

## Genetic susceptibility to heart disease in Canada: lessons from patients with familial hypercholesterolemia

Robert A. Hegele

**Abstract:** Much of the recent progress in treating patients with heart disease due to narrowed coronary arteries has resulted from studying disease evolution in patients with rare monogenic forms of disease. For instance, autosomal dominant familial hypercholesterolemia (FH, MIM (Mendelian Inheritance in Man) 143890) typically results from heterozygous mutations in *LDLR* encoding the low-density lipoprotein (LDL) receptor. Deficient *LDLR* activity results in elevated circulating LDL cholesterol, which accumulates within blood vessel walls, forming arterial plaques that can grow and eventually occlude the arterial lumen. Heterozygous *LDLR* mutations are usually detected using exon-by-exon sequence analysis (EBESA) of genomic DNA, a technology that has identified ~50 mutations in heterozygous FH (HeFH) subjects in Ontario. However, ~35% of Ontario HeFH patients had no EBESA-identified *LDLR* mutation. The diagnostic gap relates both to the genetic heterogeneity of FH and also to inadequate sensitivity of EBESA to detect certain mutation types, such as large deletions or insertions in *LDLR*. By means of a dedicated method to detect copy number variations (CNVs), additional heterozygous mutations in *LDLR* ranging from ~500 to >15 000 bases were uncovered, accounting for most of the remainder of Ontario HeFH patients. The appreciation of the key role of genomic CNVs in disease coincides with recent genome-wide mapping studies demonstrating that CNVs are common in apparently healthy people. CNVs thus represent a new level of genomic variation that is both an important mechanism of monogenic disease and a contributor to genomic variation in the general population; as well, it may have implications for evolution, biology, and possibly susceptibility to common complex diseases.

**Key words:** atherosclerosis, genetics, monogenic disorders, molecular diagnosis, DNA analysis.

**Résumé :** Une part importante des progrès récents dans le traitement des patients souffrant de maladies cardiaques découlant du rétrécissement des artères coronaires résulte de l'étude de l'évolution de la maladie chez des patients ayant des formes monogéniques rares de la maladie. Par exemple, l'hypercholestérolémie familiale (FH, MIM 143890), une condition autosomique dominante, découle habituellement de mutations hétérozygotes dans le gène *LDLR* qui code pour le récepteur des lipoprotéines de basse densité (LDL). Une activité *LDLR* déficiente entraîne une élévation du cholestérol LDL en circulation, lequel s'accumule dans les parois des vaisseaux sanguins et forme des plaques artérielles qui peuvent croître et éventuellement obstruer le lumen artériel. Les mutations *LDLR* hétérozygotes sont habituellement détectées par séquençage exon-par-exon (EBESA) de l'ADN génomique, une technologie qui a permis d'identifier ~50 mutations chez des sujets FH hétérozygotes (HeFH) en Ontario. Cependant, ~35 % des patients HeFH en Ontario ne présentent aucune mutation *LDLR* suite à une analyse EBESA. Ces insuccès diagnostiques découlent à la fois de l'hétérogénéité du FH et de la sensibilité inadéquate de l'EBESA pour la détection de certains types de mutations tels que les grandes délétions ou insertions dans le gène *LDLR*. À l'aide d'une méthode conçue pour détecter les variations dans le nombre de copies (CNV), des mutations hétérozygotes additionnelles correspondant à des changements de ~500 jusqu'à >15 000 bases ont été découvertes chez le restant des patients HeFH ontariens. Une appréciation du rôle important du CNV génomique coïncide avec les récents travaux en cartographie à l'échelle du génome qui ont montré que les CNV sont communs chez des sujets apparemment en bonne santé. Les CNV représentent ainsi un nouveau niveau de variation qui est à la fois un important mécanisme pour les maladies monogéniques et un contributeur à la variation génomique au sein de la population générale. Ceci pourrait avoir des implications pour l'évolution, la biologie et possiblement la sensibilité à des maladies complexes communes.

Received 26 September 2006. Accepted 27 November 2006. Published on the NRC Research Press Web site at <http://genome.nrc.ca> on 26 January 2007.

Corresponding Editor: R. Singh.

**R.A. Hegele.** Blackburn Cardiovascular Genetics Laboratory, Robarts Research Institute, Schulich School of Medicine and Dentistry, University of Western Ontario, 406-100 Perth Drive, London, ON N6A 5K8, Canada (e-mail: [hegele@robarts.ca](mailto:hegele@robarts.ca)).

*Mots clés* : athérosclérose, génétique, désordres monogéniques, diagnostique moléculaire, analyse de l'ADN.

