotide is labeled and sequenced in this location	The fluorescence of each base pair is captured as is it added to each amplicon
Ion Torrent	On-bead emulsion PCR	Beads are placed in individual wells for isolated sequencing	Detects proton release as each base pair is added

PCR, polymerase chain reaction.

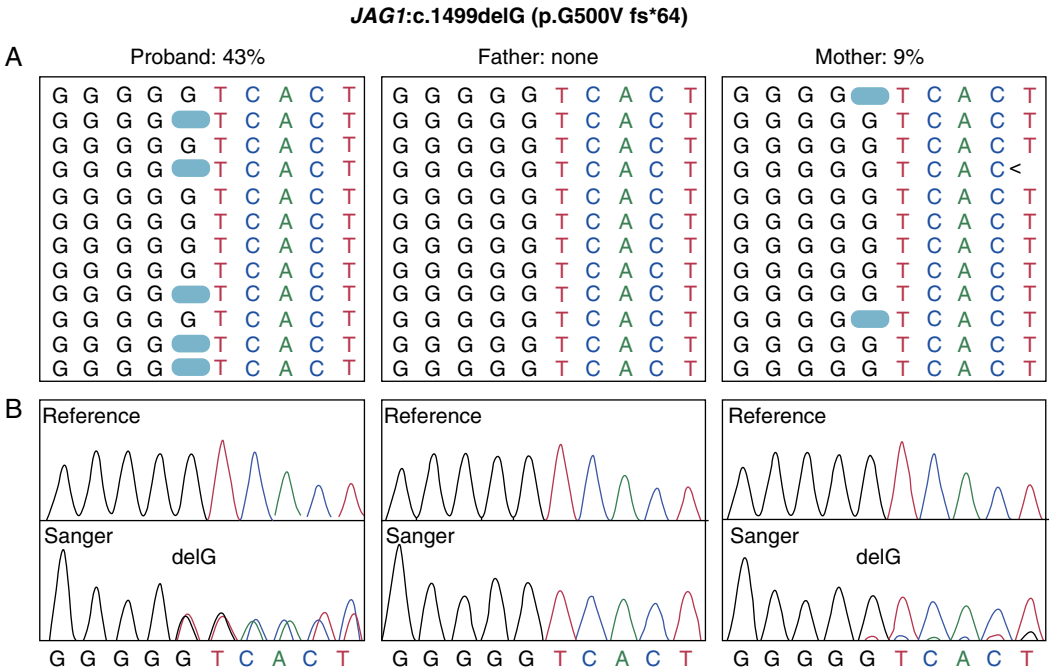


FIG. 5.8 Mosaic JAG1 mutation in maternal blood. **(A)** Next-generation sequencing detects a heterozygous JAG1:c.1499delG (p.G500Vfs*64) mutation in the proband; 9% of mutation load is detected in the mother, and the father is negative for this mutation. **(B)** Sanger sequencing results for this mutation. (Courtesy of Qin L, Wang J, Tian X, Yu H, et al. Detection and quantification of mosaic mutations in disease genes by next-generation sequencing. *J Mol Diagn*. May 2016;18(3):446–453. http://dx.doi.org/10.1016/j.jmoldx.2016.01.002. Epub 2016 Mar 2.)

TABLE 5.5
Guide to Testing Approaches

Single Gene (Sanger)	Panel Testing (NGS)	Whole Exome/Genome (NGS)
An individual has hallmark features that are indicative to one specific condition, which has a single gene associated with the phenotype; specific cancer-related variant is suspected	An individual has some features of a condition in which there are multiple genes that can be associated with the condition or conditions with overlapping presentations; multiple specific genes are tested that can be associated with a particular cancer	An individual presents with clinical features that can overlap several conditions and no specific gene or condition can be pinpointed as a starting point; the entire complement of genetic variants in a tumor needs to be identified

of detection of pathogenic variants outside the coding region of genes, although validating these as pathogenic can be challenging. Also, the bioinformatic analysis of the very large number of variants found by WGS can be more time-consuming and expensive.

