Title: Molecular genetics of 22q11.2 deletion syndrome

Bernice E. Morrow\textsuperscript{1*}, Donna M. McDonald-McGinn\textsuperscript{2}, Beverly S. Emanuel\textsuperscript{2}, Joris R. Vermeeesch\textsuperscript{3} and Peter J. Scambler\textsuperscript{4}

\textsuperscript{1}Department of Genetics, Albert Einstein College of Medicine, Bronx, NY, USA

\textsuperscript{2}Division of Human Genetics, Children’s Hospital of Philadelphia and Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, USA

\textsuperscript{3}Center for Human Genetics, Katholieke Universiteit Leuven (KU Leuven), Leuven, Belgium

\textsuperscript{4}Institute of Child Health, University College London, 30 Guilford St, London WC1N 1EH, United Kingdom

Running title: Molecular genetics of 22q11.2 deletion syndrome

*Corresponding Author: Bernice E. Morrow, Department of Genetics, Albert Einstein College of Medicine, 1301 Morris Park Ave, Bronx NY 10461, Email: Bernice.morrow@einstein.yu.edu

Phone: 718-678-1121, Fax: 718-678-1016
Key words: 22q11.2 deletion syndrome, chromosome rearrangements, congenital malformation, birth defect syndrome, pharyngeal apparatus, DiGeorge syndrome, velo-cardio-facial syndrome
ABSTRACT

The 22q11.2 deletion syndrome (22q11.2DS) is a congenital malformation and neuropsychiatric disorder caused by meiotic chromosome rearrangements. One of the goals of this review is to summarize the current state of basic research studies of 22q11.2DS. It highlights efforts to understand the mechanisms responsible for the 22q11.2 deletion that occurs in meiosis. This mechanism involves the four sets of low copy repeats (LCR22) that are dispersed in the 22q11.2 region and the deletion is mediated by non-allelic homologous recombination events. This review also highlights selected genes mapping to the 22q11.2 region that may contribute to the typical clinical findings associated with the disorder and explain that mutations in genes on the remaining allele can uncover rare recessive conditions. Another important aspect of 22q11.2DS is the existence of phenotypic heterogeneity. While some patients are mildly affected, others have severe medical, cognitive and/or psychiatric challenges. Variability may be due in part to the presence of genetic modifiers. This review discusses current genome-wide efforts to identify such modifiers that could shed light on molecular pathways required for normal human development, cognition or behavior.
Clinical aspects of 22q11.2 deletion syndrome

The chromosome 22q11.2 region is susceptible to meiotic chromosome rearrangements leading to congenital malformation syndromes. The best characterized among these is the 22q11.2 deletion syndrome (22q11.2DS; velo-cardio-facial syndrome (MIM#192430) or DiGeorge syndrome (MIM#188400). The disorder is the most common among microdeletion syndromes, occurring in ~1/4000 live births (Botto et al., 2003; Devriendt, Fryns, Mortier, van Thienen, & Keymolen, 1998; Goodship, Cross, LiLing, & Wren, 1998) and 1/1000 fetuses (Grati et al., 2015; Wapner et al., 2012). It occurs as a de novo 1.5-3 million base pair (Mb) deletion in most individuals (Lindsay et al., 1995; Morrow et al., 1995), although approximately 5% are inherited (McDonald-McGinn et al., 2001). When inherited, there is a 50% recurrence risk with 100% penetrance and wide expressivity. Affected individuals have mild to serious medical features including most often congenital heart disease, immunodeficiency, autoimmune disease, palatal abnormalities, hypocalcemia (often associated with hypoparathyroidism), thyroid disease, gastrointestinal differences, renal anomalies, skeletal anomalies, thrombocytopenia and characteristic facial features (Bassett et al., 2011; McDonald-McGinn et al., 2015). The majority have deficits in cognitive abilities and 25% develop schizophrenia (Bassett & Chow, 1999; Murphy, Jones, & Owen, 1999). Other significant behavioral anomalies occur including attention deficit disorder and anxiety (Biswas & Furniss, 2016; Campbell, McCabe, Melville, Strutt, & Schall, 2015; Niarchou, Martin, Thapar, Owen, & van den Bree, 2015; Vangkilde et al., 2016; Vorstman et al., 2015). Further medical and psychiatric disorders occur less often in affected individuals but at higher frequencies than the general population (Bassett et al.,
Studies of the condition provide opportunities to dissect out genetic and environmental risk factors or modifiers within the 22q11.2 region and elsewhere in the genome. Since the condition results from the deletion, it is necessary to understand why the deletion occurs.