[Traduit par la Rédaction]

## Introduction

Atherogenesis is a complex process that involves the contributions of several pathophysiological subsystems and is responsible for >50% of deaths in Westernized societies (Lusis 2000). Dissecting the genetic component of atherosclerosis has become possible with current molecular technologies. A key pathway that contributes to atherosclerosis is the lipid biosynthetic and transport pathway. Cholesterol is carried through the blood stream in spherical lipoprotein particles: the 2 main species of cholesterol-carrying particles are called low-density lipoprotein (LDL or “bad cholesterol”) and high-density lipoprotein (HDL or “good cholesterol”) (Lusis 2000). LDL cholesterol accumulates within blood vessel walls, forming arterial plaques that can grow and eventually occlude the arterial lumen, whereas HDL facilitates removal of cholesterol from blood vessel walls. Within certain families and communities, the effect of a single candidate gene on atherosclerosis susceptibility may be profound, as exemplified by mutations in *LDLR* encoding the low-density lipoprotein (LDL) receptor, which cause heterozygous familial hypercholesterolemia (HeFH) and premature atherosclerosis (Yuan et al. 2006). By defining important molecular pathways and targets, the study of monogenic forms of cholesterol metabolism has helped to develop effective new therapies, such as statin drugs that reduce plasma LDL cholesterol not only in HeFH patients but also in individuals in the general population at risk of heart attack and stroke. In addition to providing biochemical and biological insights, HeFH also demonstrates the importance of studying a range of types of genomic variation, especially large-scale copy number variations (CNVs). CNVs are increasingly recognized as a ubiquitous source of genomic variation that contributes substantially to interindividual variation in the general population (Feuk et al. 2006a, 2006b), to the range of causes of monogenic diseases such as HeFH, and perhaps also to susceptibility to common complex diseases.

### Familial hypercholesterolemia: a monogenic model of atherosclerosis

HeFH (MIM (Mendelian Inheritance in Man) 143890) affects ~ 1:500 people, with higher rates among people of Quebecois, Christian Lebanese, and Dutch South Afrikaner extraction because of genetic founder effects (Yuan et al. 2006). HeFH is an autosomal dominant disease characterized by markedly elevated plasma concentrations of LDL cholesterol, typically >95th percentile for age and sex (Brown and Goldstein 1986). Without adequate diagnosis and intervention, HeFH subjects can have an increased risk of fatal coronary heart disease (CHD) that is up to 100-fold higher than in the general population (Yuan et al. 2006). Among people with HeFH in Quebec, CHD onset is generally at around 40 years of age in men and about 50 years in

women (Yuan et al. 2006). Fortunately, treatment with statin drugs substantially reduces this risk and can improve clinical outcome (Williams et al. 1993). HeFH is not only relatively common and associated with a high risk of early CHD but is also easily treatable; it is therefore a genetic disorder that meets the World Health Organization (WHO) criteria for systematic screening (WHO 1999). However, WHO has estimated that HeFH is properly diagnosed in only ~15% of affected Canadians (WHO 1999). Since as many as 30% of patients do not survive their first myocardial infarction (MI), early detection of HeFH, using a variety of screening methods, has the potential to save many lives and prevent early morbidities among Canada's estimated 60 000 to 80 000 patients with HeFH (Yuan et al. 2006).

### Clinical phenotype of HeFH

Familial clustering of tendon xanthomas, high serum cholesterol, and early MIs was first noted almost 70 years ago (Müller 1939). In patients with HeFH, the liver's capacity to catabolize LDL cholesterol in a regulated manner is impaired because of the deficiency of LDL receptor function. LDL cholesterol residence time in plasma is therefore prolonged, and the propensity of the lipoprotein particles to undergo oxidation is increased. The modified LDL particles are taken up by macrophages by means of an unregulated scavenger receptor, which causes cholesterol-laden foam cells to form, leading to clinical manifestations. Cholesterol deposits within the skin of the eyelids are called xanthelasmas; those in connective tissue within and surrounding extensor tendons, especially the Achilles and extensor tendons of the hands, are called xanthomas; and deposits along the corneal margin are called arcus cornealis. The most serious deposits occur within arteries, where they can cause CHD and stroke (Yuan et al. 2006).

### Molecular genetics of HeFH

Brown and Goldstein (1986) discovered the LDL receptor, a cell-surface glycoprotein that binds the apolipoprotein (apo) B moiety on the LDL particle during receptor-mediated endocytosis. HeFH patients have 1 mutated copy of the *LDLR* gene, which normally is located on chromosome 19p13 and comprises 18 exons (Brown and Goldstein 1986). The catabolic defect in HeFH patients results in a doubling of plasma LDL cholesterol concentration. In extremely rare cases, about 1 in a million, patients have homozygous familial hypercholesterolemia (HoFH) because of mutations in both copies of their *LDLR* gene, with plasma LDL cholesterol concentrations increased up to 10-fold, prominent eye, skin, and tendon deposits, and vascular atherosclerotic disease in childhood (Brown and Goldstein 1986). HeFH most commonly results from a mutation in the *LDLR* (MIM 606945). At the molecular level, FH is now

commonly diagnosed using exon-by-exon sequence analysis (EBESA) of *LDLR* from genomic DNA (gDNA) (Wang et al. 2001, 2005).