Given the complexities and the unique approaches needed based on each individual's presenting circumstance, it can be difficult to determine what is the best place to start when deciding on a testing approach. Table 5.5 provides a guide that can be considered.

Limitations

NGS is limited by the analysis and bioinformatics pipeline used to analyze the data. Most NGS testing strategies are limited in their ability to detect copy number variations; however, the size and limits are specific to the bioinformatics software and the testing platform. In addition, using an NGS platform for panel testing versus whole exome/genome testing determines the amount of surface area on the platform that is available for each region of interest. As a result, in a panel setting, where there are smaller regions of interest, a targeted capture approach provides sufficient surface area to capture reads from all available regions of this specific target. In contrast, whole exome/genome sequences require the capture of more regions of interest, limiting the area available on the platform for reads of each target. This can mean that coverage for a particular region would be better represented in a targeted panel than a whole exome/genome test.

Another limitation to NGS testing is the potential for incidental findings. An incidental finding is a variant identified in gene that is unrelated to the original reason for testing, yet potentially medically significant. For example, whole exome sequencing might be done for an individual because of a personal history of heart defects, and a pathogenic mutation is identified in a gene associated with colon cancer. The likelihood of identifying an incidental finding increases with the number of target regions available on the testing platform; a whole exome/genome test will yield more potential incidental findings than a panel. To prepare the individuals tested, laboratories using this testing approach provide a statement on how they handle incidental findings and will provide a consent form for the individual to indicate whether he/she would/would not like to know if an incidental finding was identified. The American College of Medical Genetics and Genomics (ACMG) has issued a guideline regarding the disclosure of incidental findings. In summary, ACMG suggests disclosure of pathogenic mutations in genes with a direct association with a known clinical condition that also has a medical surveillance, management, treatment, or cure available. The guideline then provides a list of suggested genes that meet these criteria; however, it is left up to the individual testing laboratory whether they would like to adhere to this recommendation; some have chosen to expand the list to additional genes.

Lastly, as NGS testing has been an emerging clinical technology, many clinicians and certifying bodies require or suggest that any results identified via NGS be clinically confirmed using an orthogonal technology. Currently, most laboratories confirm variants identified via NGS by Sanger sequencing.

PRACTICAL EXAMPLES

Somatic: BCR-ABL gene fusion.

The Philadelphia chromosome is a reciprocal translocation involving chromosomes 9 and 22 that is commonly identified in chronic myelogenous leukemia (CML). The break points of the translocation create a fusion of two genes: ABL1 on chromosome 9 and BCR on chromosome 22. The fusion gene encodes a tyrosine kinase signaling protein that leads to uncontrollable cell division. Approximately 95% of individuals with a clinical diagnosis of CML have this chromosome anomaly, enabling this to be a sensitive

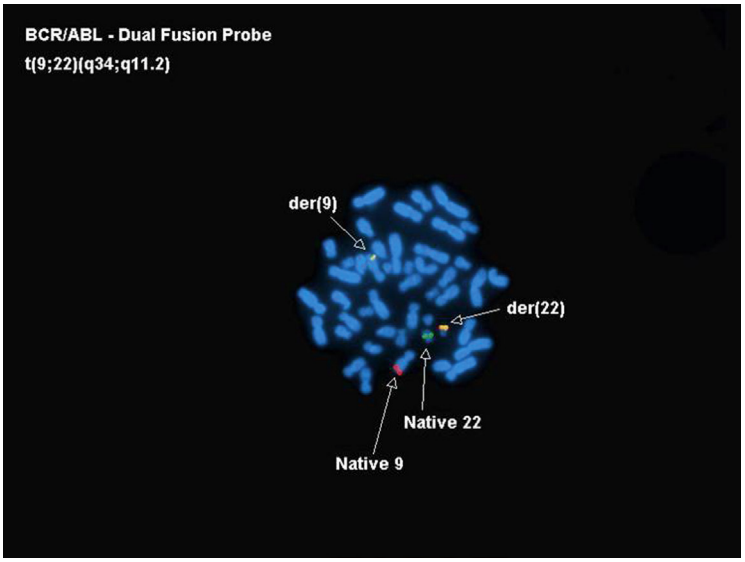


FIG. 5.9 Fluorescence in situ hybridization (FISH) probe analysis for BCR/ABL fusion probe. (Courtesy of Dr. Fady Mikhail, MD, PhD, FACMG. University of Alabama at Birmingham Cytogenetics Laboratory.)

test for clinical confirmation.[9] Identification of this translocation, however, is not specific, as it can also be found at lower detection rates in individuals with a diagnosis of acute lymphoblastic leukemia as well as acute myelogenous leukemia.