**Mechanism for meiotic 22q11.2 chromosome rearrangements**

The 22q11.2 deletion usually occurs by meiotic non-allelic homologous recombination (NAHR) events between low copy repeats on chromosome 22q11.2 termed LCR22 (Edelmann, Pandita, & Morrow, 1999; Edelmann, Pandita, Spiteri, et al., 1999; Shaikh et al., 2000). Low copy repeats are also referred to as segmental duplications (Bailey et al., 2002). There are eight LCR22s that span the 22q11.2 region termed LCR22A through -H. There are four LCR22s that map to the 3 Mb region associated with the disorder and they are termed LCR22A, -B, -C and -D. In over 90% of patients, the region between LCR22A-D is hemizygotously deleted. This is referred to as the LCR22A-D deletion type. The 1.5 Mb deletion between LCR22A-B is the second most common deletion type, while the 2.0 Mb LCR22A-C deletion is least frequent [(Carlson et al., 1997); Figure 1]. LCR22A and LCR22D are the largest in size and have the most homology to each other, making them good targets for NAHR events. The LCR22s are composed of modules harboring pseudogenes that formed during primate evolution by the processes of gene duplication and gene conversion events (Babcock et al., 2003; Babcock et al., 2007; Pavlicek, House, Gentles, Jurka, & Morrow, 2005). Each LCR22 is a complex mosaic of modules that has been challenging to precisely map in humans. This complexity has made it exceedingly difficult to identify the position...
Efforts to determine the 22q11.2 deletion endpoints

There has been much interest to clone and sequence the position of the chromosome 22q11.2 breakpoints leading to the typical LCR22A-D deletion. Due to the complexity of LCR22 structure, usage of traditional sequencing methods has only been partially successful, as eluded to above (X. Guo et al., 2016). This technical challenge can potentially be alleviated by the application of optical genome mapping techniques. One approach utilizes long single molecules imaged in nanochannel arrays to generate molecular maps (Lam et al., 2012; Mak et al., 2016). Another approach is molecular combing with targeted probes flanking and targeting the LCR22 subunits, which will visualize the probe patterns along DNA fibers. The stretching of chromatin is a particular process needed to form the fibers (Michalet et al., 1997). By leveraging the increased sensitivity afforded by long single molecule optical maps, coupled with low coverage whole-genome sequence, the previously unmappable architecture and variation of the LCR22s and surrounding regions might be elucidated. This will provide enhanced insight into the mechanisms responsible for the deletion and its associated phenotypes. Further, it will likely provide an approach to deletion breakpoint detection in a clinical setting.