### Molecular heterogeneity of HeFH

Genetic studies over the past 20 years, primarily using EBESA discovery platforms, have shown that various *LDLR* mutations out of a total of ~800 are found in most HeFH patients (Austin et al. 2004). However, this method seems to find mutations in only a portion of patients with clinically diagnosed FH (Wang et al. 2001, 2005). Part of the gap in detecting mutations in HeFH is related to its genetic heterogeneity (Rader et al. 2003). For instance, a phenocopy of HeFH called HCHOLAD2 (MIM 144010), which results from a missense mutation in *APOB* affecting the LDL receptor-binding domain of apo B-100 (MIM 107730) (Rader et al. 2003), accounts for 5%–10% of the HeFH phenotype. A rare HeFH subtype called HCHOLAD3 (MIM 603776) results from gain-of-function mutations in *PCSK9* (MIM 607786), which encodes a protease that degrades the LDL receptor intracellularly (Rader et al. 2003). A similarly rare FH phenotype with autosomal recessive inheritance is called HCHOLAR1 (MIM 603813), which results from mutations in *ARH* (MIM 605747) encoding a putative adaptor for the LDL-receptor (Rader et al. 2003). However, even after accounting for genetic heterogeneity, EBESA has found no *LDLR* mutation in many patients with clinically diagnosed FH.

### EBESA mutations in Ontario HeFH patients

In a large number of HeFH patients from Ontario we used EBESA of *LDLR* and showed that only ~50% of patients had mutations that were resolved at the single-base level, including single nucleotide mutations that caused mainly single amino acid changes, premature termination codons, and RNA splice site changes. In addition, a few small insertions and deletions were found. The spectrum of mutations in Ontario HeFH patients found by EBESA is shown above the genomic map in Fig. 1A. Sequencing of *APOB* showed that an additional ~7% of HeFH patients were heterozygous for the R3500Q receptor-binding domain mutation, which is consistent with other study samples (Yuan et al. 2006). No mutations were found in either *PCSK9* or *ARH*. Thus, about one-third of Ontario HeFH patients had no mutation in any gene detectable with the EBESA genomic DNA screening platform.

### Heterogeneity of genomic mutation types in HeFH

Another possible explanation for the gap in HeFH molecular diagnosis is inadequate sensitivity of EBESA to detect certain mutation types, such as CNVs (e.g., large genomic DNA deletions or insertions). For instance, a large genomic DNA deletion that removes a region of 1 chromosome spanning several *LDLR* exons would effectively create hemizygoty for the normal-sequence single exons present on the unaffected allele. Since only 1 allele from an affected individual would be amplified and sequenced, the resulting EBESA profile would be indistinguishable from homozygosity

for 2 normal *LDLR* alleles. MLPA is a new analytical method (Sellner and Taylor 2004) that detects CNVs, specifically larger gDNA deletions or insertions involving whole exons, that would otherwise be overlooked by EBESA. We applied the MLPA method in the ~35% of HeFH patients who did not have mutations detected by EBESA and found that about two-thirds of them had an abnormal *LDLR* MLPA pattern (Wang et al. 2005). We then showed, using genomic DNA sequence analysis of the deletion break point, that the abnormal MLPA pattern in 5 patients corresponded to a specific deletion, namely the French Canadian *LDLR*  $\Delta >15$  kb, involving the 5'-upstream translated region of *LDLR*. While *LDLR* missense mutations are most common in HeFH, MLPA abnormalities – mainly deletions – appear also to be very prevalent (Wang et al. 2005). The spectrum of mutations in Ontario HeFH patients found by MLPA is shown below the genomic map in Fig. 1B; the results suggested that MLPA might help to diagnose HeFH by detecting *LDLR* CNVs.