There are benefits in therapy when an individual is identified to have the BCR-ABL translocation. In the 1990s, the drug imatinib mesylate was developed as a tyrosine kinase inhibitor able to significantly limit the growth of tumor cells with the translocation. Since this time additional therapeutic agents have been created, such as dasatinib and nilotinib, with a similar mechanism of action.

Possible testing strategy

To test for this gene fusion, live cells can be submitted for culture and karyotyping. This will detect the Philadelphia chromosome and potentially other chromosomal abnormalities (Fig. 5.10). Alternatively, FISH analysis can be performed if a faster turnaround is needed, with the fusion gene identified by labeling the ABL and BCR probes with different colors (Fig. 5.9).

Germline: Klinefelter syndrome

Klinefelter syndrome is a sex chromosome aneuploidy that affects 1 in every 1000 newborn males. Klinefelter syndrome presents with a spectrum of phenotypes; in fact, it is suspected that this syndrome is underdiagnosed given how mild the phenotype may be.

The most common findings in this condition are caused by low testosterone production:

- small testes
- delayed or absent puberty
- cryptorchidism
- hypospadias
- micropenis
- infertility
- gynecomastia
- reduced or absent facial and body hair

Men with Klinefelter syndrome have at least one extra X chromosome. Instead of a typical 46,XY karyotype, a male with Klinefelter may present as 47,XXY, 48,XXXY, or even more copies of the X chromosome. A karyotype analysis is the simplest way to make this diagnosis.

Men with Klinefelter syndrome are also at an increased risk for some cancers and are at a decreased risk for prostate cancer. The underlying cause for these increased risks has not been established. The incidence of each cancer risk using a standardized mortality ratio compared with that of general population[3] is listed in Table 5.6.

A karyotype or FISH for the sex chromosomes would likely be the first testing approach used to confirm this diagnosis. FISH analysis would also be most helpful to determine if there is also mosaicism. A microarray is another testing approach that could establish this