Clinical testing for the 22q11.2 deletion

There are several different molecular tools that have been developed for clinical
purposes to identify the 22q11.2 deletion. Classic testing by fluorescence in situ hybridization (FISH) utilizes commercial probes such has N25 or TUPLE (Figure 1). The two FISH probes map to the LCR22A-B region, thus the probes will identify typical deletions, but would miss atypical nested deletions that occur outside the A-B region. Multiplex Ligation-dependent Probe Amplification (MLPA) has probes that are spaced throughout the 22q11.2 region and can be used to identify typical and atypical deletions (Fernandez et al., 2005; Vorstman et al., 2006). MLPA is most useful to confirm a suspected diagnosis because it only probes the 22q11.2 region. Clinical genome-wide microarray testing (single nucleotide polymorphism/SNP genotyping microarray or microarray comparative genome hybridization-CGH) is useful when it is not possible to make a definitive diagnosis based upon clinical presentation or when MLPA is not clinically available. One of the unexpected discoveries of genome-wide testing is the expansion of the known phenotypic spectrum in patients with 22q11.2DS. Further, it has identified atypical nested 22q11.2 deletions such as the LCR22B-D or C-D deletions. These patients have some related features to those with the typical deletions, but they occur with reduced penetrance (Racedo et al., 2015; Rump et al., 2014). What is truly fascinating is that the LCR22A-B deleted region does not overlap with the LCR22B-D or C-D region, ascribing haploinsufficiency of different genes to similar phenotypes. The existence of atypical deletions underscores the importance of evaluating deletion size as well as of understanding the function of genes throughout the 22q11.2 region to delineate their role in individual abnormalities and the overall phenotype of the disorder.
Genes on 22q11.2 that cause congenital malformations

There are 46 known protein coding genes, seven miRNA and ten non-coding genes that map to the 3 Mb, 22q11.2 region, as well as additional predicted coding and non-coding genes as shown in Figure 1. Among the genes mapping to the LCR22A-B region is \textit{TBX1}, which encodes a T-box type of transcription factor. This gene has received much attention for a major role in medical aspects of the disorder. Heterozygous mutations within \textit{TBX1} have been identified in patients with similar defects as occurs in those with 22q11.2DS (Castellanos, Xie, Zheng, Cvekl, & Morrow, 2014; Gong et al., 2001; Griffin et al., 2010; Ogata et al., 2014; Pan et al., 2015; Stoller & Epstein, 2005; Torres-Juan et al., 2007; Xu et al., 2014; Yagi et al., 2003; Zweier, Sticht, Aydin-Yaylagul, Campbell, & Rauch, 2007). Supporting this, and occurring chronologically before mutations in patients were discovered, mouse models in which \textit{Tbx1} was inactivated were generated (Jerome & Papaioannou, 2001; Lindsay et al., 2001; Merscher et al., 2001). Inactivation of one allele resulted in mild cardiovascular defects, but inactivation of both alleles resulted in cleft palate, thymic and parathyroid gland aplasia as well as cardiovascular defects (Jerome & Papaioannou, 2001; Lindsay et al., 2001; Merscher et al., 2001). A series of alleles was generated with varying expression levels of \textit{Tbx1} and it was found that the gene is very sensitive to changes in copy number, such as what occurs when one allele is deleted in humans (Baldini, 2006; Z. Zhang & Baldini, 2008). Of interest, it was discovered that different tissues and organs have different degrees of sensitivity to \textit{Tbx1} dosage and suggests that altered regulation of its expression could explain phenotypic variability in patients (Z. Zhang & Baldini, 2008). Further, overexpression of \textit{TBX1} is able to genetically rescue
malformations that occur in mouse deletion models equivalent to the LCR22A-B deletion (Lindsay et al., 2001; Merscher et al., 2001). The discovery of TBX1 as a major player in medical aspects of 22q11.2DS has allowed the exploration of genetic pathways upstream and downstream of TBX1 needed for normal embryonic development (Ivins et al., 2005; Liao et al., 2008).