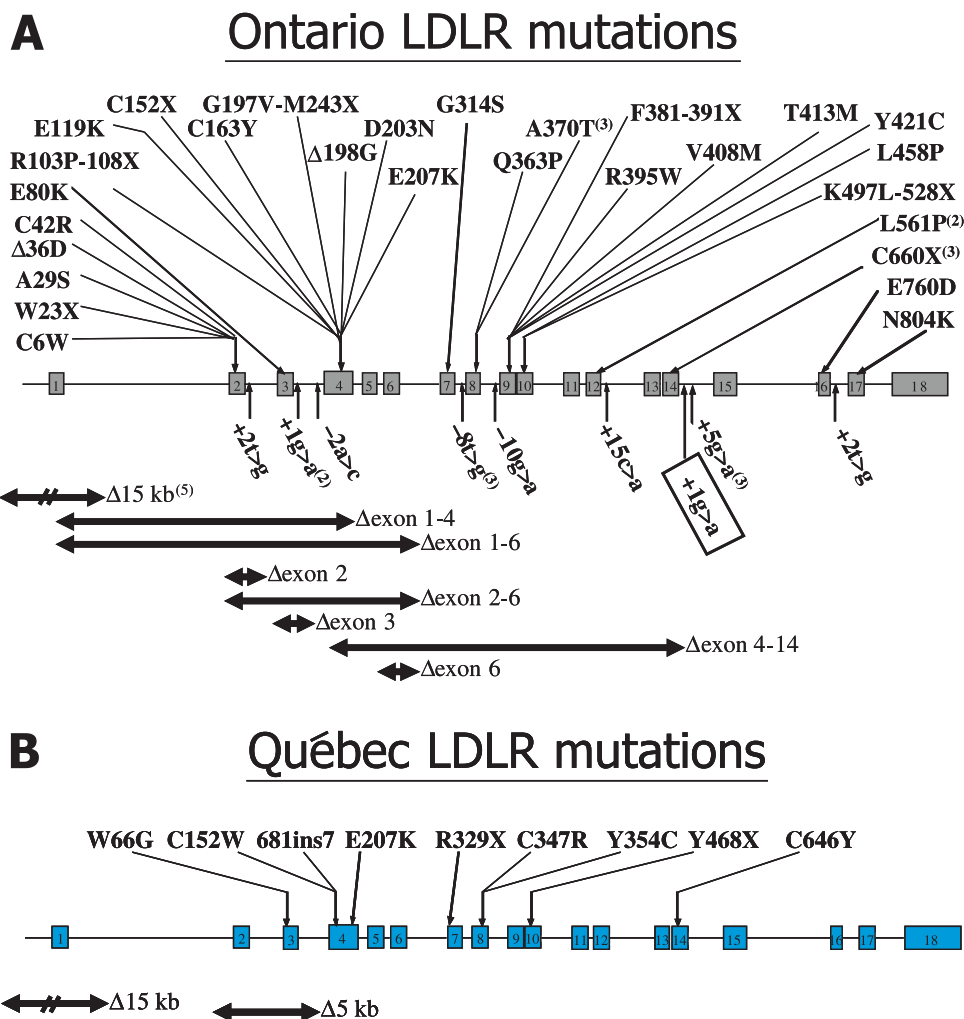
### Biochemical attributes according to molecular diagnosis

We stratified patients into molecular subgroups: (i) heterozygotes for missense mutations; (ii) heterozygotes for splicing mutations; (iii) heterozygotes for either nonsense mutations or in-frame deletions; (iv) abnormal MLPA pattern; (v) heterozygotes for *APOB* R3500Q; and (vi) no *LDLR* mutation and normal *APOB* receptor-binding domain sequence (Fig. 2). The 12 HeFH patients with abnormal *LDLR* MLPA patterns had significantly higher mean levels of untreated plasma LDL cholesterol than the other subgroups. The 9 HeFH patients who had no mutation after analysis had a mean level of untreated plasma LDL cholesterol similar to that of the 21 HeFH patients with missense mutations and the 5 HeFH patients with *APOB* R3500Q (Fig. 2). This is an example of genotype–phenotype correlation and indicates that heterozygous *LDLR* CNVs are associated with a more severe biochemical phenotype than the other types of mutations.

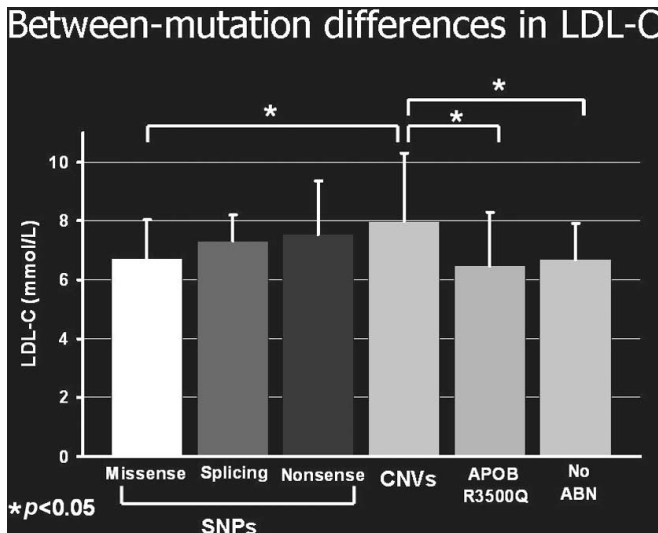
### HeFH patients with no detected mutation by sequence or CNV analysis

MLPA reduced the proportion of HeFH patients who had no apparent genomic DNA abnormality to ~10%. What might be the basis of the disease phenotype in these remaining HeFH subjects? Some of the patients might have had further distinctive *LDLR* mutation types, such as mutations deep within introns or beyond the region of the promoter and 3'-untranslated regions accessible by our EBESA and MLPA primers. Alternatively, some patients may have had another genetic etiology for their HeFH phenotype, as exemplified by the 5 patients with *APOB* R3500Q. We used EBESA to examine *PCSK9* and *ARH* in the 9 HeFH patients with no *LDLR* mutation and normal *APOB* codon 26 sequence. Mutations in *APOB*, *PCSK9*, and (or) *ARH* might still be found in some of these FH patients using MLPA to detect CNVs. Alternatively, mutations in some patients may be found in novel genes that are not known to be causative in HeFH.