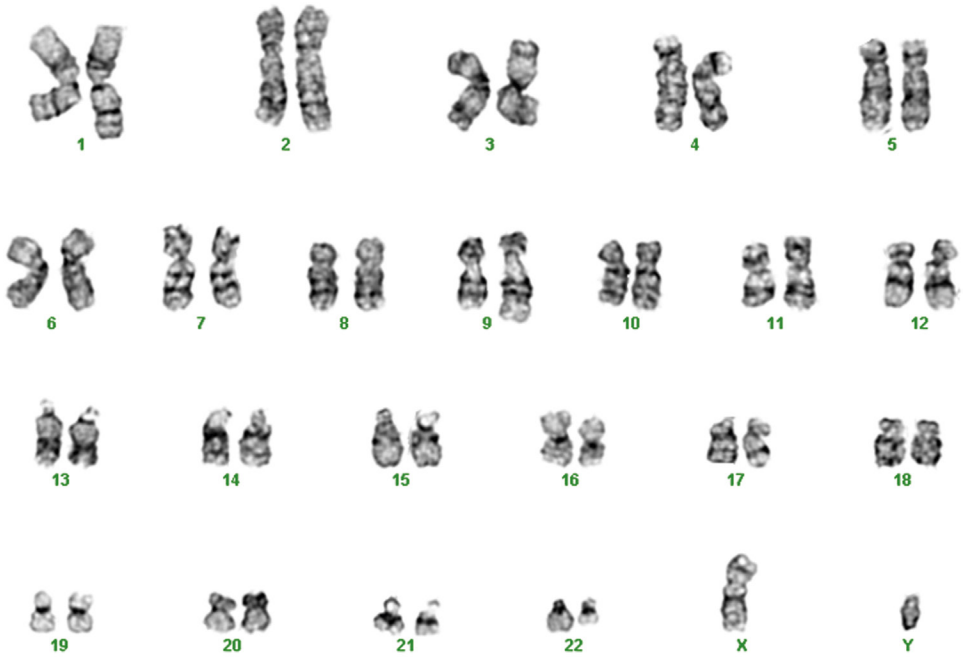


FIG. 5.10 Karyotype analysis for chromosome 9 and 22 translocation associated with BCR/ABL fusion. (Courtesy of Dr. Fady Mikhail, MD, PhD, FACMG. University of Alabama at Birmingham Cytogenetics Laboratory.)

diagnosis; however, a microarray analysis would not be able to confidently confirm if the patient is mosaic (Fig. 5.11).

GERMLINE TESTING: LI-FRAUMENI SYNDROME

Li-Fraumeni syndrome is a rare, autosomal dominant condition caused by mutation in the *TP53* gene. The *TP53* gene encodes a p53 protein that is unable to regulate cell growth or division. Mutations in *TP53* can be found in 50% of all tumors as somatic mutations. Individuals with Li-Fraumeni syndrome (LFS) have a greatly increased risk of developing cancer. The hallmark cancers associated with LFS are sarcomas, breast cancer, brain tumors (astrocytomas, glioblastomas, medulloblastomas, choroid plexus carcinomas), and adrenocortical carcinomas. Those with a germline mutation in the *TP53* gene are estimated to have a 50% chance of developing an LFS-related tumor by age 30 and >90% lifetime risk[10] (Table 5.7).

Although the Chompret criteria for clinical diagnosis of LFS provides a guideline for making a clinical diagnosis, identifying a causative gene mutation can be helpful in recognizing asymptomatic family members who require surveillance. In addition, a few clinical predictions can be made for those with certain mutations identified (genotype-phenotype correlation). Table 5.8 lists genotype-phenotype correlations that have been identified.

TABLE 5.6
Incidence of Cancer Rates

Cancer Type	Standardized Mortality Ratio
All cancers	1.2
Lung cancer	1.5
Breast cancer	57.8
Non-Hodgkin lymphoma	3.5
Prostate cancer	0

TESTING STRATEGY

To plan an appropriate testing strategy, it is important to identify the mutational spectrum seen in the condition in