In mammals, Tbx1 is expressed in the embryonic cells that will form the craniofacial region, thymus and parathyroid glands, aortic arch as well as cardiac outflow tract. These are the structures affected in 22q11.2DS and in Tbx1 loss of function mouse embryos. The Tbx1 expressing cells lie within the embryonic pharyngeal apparatus (Figure 2), which is a structure that will form gills in fish but becomes dramatically remodelled in mammals. Tbx1 is expressed and functions in all three germ layers of the pharyngeal apparatus, namely the endoderm (Arnold et al., 2006; Jackson, Kasah, Mansour, Morrow, & Basson, 2014; Z. Zhang et al., 2005), mesoderm (Z. Zhang, Huynh, & Baldini, 2006) and ectoderm (Z. Zhang et al., 2005). These cells communicate signals to neural crest cells. Neural crest cells comprise a population of cells that delaminates and migrates from the neural tube. These cells make a major contribution to morphogenesis of the structures affected in 22q11.2DS (Calmont et al., 2009). This illustrates the complexity of the tissue interactions that are required for morphogenesis of the derivative structures from the pharyngeal apparatus (Papangeli & Scambler, 2013). Recently, direct downstream transcriptional target genes of TBX1 protein were discovered (Baldini, Fulcoli, & Illingworth, 2017; Fulcoli et al., 2016). These genes might shed light into the molecular pathogenesis of defects that occur in mice.
Besides *TBX1*, there is another gene that may contribute to the medical findings in 22q11.2DS and the gene is termed *CRKL* (CRK like proto-oncogene, adaptor protein). *CRKL* maps to the LCR22C-D region and it encodes a cytoplasmic adaptor protein involved in growth factor signalling. *Crkl* is required for development of the thymus and parathyroid glands, aortic arch and heart (Guris, Duester, Papaioannou, & Imamoto, 2006; Guris, Fantes, Tara, Druker, & Imamoto, 2001; Moon et al., 2006). Of interest, inactivation of both alleles of *Crkl* in the mouse results in similar defects as in *Tbx1* null mutant mice. Mice that have one allele of *Tbx1* and *Crkl* inactivated together have similar defects as in 22q11.2DS patients (Guris et al., 2006; Guris et al., 2001). This supports the idea that the two genes act in the same genetic pathway. Unlike *Tbx1*, whose expression is restricted in specific cells and tissues, *Crkl* is ubiquitously expressed in all cells (Guris et al., 2001). It is hypothesized that *Tbx1* acts upstream of the *Fgf8* (Fibroblast growth factor 8) growth factor gene, and this activates *CRKL* in neural crest cells and leads to activation of downstream signalling (Moon et al., 2006). Recently, it was found that hemizygosity of *CRKL* contributes to genitourinary tract development in humans and animal models, including the kidney (Haller, Mo, Imamoto, & Lamb, 2017; Lopez-Rivera et al., 2017). Although these are not considered the most common of defects in 22q11.2DS, they occur more commonly than the general population. Besides coding genes, non-coding genes, such as miRNAs might also contribute to the etiology of 22q11.2DS.

The miRNAs are small non-coding RNAs that regulate expression of target genes by binding to specific sites in messenger RNAs causing repression of translation or degradation. Another gene of note required for both cardiovascular and
brain development or function is DGCR8 [DiGeorge Critical Region 8; (Chapnik, Sasson, Blelloch, & Hornstein, 2012; Earls et al., 2012; Sellier et al., 2014)]. This gene encodes a subunit of a microprocessor complex that mediates the biogenesis of miRNAs (Gregory et al., 2004). Further, there are several miRNA genes that map to the 22q11.2 region (Figure 1) and they have not been studied in detail, as of yet. Another potentially relevant gene is HIC2 (Hypermethylated In Cancer 2). HIC2 maps to a small unique genomic region near the distal end of LCR22D and it is not usually deleted in patients. However, some individuals have atypical longer deletions in LCR22D that extend beyond HIC2 (Dykes et al., 2014). HIRA is also of interest, because it encodes a component of a protein complex that deposits the variant histone H3.3 at regulatory regions of genes, thereby modulating gene expression (Dilg et al., 2016; Farrell et al., 1999; Majumder, Syed, Joseph, Scambler, & Dutta, 2015; H. Zhang et al., 2017). Supporting a possible relevant function, conditional inactivation of Hira in mouse mesoderm cells causes cardiac defects (Manchineella, Thrivikraman, Basu, & Govindaraju, 2016).