**Fig. 1.** Genetic map of *LDLR* mutations in Canadian HeFH patients. Linear map showing the genomic structure of the *LDLR* gene in the centre of each panel. Filled numbered boxes represent exons. Arrows point to the mutation position. Above the map are single nucleotide and small insertion-deletion (indel) mutations. Single-letter amino codes are shown, along with mutated codon number. Splicing mutations are indicated directly under the linear map, on the diagonal. The position relative to the intron-exon boundary is shown numerically, and the mutated nucleotide is shown in lower case. Larger deletions are shown as thick bidirectional arrows further below the linear map, with the span of the deleted region indicated by text and delimited by the arrow heads. (A) Ontario *LDLR* mutations. In the 80 Ontario HeFH patients studied, 58 had mutations in the *LDLR* gene, 5 had mutations in the *APOB* gene. Among the 58 HeFH Ontario patients with *LDLR* mutations, 40 had single nucleotide mutations, and 18 had insertion-deletion (indel) mutations. There were 45 unique *LDLR* gene mutations in total with no predominant recurring mutation. When mutations occurred in more than 1 unrelated proband, the number of recurrences is indicated within superscripted parentheses. Of the 31 single-nucleotide mutations 19 were missense mutations (C6W, A29S, C42R, E80K, E199K, C163Y, D903N, E207K, G314S, Q363P, A370T [3 probands], R395W, V408 mol/L, T413 mol/L, Y421C, L458P, L561P [2 probands], E760D, and N804K); 3 were nonsense mutations (W23X, C152X, and C660X [2 probands]); and 9 were splicing mutations (intron 2 +2t > g; intron 3 +1 g > a [2 probands] and -2a > c; intron 7 -8t > g [3 probands]; intron 8 -10 g > a; intron 12 +15c > a; intron 14 +1 g > a [current proband's mutation, boxed] and +5 g > a [3 times]; and intron 16 +2t > g). Among the 14 unique indel mutations 2 were inframe deletions detected by nucleotide sequence analysis ( $\Delta$ 36D and  $\Delta$ 198G); 4 were small frameshift deletions causing premature truncations (R103P-108X, G197V-M243X, F381-391X, and K497L-528X); and 8 were larger deletions of ~300 to 15 000 nucleotide base pairs detected only by MLPA (multiplex ligation-dependent amplification analysis), with the span of the deleted region shown by bidirectional arrows:  $\Delta$ 15 kb, including exon 1 (5 times; also very common in Quebec HeFH, see (B));  $\Delta$ exon 1-4;  $\Delta$ exon 1-6;  $\Delta$ exon 2;  $\Delta$ exon 2-6;  $\Delta$ exon 3;  $\Delta$ exon 4-14; and  $\Delta$ exon 6. Each mutation was clearly heterozygous by DNA sequencing or MLPA and absent from the genomes of 100 normolipidemic subjects. (B) Quebec *LDLR* mutations. Reports of large clinical cohorts consistently showed that >90% of HeFH patients in Quebec have 1 of 11 *LDLR* mutations, most of which are highly recurrent in the population. Of the 8 Quebec mutations involving single nucleotides, 6 were missense mutations (W66G, C152Y, E207K [also found in Ontario], C347R, Y354C, and C646Y), and 2 were nonsense mutations (R329X and Y468X). Of the 3 indel mutations, 1 was a small frameshift deletion with premature truncation (681ins7), and 2 were larger deletions:  $\Delta$ 15 kb, including exon 1 (also found in Ontario patients) and  $\Delta$ 5 kb.



**Fig. 2.** Genotype–phenotype correlations in HeFH. The figure shows mean level  $\pm$  standard deviation (SD) of plasma LDL cholesterol in Ontario HeFH subjects classified according to genotype, including most of the mutations shown in Fig. 1. The lowest mean LDL cholesterol (LDL-C) concentration was seen in subjects with heterozygous *LDLR* missense mutations, but also in those with the *APOB* R3500Q mutation and individuals with HeFH but no genomic DNA mutation detected in any candidate gene screened using any method. HeFH individuals with *LDLR* splicing mutations and nonsense mutations had sequentially higher mean plasma LDL cholesterol concentrations, and subjects with CNVs involving the *LDLR* locus had the highest mean LDL cholesterol concentrations.



### Technology drives discovery of specific mutation types

Technology is an important determinant of the mutation type detected. For instance, in the era before polymerase chain reaction (PCR), the most prevalent *LDLR* mutation type found in FH was large-scale gDNA rearrangement, such as large insertions and deletions (Brown and Goldstein 1986). One reason for this was the relative enrichment within *LDLR* of sequences that were predisposed to imperfect recombination events (Hobbs et al. 1990); but another reason was that the first *LDLR* mutations found in the late 1980s were detected by Southern blotting. Southern blots were ideal for detecting large gDNA rearrangements involving  $>200$  nt of target gDNA, which was the resolution limit for DNA fragments. However, detection of small mutations was beyond the Southern blots' resolving capacity, unless a mutation altered a recognition site for a restriction endonuclease. Since Southern blotting was optimal for detecting large *LDLR* rearrangements, by the early 1990s these represented a substantial proportion of FH mutations found (Hobbs et al. 1990).