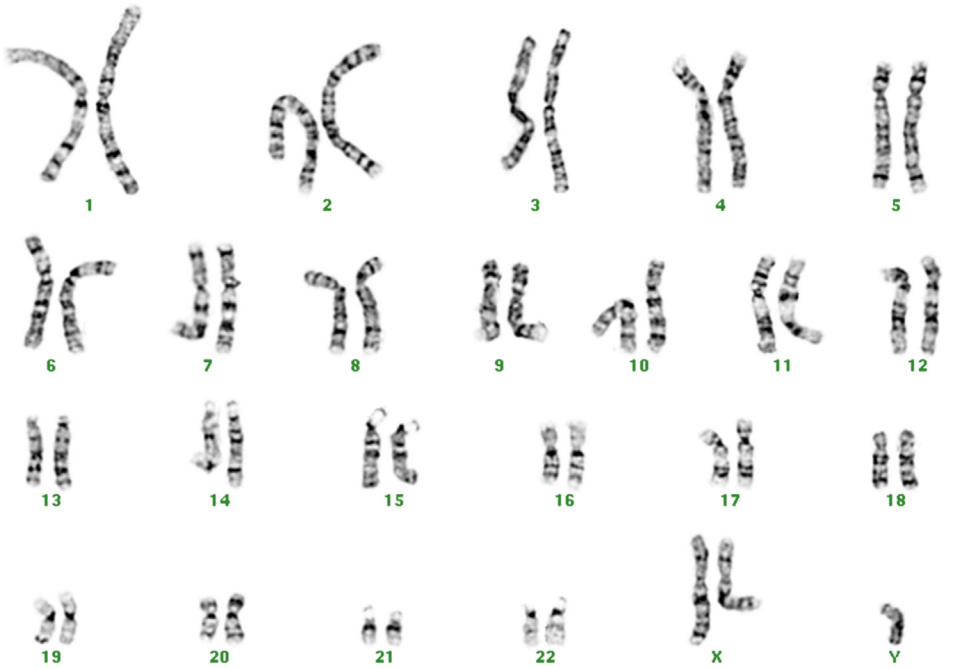


FIG. 5.11 Abnormal karyotype featuring Klinefelter syndrome. (Courtesy of Dr. Fady Mikhail, MD, PhD, FACMG. University of Alabama at Birmingham Cytogenetics Laboratory.)

TABLE 5.7
List of Common Tumors or Cancers Seen in Li-Fraumeni Syndrome as well as Their Lifetime Risk of Development[10]

Location	Relative Risk (Confidence Interval 95%)
Bone	107 (49–203)
Connective tissue	61 (33–102)
Brain	35 (19–60)
Pancreas	7.3 (2–19)
Breast	6.4 (4.3–9.3)
Colon	2.8 (1–6)
Liver	1.8 (2.1–64)

TABLE 5.8
Genotype and Phenotype Correlations

Genotype	Phenotype
Pathogenic missense mutations	Earlier onset of cancer
Total or partial gene deletions	Classic Li-Fraumeni syndrome phenotypes
Microdeletions involving tp53	Lower risk for tumor formation
Mutations in the DNA-binding loop that contact the minor groove of DNA	Increased risk for brain tumors
Pathogenic mutations within the loops opposing the protein-DNA contact surface	Increased risk for adenoid cystic carcinoma

question. 80% of families with features of LFS are found to have mutations in the TP53 gene. In addition, 95% of those with TP53 mutations are found to have variants identified via sequencing analysis and only 1% are found to have deletions/duplications in this gene. As a result, single gene Sanger sequencing would be an acceptable approach depending on a patient's clinical presentation. An adenoid cystic carcinoma tumor is more characteristic for LFS than a woman presenting with breast cancer. As a result, for one patient, a Sanger sequencing test for the TP53 gene may be the best starting point, whereas