**Genes on 22q11.2 that affect the brain or behavior.**

The relative importance of individual genes on chromosome 22q11.2 that contribute to cognitive and behavioral problems in patients is still being elucidated. Genes in the nested, LCR22A-B region have been more extensively evaluated than those in the distal, LCR22B-D region, because patients have been identified with the LCR22A-B deletion that are affected with schizophrenia (Karayiorgou et al., 1995). Identifying which gene(s) contribute to these anomalies has been challenging because
the great majority of genes in the LCR22A-B region are expressed in the brain (Maynard et al., 2003). In humans, the gene most well-studied for such deficits is \textit{COMT} (Catechol-O-methyltransferase). The COMT protein has a critical role in dopamine neurotransmitter metabolism. Further, there is a commonly occurring coding variant, referred to as Val158Met that affects enzyme activity (Li et al., 1996). While some studies support a role of reduced enzyme activity of \textit{COMT} to cause cognitive and/or behavioural deficits, others do not, and the importance of this functional variation in \textit{COMT} remains controversial (Armando, Papaleo, & Vicari, 2012; Franconi et al., 2016). Another gene that has been extensively studied in humans is \textit{PRODH}, encoding proline dehydrogenase that converts proline to glutamate, involved in neurotransmission. As is true for \textit{COMT}, the role of \textit{PRODH} is not yet definitive (Carmel et al., 2014; Radoeva et al., 2014; Zarchi et al., 2013). Commonly occurring single nucleotide polymorphisms in additional genes have been identified in patients with schizophrenia, but most of the studies were small in size and some of the findings failed replication. A gene of interest outside the LCR22A-B region, mapping to the LCR22C-D region is \textit{PI4KA} (Phosphatidylinositol 4-kinase alpha; (Ikeda et al., 2010; Vorstman et al., 2009), however, more studies need to be done to understand its connection to behavioral phenotypes.

Genes in the LCR22A-B region have been carefully evaluated in mouse models for a possible role in brain development or function (Karayiorgou & Gogos, 2004; Wang, Bryan, LaMantia, & Mendelowitz, 2017). Genes of interest required for brain or behavioral function in mouse models are \textit{Zdhhc8} (Zinc finger DHHC-type containing 8; (Mukai et al., 2004) ), \textit{Ranbp1} (Ran GTPase-binding protein; (Paronett, Meechan,
Genes important for mitochondrial function are newly appreciated in mouse models, to have a role in the etiology of schizophrenia (Devaraju & Zakharenko, 2017; Meechan, Maynard, Tucker, & LaMantia, 2011), and specifically, for example, Mrpl40 [Mitochondrial large ribosomal subunit protein 40; (Devaraju et al., 2017)]. Further work needs to be done to understand the relative importance of these genes in humans.