The invention of PCR made it easier to quickly detect small mutations from gDNA using relatively few resources, with less expense, effort, and radiation requirements than Southern blotting. Thus, most of the  $\sim 800$  *LDLR* mutations in FH that are in current databases are of the single-

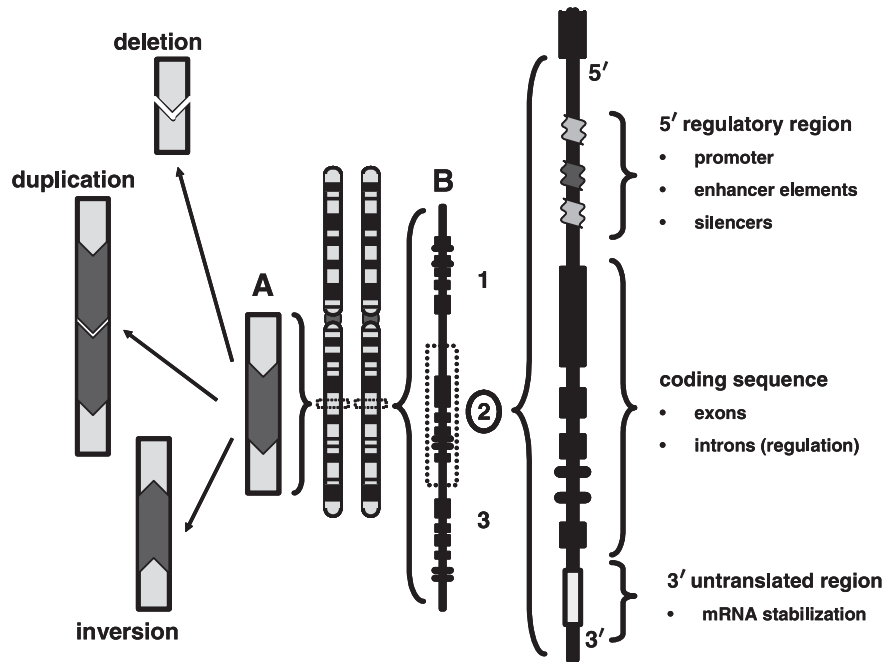
nucleotide variety discovered by PCR-based methods such as EBESA (<http://www.ucl.ac.uk/fh/>). However, PCR-based methods of interrogating gDNA are practically limited to regions of 200–700 bp. A large deletion affecting 1 *LDLR* allele may never be detected by PCR and sequence analysis, since there is no target sequence upon which amplification primers can anneal. Only the normal allele is amplified, creating effective hemizygosity for that region of genomic DNA. Analysis of the products of such a reaction in the search for a DNA change will yield a false-negative result.

MLPA is a PCR-based method that optimally detects alterations over a relatively large gDNA region without radioactivity. As a semiquantitative method, it has internal standards and controls that create an additional burden for each analysis, but this need is met by the high throughput afforded by rapid, automated fragment detection. Our cost for *LDLR* MLPA analysis per patient sample, including reagents, duplicate analyses, controls, and labour, was  $\sim$ \$150 CDN, which is probably cost-effective considering the increment in molecular diagnosis. There is no optimal procedure for data normalization and for detection of abnormal data points for MLPA (Sellner and Taylor 2004). For future clinical applications, performance of *LDLR* MLPA will need to be assessed by blinded analysis of HeFH samples that have well-defined genomic DNA rearrangements. Various reference samples, including healthy controls and subjects with other mutation types, such as CNVs, will need to be assessed to create reference standards.

### CNVs: from perceived rarity to common form of genomic variation

In contrast to single-nucleotide polymorphisms, which have been considered to be the predominant form of genomic DNA variation in quantitative terms (International HapMap Consortium 2003), large chromosomal alterations have always been thought to be relatively rare. However, recent work has shown that submicroscopic large-scale chromosomal alterations, such as insertions or duplications, deletions, inversions, and translocations, collectively referred to as CNVs, are more ubiquitous and common than was previously appreciated (Feuk et al. 2006a, 2006b). In particular, integration of CNVs into maps of other types of variation signals a new chapter in human genomic research. CNVs include variants involving  $\geq 1$  kb of gDNA (Feuk et al. 2006a, 2006b) such as (i) structural changes that are qualitatively analogous to those that previously were resolved cytogenetically, such as duplications, deletions, and translocations; and (ii) nonquantitative variants that primarily affect qualitative genomic attributes of large chromosomal segments, such as inversions with no net change in copy number. It is now relatively common to find that 1 of a homologous pair of chromosomes can be up to a million nucleotides and up to 20 genes shorter than the other. While this understanding is nascent, the biological and medical implications of these dramatic variations could affect our conventional concepts of “common”, “normal”, “healthy”, and “diseased”.