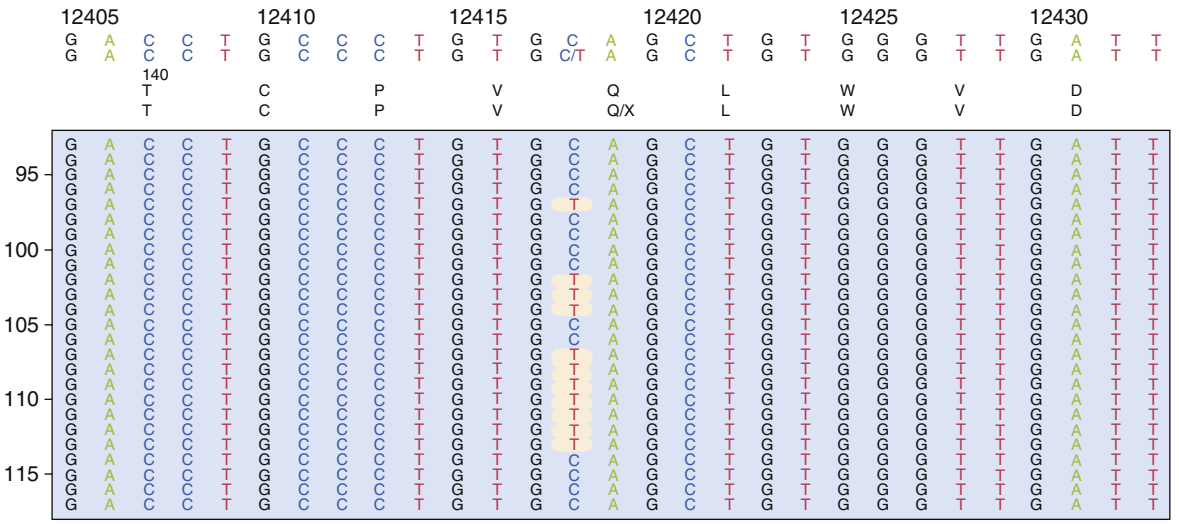


FIG. 5.12 Detection of a TP53 mutation in a patient with Li-Fraumeni syndrome. The reference sequence is shown at the top and the patient sequence below. A heterozygous C>T mutation (c.430C>T; p.Gln144X) is visible in 11 of the 25 reads shown. (Courtesy Jo Morgan and Graham Taylor, Leeds Institute of Molecular Medicine, St. James's Hospital, Leeds, UK. From Turnpenny P, Ellard S. *Emery's Elements of Medical Genetics*. Philadelphia, PA: Elsevier; 2017.)

another patient may require a multigene NGS breast cancer panel that includes the *TP53* gene. If the sequencing analysis is negative, one may consider deletion/duplication analysis as well. If the *TP53* gene is the only gene considered, MLPA analysis may be helpful for deletion/duplication analysis. If, however, several genes are under consideration for diagnosis, a microarray may be useful to look for deletions/duplications within all genes under consideration (Fig. 5.12).

TESTING ON THE HORIZON

As technology advances and our knowledge of cancer syndromes improve, genetic testing is constantly improving. In addition, new genes and gene associations with particular cancers are also being discovered. Therefore, individuals who receive genetic testing with negative results may benefit in follow-up testing every few years as testing strategies improve.

Cancer Screening Tests

Direct to consumer

Direct to consumer (DTC) tests can be ordered directly by patients, without the involvement of a health professional. While routine clinical testing options provide diagnostic results for specific genes and/or mutations, a DTC test analyzes SNPs and their statistical association with certain conditions, including cancer. Food and Drug Administration rules prevent return of clinical results based on DTC tests, unless a specific test has received approval. Nevertheless, a consumer can obtain his/her raw genotypic data, and there are services that can annotate these with potential medical conditions or risks.

ctDNA screening tests

Another testing approach on the horizon is the use of noninvasive cancer tests, the so-called liquid biopsy. Cell-free DNA-based testing involves NGS of circulating DNA derived from degenerating tumor cells in an individual's blood. One emerging test detects the presence of nine driver cancer genes and can provide results in 2–3 weeks.

REFERENCES

1. Bickmore W. Karyotype analysis and chromosome banding. *eLS*. 2001:1–5.
2. Hasle H, Clemmensen IH, Mikkelsen M. Risks of leukaemia and solid tumors in individuals with Down's syndrome. *Lancet*. 2000;355(9199):165–169.
3. Swerdlow A, Shoemaker M, Higgins C, et al. Cancer incidence and mortality in men with Klinefelter syndrome: a cohort study. *J Natl Cancer Inst*. 2005;97(16):1204–1210.
4. Schneider K. *Counseling about Cancer: Strategies for Genetic Counselors*. 3rd ed. Hoboken, New Jersey: John Wiley & Sons Inc; 2012.

5. Bumgarner R. DNA microarrays: types, applications, and their future use. *Curr Protoc Mol Biol*. 2013;0 22(22.1). http://dx.doi.org/10.1002/0471142727.mb2201s101.
6. Miller D, Adam M, Aradhya S, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet*. 2010;86:749–764.
7. Van Veghel-Plandsoen M, Wouters C, Kromosoteo J, et al. Multiplex ligation-depending probe amplification is not suitable for detection of low-grade mosaicism. *Eur J Hum Genet*. 2011;19(9):1009–1012.
8. Hui P. Next generation sequencing: chemistry, technology and applications. *Top Curr Chem*. 2012. http://dx.doi.org/10.1007/128_2012_329.
9. Melo JV. BCR-ABL gene variants. *Baillieres Clin Haematol*. 1997;10(2):203–222.
10. Schneider K, Zelley K, Nichols KE, et al. Li-Fraumeni Syndrome. www.ncbi.nlm.nih.gov/books/NBK1311/.