Uncovering of recessive mutations on 22q11.2

Over the years, some subjects have been identified with unusual clinical findings co-occurring with the main phenotypes of 22q11.2DS. One possible explanation is that there is a mutation in a gene on the remaining, haploid allele of 22q11.2 resulting in the co-occurrence of a recessive disorder. This explanation is exactly what has been found to occur. Several genes within the deleted region have been identified as harboring autosomal recessive mutations and they are listed in Table 1. Five genes, PRODH (Afenjar et al., 2007; Bender et al., 2005; Efron, 1965; Goodman, Rutberg, Lin, Pulver, & Thomas, 2000; Jacquet et al., 2003; Jacquet et al., 2002), SLC25A1 [Solute family carrier 25 member 1 (Edvardson et al., 2013; Majd, King, Smith, & Kunji, 2018; Nota et al., 2013)], CDC45 [Cell division cycle 45; (Fenwick et al., 2016)], GP1BB [Glycoprotein 1b platelet subunit beta; (Budarf et al., 1995; Hayashi & Suzuki, 2000; Kato et al., 2003; Kunishima et al., 2013; Lawrence, McDonald-McGinn, Zackai, & Sullivan, 2003; Ludlow et al., 1996; Roth, 1996; Tang et al., 2004)] and TANGO2 [Transport and golgi
organization 2 homolog; (Kremer et al., 2016; Lalani et al., 2016)], map to the LCR22A-B region. Each recessive disorder has their own characteristic clinical findings as listed in Table 1. SCARF2 [Scavenger receptor class F member 2; (Bedeschi et al., 2010)] maps to the LCR22B-C region, while SNAP29 [Synaptosome associated protein 29; (T. Hsu et al., 2017; McDonald-McGinn et al., 2013; Sprecher et al., 2005)] and LZTR1 [Leucine zipper like transcription regulator 1; (Johnston et al., 2018)] map to the LCR22C-D region. Somatic loss of heterozygosity of one allele and a loss of function mutation in LZTR1 on the other allele predispose individuals to an inherited disorder of multiple schwannomas (Piotrowski et al., 2014). These findings of recessive disorders and a possible somatic dominant condition, confirm that extreme phenotypic abnormalities observed in a subset of patients with 22q11.2DS can be due to mutations on the non-deleted allele, leading to unmasking of additional conditions. It is also possible that less severe mutations or variations on the remaining allele could contribute to phenotypic variability in 22q11.2DS.

**Phenotypic variability and 22q11.2DS**

One of the most challenging features of 22q11.2DS is its phenotypic variability, despite the fact that most patients have the same size, 3 Mb deletion. For example, 60-70% have congenital heart disease, which includes many different aortic arch and/or cardiac outflow tract anomalies (Goldmuntz et al., 1993; McDonald-McGinn et al., 2015; Swillen & McDonald-McGinn, 2015). Other common features include cleft of the muscular palate/velopharyngeal dysfunction, endocrine abnormalities including hypocalcemia, thyroid disease and growth hormone deficiency,
immunodeficiency/autoimmune disease, gastroenterological issues, skeletal abnormalities such as scoliosis, as well as, neuropsychiatric and cognitive deficits. As mentioned above, schizophrenia occurs in 25% of individuals with the disorder (Bassett & Chow, 1999; Murphy et al., 1999). This is the most common pathogenic copy number variation associated with schizophrenia. One possible explanation is the occurrence of other genetic variations in the genome, in addition to the deletion that modify the overall phenotype. Another possible explanation is the occurrence of stochastic events during embryogenesis and a third is the existence of as of yet, unknown environmental exposures in pregnancy. Of interest, monozygotic twins with 22q11.2DS typically have discordant features (Fryer, 1996; Goodship, Cross, Scambler, & Burn, 1995; Halder, Jain, Chaudhary, & Varma, 2012; Hatchwell, 1996; Lu, Chung, Hwang, & Chien, 2001; McDonald-McGinn et al., 2001; Rauch et al., 1998; Singh, Murphy, & O'Reilly, 2002; Vincent et al., 1999; Yamagishi et al., 1998). This implicates stochastic events, epigenetic variation in pregnancy or somatic mutations as also contributing to variation.

Among the possibilities, identifying genetic modifiers on the remaining 22q11.2 allele or elsewhere in the genome are feasible to test in sufficient sized cohorts of affected individuals. This may be useful for identifying modifiers of congenital heart disease because this phenotype is usually available for patients and it is identified soon after birth. For this, firstly, mutations in the coding region of TBX1 were ruled out influence cardiovascular phenotypes in a large cohort of 22q11.2DS patients (T. Guo et al., 2011).