**Fig. 3.** Overview of CNV types. The left side of the figure shows some of the types of CNVs, and the right side shows some of the consequences of CNVs at the level of the gene. Segment A represents the normal structure of a region of 1 of a homologous pair of chromosomes. The type of genomic alterations of the normal structure that lead to deletion, duplication, and inversion are shown at the far left. Segment B represents the normal structure of a chromosomal locus that contains a cluster of 3 genes (1, 2, and 3). Segment C shows the detailed normal structure of gene 2, including key functional elements, such as the 5' regulatory region, with promoter, enhancer, and silencer elements, 3' untranslated regions that can regulate message stability, and also the intron–exon structure of the gene. Any of these structural and functional elements can become involved in a CNV, with a range of functional consequences depending on the size of the variant and the affected functional domains.



### CNVs and their potential to explain genetic susceptibility to common diseases

The size and ubiquity of CNVs suggests a potential role for susceptibility to common complex, polygenic diseases such as diabetes, cancer, schizophrenia, and cardiovascular disease, if only for the somewhat simplistic reason that certain CNVs span regions containing many genes. While this is promising as a hypothesis for human genetic research, the ultimate proof of the involvement of CNVs in common disease phenotypes will require large-scale studies involving well-phenotyped cohorts and comprehensive, robust methods to classify individuals according to their CNV status.

### Quantitative mechanisms for CNVs in disease

The potential etiologic mechanisms for CNV in disease have both a quantitative and a qualitative dimension (Fig. 3). CNVs can result in phenotypic diversity by altering transcriptional, and presumably translational, levels of genes and their products. For instance, haploinsufficiency occurs when 1 copy of a dosage-sensitive gene has been deleted, producing deficient function that cannot be rescued. Haploinsufficiency also applies to genomic deletions that do not result in a monogenic disease, being instead

present in healthy and apparently normal individuals but requiring additional genetic and (or) environmental factors for the abnormal phenotype to be expressed later in life. This has already been demonstrated for certain CNV genes both at the transcriptional (Hollox et al. 2003; McCarroll et al. 2006) and translational (Gonzalez et al. 2005) levels. Also, gene dosage increases can also be pathological in humans (Aitman et al. 2006). Correlation of CNV with mRNA and protein levels and phenotypic effects might further be modulated by tissue specificity and (or) developmental factors.

### Qualitative aspects of disease-causing CNVs

The mechanisms whereby CNVs produce disease also depend on the nature of the genomic region affected and have been discussed at length (Feuk et al. 2006a, 2006b). They include the following: (i) involvement of several genes at a gene cluster, with the potential for multiple effects on intermediate pathways and possibly phenotypes; (ii) involvement of a single gene with increased gene dosage or allelic loss leading to increased or decreased gene expression, respectively; (iii) involvement of a single gene in the presence of a risk allele, with loss of wild-type allele leading to hemizygosity for the risk allele but gain of a wild-type allele perhaps decreasing risk; (iv) various consequences from involvement of noncoding regulatory regions, e.g., loss of a

cis-enhancer element or binding site leading to decreased gene expression, loss of a silencer binding site leading to increased gene expression, or loss of a 3'-untranslated sequence leading to decreased mRNA stability.

### Potential clinical implications of CNVs

The probability of finding normal genomic variants in screened samples is now markedly increased, given the sensitivity of the new methods and the ubiquity of CNVs of >1 kb in the human genome. We have already shown how testing for CNVs improves the diagnostic range of human mutations in HeFH (Wang et al. 2005). In a clinical research setting, how will the management and counseling of a patient and his or her family unfold with, and without, taking CNVs into account? Will ethical issues arising from analysis of CNVs simply mirror past issues encountered when using cytogenetic methods? Could past medical genetic diagnosis be revised in light of new knowledge afforded by CNVs? Should archived specimens be re-evaluated in light of the CNV map? These questions and others (Daar et al. 2006) will require attention over the coming months and years.

### Conclusion

The ability to screen for pathological CNVs has markedly increased the number of *LDLR* mutations detected in Ontario HeFH subjects. Furthermore, the CNV map in healthy "normal" individuals adds a new dimension to the study of the human genome with immediate implications for biology and perhaps longer-term implications for medicine. Future genomic mapping experiments and genome-wide association analyses, and their respective detection technologies, will need to account for the presence of CNVs. Current platforms to study the genome may need to be redesigned to either maximize detection of CNVs or minimize their interference with the detection of other forms of genomic variation. As the "personal genome" moves closer to becoming a reality, it will be important to account for CNVs along with other types of mutations and variants and then to interpret their biological meaning for any individual. Finally, the new CNV discoveries should be placed within the wider context of contemporary discourse on genomic variation studies and their biological, health, and societal implications. However, this could wait until more extensive characterization is complete to fully understand the implications and potential applications of human genomic CNVs.

### Acknowledgements

Supported by a the Jacob J. Wolfe Chair in Functional Genomics, the Edith Schulich Vinet Canada Research Chair (Tier I) in Human Genetics, a Career Investigator award from the Heart and Stroke Foundation of Ontario, and operating grants from the Canadian Institutes for Health Research, the Heart and Stroke Foundation of Ontario, the Ontario Research and Development Challenge Fund (99-0507), and Genome Canada through the Ontario Genomics Institute.