In terms of genetic factors, one obvious category to investigate is the presence of second-hit large-sized copy number variations (CNVs). The CNVs might contain genes
whose function is sensitive to copy number for various phenotypes. Some may act in the same genetic pathway as 22q11.2 genes, and either exacerbate or suppress individual phenotypes. Affymetrix 6.0 arrays on almost one-thousand subjects with 22q11.2DS revealed a commonly occurring duplication CNV containing the SLC2A3 (Solute carrier family 3 member 2) gene in individuals with 22q11.2DS with congenital heart disease compared to those with the condition but with normal cardiac and/or aortic arch structures (Mlynarski et al., 2015). Analysis of rare CNVs that alter risk to cardiovascular anomalies has been more challenging, with no specific increased burden of CNVs, but it is possible, when taken together that CNVs harboring genes important for cardiovascular development play a role in this sensitized population (Mlynarski et al., 2016). Recently, rare CNVs were shown to contribute to schizophrenia and 22q11.2DS (Bassett et al., 2017).

Whole exome sequencing on 184 individuals with 22q11DS, in which roughly half had severe cardiac anomalies and half had no defect, identified genes important for chromatin modification as altering risk (T. Guo et al., 2015) as well as additional genes with rare possible damaging variants, for cardiac development (Lin, Zhang, Cai, Morrow, & Zhang, 2017). Examination of whole genome sequence (WGS) in >1,500 subjects with 22q11.2DS is currently underway to identify genetic factors for congenital heart disease as well as for schizophrenia (Gur et al., 2017). The availability of WGS will enable scientists to identify common and rare single nucleotide variations in coding and non-coding regions, as well as large or smaller second-hit CNVs that serve as modifiers. It can be hypothesized that some of the genes or regulatory regions that will be uncovered likely interact with genes that are hemizygously deleted on the 22q11.2
CONCLUSIONS

Following the discovery of the 22q11.2 deletion in patients, almost 40 years ago, there is still much to learn about molecular aspects of the condition. This includes efforts to determine the mechanism of chromosome 22q11.2 deletions that could lead to development of improved clinical screening methods and understanding the genes responsible for typical as well as atypical phenotypes. Further, 22q11.2DS serves as a model to understand how genetics and the environment can modify the phenotype.
## TABLE

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Map position (hg19 assembly)</th>
<th>Disease name</th>
<th>MIM #</th>
<th>Phenotype</th>
<th>PMID</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRODH</strong></td>
<td>A-B; (chr22:18,900,287-18,924,066)</td>
<td>Hyperprolinemia Type 1</td>
<td>239500</td>
<td>Neurological deficits, psychomotor delay, hypotonia, seizures</td>
<td>11196113; 12217952; 12525555; 14290545; 15662599; 17412540</td>
</tr>
<tr>
<td><strong>SLC25A1</strong></td>
<td>A-B (chr22:19,163,088-19,166,338)</td>
<td>D2L2AD syndrome</td>
<td>615182</td>
<td>Encephalopathy, severe muscular weakness, seizures, respiratory distress, failed psychomotor development, early death</td>
<td>23393310, 23561848, 29031613</td>
</tr>
<tr>
<td><strong>CDC45</strong></td>
<td>A-B (chr22:19,467,414-19,508,135)</td>
<td>Meier-Gorlin syndrome 7</td>
<td>617063</td>
<td>Craniosynostosis, imperforate anus, limb abnormalities, short stature, absent patellae and microtia</td>
<td>27374770</td>
</tr>
<tr>
<td><strong>GP1BB</strong></td>
<td>A-B; (chr22:19,705,992-19,712,297)</td>
<td>Bernard-Soulier syndrome</td>
<td>231200</td>
<td>Hematologic disease; thrombocytopenia; increased megakaryocytes</td>
<td>8952885; 10805283; 23566026</td>
</tr>
<tr>
<td><strong>TANGO2</strong></td>
<td>A-B (chr22:20,008,631-20,053,447)</td>
<td>MECRCN syndrome</td>
<td>616878</td>
<td>Recurrent metabolic crises with encephalocardiomyopathy that includes rhabdomyolysis, neurodegeneration, hearing loss, thyroid disease, hypoglycemia, seizures, dystonia and sudden death</td>
<td>26805781; 26805782</td>
</tr>
<tr>
<td><strong>SCARF2</strong></td>
<td>B-C (chr22:20,778,874-20,792,146)</td>
<td>van den Ende-Gupta syndrome</td>
<td>600920</td>
<td>Contractural arachnodactyly, hooked clavicles, joint dislocations and blepharophimosis</td>
<td>22140376</td>
</tr>
<tr>
<td><strong>SNAP29</strong></td>
<td>C-D (chr22:21,213,292-21,245,501)</td>
<td>CEDNIK Syndrome</td>
<td>609528</td>
<td>Cerebral dysgenesis, neuropathy, ichthyosis and keratoderma</td>
<td>15968592; 23231787; 29051910</td>
</tr>
<tr>
<td><strong>LZTR1</strong></td>
<td>C-D (chr22:21,336,558-21,353,326)</td>
<td>Autosomal recessive Noonan syndrome 10; somatic: multiple schwannomas</td>
<td>616564</td>
<td>Congenital heart disease, ptosis, short stature and webbed neck</td>
<td>29469822; 24362817</td>
</tr>
</tbody>
</table>

## ACKNOWLEDGEMENTS

We would like to thank the families with 22q11.2DS who have participated in molecular genetic studies. This work was supported by grants NIH R01 HL084410 (BSE, BEM,
DMM), P01 HD070454 (BSE, BEM, DMM), MH087636 (BSE, DMM), U01 MH101720 (BSE, DMM, JV) and a Foundation Leducq grant (BEM, PJS). This work was also funded by the British Heart Foundation, PG/09/065/27893 (PJS). This work was made possible by grants from the KUL PFV/10/016 SymBioSys, GOA/12/015 G.0E1117N, the Jerome Lejeune Foundation (project #1665) to JRV.

REFERENCES


analysis using human and mouse genetic approaches. *PLoS One, 2*(11), e1234. doi:10.1371/journal.pone.0001234


### TABLE

Table 1. Recessive mutations in genes on the 22q11.2 region. Each gene is listed with respect to the map location and to the position of the LCR22s. The genomic coordinates in the hg19 genome assembly is listed in the second column. The name of the genetic syndrome caused by recessive mutations is listed as is the Online Mendelian Inheritance in Man (MIM) number. Characteristic clinical features of each disorder are shown. The references used to create the table are listed by PMID (PubMed ID) in the column on the right and are cited in the text.

### FIGURE LEGENDS

**Figure 1. The 22q11.2 region (McDonald-McGinn et al., 2015).** This image was previously published (McDonald-McGinn et al., 2015), with slight modifications. The 3
Mb 22q11.2 region (hg19 assembly, coordinates) is shown as a line that spans the 22q11.2 region, distal to the centromere (Cen) and harbors four sets of low copy repeats (LCR22), termed LCR22A, -B, -C and –D (green boxes). The position of clinical diagnostic probes, N25 and TUPLE, for clinical fluorescence *in situ* hybridization (FISH) testing are shown (aqua boxes). Most of the known coding and non-coding genes that span the interval are shown below the line representing the 22q11.2 region. Genes associated with recessive genetic conditions are indicated in gray boxes. Non-coding genes are indicated with a star. The size and position of 22q11.2 deletions are indicated (aqua boxes).

**Figure 2. Embryonic development of the pharyngeal apparatus into adult structures.** Cartoon of a lateral view of a mouse embryo at stage day (E) 10.5 in development. The head is shown on top and the tail can be visualized as it twists around the body of the embryo. The pharyngeal apparatus is highlighted in light blue with individual arches (PA) and their derivative structures indicated on the left side (arrows). The outflow tract (OFT), right ventricle (RV), forelimb (FL) and hindlimb (HL) are indicated. Cells migrate from the pharyngeal apparatus into the cardiac OFT to form the aorta and pulmonary trunk during later stages of embryogenesis.