### References

- Aitman, T.J., Dong, R., Vyse, T.J., Norsworthy, P.J., Johnson, M.D., Smith, J., et al. 2006. Copy number polymorphism in *Fcgr3* predisposes to glomerulonephritis in rats and humans. *Nature (London)*, **439**: 851–855. doi:10.1038/nature04489. PMID: 16482158.
- Austin, M.A., Hutter, C.M., Zimmern, R.L., and Humphries, S.E. 2004. Genetic causes of monogenic heterozygous familial hypercholesterolemia: a HuGE prevalence review. *Am. J. Epidemiol.* **160**: 407–420. doi:10.1093/aje/kwh236. PMID:15321837.
- Brown, M.S., and Goldstein, J.L. 1986. A receptor-mediated pathway for cholesterol homeostasis. *Science (Washington, D.C.)*, **232**: 34–47. doi:10.1126/science.3513311. PMID:3513311.
- Daar, A.S., Scherer, S.W., and Hegele, R.A. 2006. Implications of copy-number variation in the human genome: a time for questions. *Nat. Rev. Genet.* **7**: 414. doi:10.1038/nrg1884.
- Feuk, L., Carson, A.R., and Scherer, S.W. 2006a. Structural variation in the human genome. *Nat. Rev. Genet.* **7**: 85–97. doi:10.1038/nrg1767. PMID:16418744.
- Feuk, L., Marshall, C.R., Wintle, R.F., and Scherer, S.W. 2006b. Structural variants: changing the landscape of chromosomes and design of disease studies. *Hum. Mol. Genet.* **15**: R57–R66. doi:10.1093/hmg/ddl057. PMID:16651370.
- Gonzalez, E., Kulkarni, H., Bolivar, H., Mangano, A., Sanchez, R., Catano, G., et al. 2005. The influence of *CCL3L1* gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science (Washington, D.C.)*, **307**: 1434–1440. doi:10.1126/science.1101160. PMID:15637236.
- Hobbs, H.H., Russell, D.W., Brown, M.S., and Goldstein, J.L. 1990. The LDL receptor locus in familial hypercholesterolemia: mutational analysis of a membrane protein. *Annu. Rev. Genet.* **24**: 133–170. doi:10.1146/annurev.ge.24.120190.001025. PMID:2088165.
- Hollox, E.J., Armour, J.A., and Barber, J.C. 2003. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. *Am. J. Hum. Genet.* **73**: 591–600. doi:10.1086/378157. PMID:12916016.
- International HapMap Consortium. 2003. The International HapMap Project. *Nature (London)*, **426**: 789–796. doi:10.1038/nature02168. PMID:14685227.
- Lusis, A.J. 2000. Atherosclerosis. *Nature (London)*, **407**: 233–241. doi:10.1038/35025203. PMID:11001066.
- McCarroll, S.A., Hadnott, T.N., Perry, G.H., Sabeti, P.C., Zody, M.C., Barrett, J.C., et al. 2006. Common deletion polymorphisms in the human genome. *Nat. Genet.* **38**: 86–92. doi:10.1038/ng1696. PMID:16468122.
- Müller, C. 1939. Xanthomata, hypercholesterolemia, angina pectoris. *Acta Med. Scand.* **89**: 75–84.
- Rader, D.J., Cohen, J., and Hobbs, H.H. 2003. Monogenic hypercholesterolemia: new insights into pathogenesis and treatment. *J. Clin. Invest.* **111**: 1795–1803. doi:10.1172/JCI200318925. PMID:12813012.
- Sellner, L.N., and Taylor, G.R. 2004. MLPA and MAPH: new techniques for detection of gene deletions. *Hum. Mutat.* **23**: 413–419. doi:10.1002/humu.20035. PMID:15108271.
- Wang, J., Huff, E., Janecka, L., and Hegele, R.A. 2001. Low density lipoprotein receptor (*LDLR*) gene mutations in Canadian subjects with familial hypercholesterolemia, but not of French descent. *Hum. Mutat.* **18**: 359. doi:10.1002/humu.1205. PMID:11668627.
- Wang, J., Ban, M.R., and Hegele, R.A. 2005. Multiplex ligation-dependent probe amplification of *LDLR* enhances molecular diagnosis of familial hypercholesterolemia. *J. Lipid Res.* **46**: 366–372. PMID:15576851.

- World Health Organization. 1999. DoNDP, familial hypercholesterolaemia: report of a second WHO consultation. *Edited by Human Genetics*. WHO, Geneva.
- Williams, R.R., Hunt, S.C., Schumacher, M.C., Hegele, R.A., Lepert, M.F., Ludwig, E.H., and Hopkins, P.N. 1993. Diagnosing heterozygous familial hypercholesterolemia using new practical criteria validated by molecular genetics. *Am. J. Cardiol.* **72**: 171–176. doi:10.1016/0002-9149(93)90155-6. PMID:8328379.
- Yuan, G., Wang, J., and Hegele, R.A. 2006. Heterozygous familial hypercholesterolemia: an underrecognized cause of early cardiovascular disease. *CMAJ*, **174**: 1124–1129. PMID:16606962.