

Duplications, deletions, and single-nucleotide variations: the complexity of genetic arithmetic

Vincent M. Riccardi, MD¹ and James R. Lupski, MD, PhD^{2,3}

A “genetic disease” in a very important sense is multiple disorders, manifested by the multiple genetic variations that underlie a particular genetic disease. Sometimes the disorder is caused by subtraction, the loss of one or more genetic components. Loss of gene function may be due to a single-nucleotide variation, a change in the DNA alphabet within the gene; a whole-gene deletion as a form of copy number variation (CNV), the DNA of part or all of the gene being absent entirely; or an entire chromosome or segment thereof is lost. Sometimes, the genetic disease is caused by addition, the gain of one or more genetic components. An additional chromosome 21 causes Down syndrome. We are also aware of the pathology caused by more-subtle additions of genetic material to an individual’s genome, particularly gene duplication (e.g., Charcot-Marie-Tooth disease type 1A) and expansion of tandem repeats (e.g., Huntington disease). Furthermore, the genetic changes may involve only a portion of the affected person’s cells and the patient is mosaic for the gain or loss.

One of the first patients with a neurofibromatosis type 1 (NF1) microdeletion was reported in 1990 and 1992.¹ Another case was published in 1990.² Since then, many cases have been studied intensely in Germany^{3,4} and elsewhere.⁵ Detailed clinical, cell biology, and genomic data are now readily available on many patients with one of three *NF1* microdeletion disorders, one of which usually involves mosaicism, indicating both meiotic and postzygotic mitotic origins of the deletions. These microdeletion cases and recently reported microduplication cases^{6,7} are now poised to teach us something.

In the May 2012 issue of *Genetics in Medicine*, Moles et al.⁷ reported rigorous genomic data on seven patients with “microduplications” of that portion of chromosome band 17q11.2, which, when deleted, accounts for several percent of patients specified to have the disorder NF1. Their data corroborated and complemented those from an earlier publication involving the “same” duplication in seven persons in a Belgian family.⁶ The 14 patients (from eight families) included 12 females and two males; in the Belgian family, six of the seven persons with the duplication were females. Clinical features ranged from “none” to moderately severe but nonspecific “developmental delay/learning disability, facial dysmorphisms, and seizures,” and in the Belgian family, “enamel hypoplasia and early-onset baldness.” Other than noting that no person with the

microduplication had a neurofibroma, specific comparisons with the NF1 clinical spectrum were not tabulated.

From March 2004 through April 2011, Moles et al.⁷ performed microarray analysis on 48,817 individuals, most often with a clinical indication of “intellectual disability, developmental delay, or multiple congenital anomalies.” The yield was one *NF1* microduplication per 6,974 specimens per year. Although this duplication CNV would be considered a “rare” variant, this high yield suggests that there are many other similar patients waiting to be identified.

NF1 (OMIM no. 162200) is one of the most intensely studied human genetic diseases in the 21st century. NF1 is quite common, with about 1 in 3,000 people having the disorder, almost as common as Charcot-Marie-Tooth neuropathy (CMT; OMIM no. 178200), which occurs in 1 in 2,500 Scandinavians. A duplication CNV—the CMT1A duplication⁸—accounts for the majority of patients with CMT1. NF1 and CMT occur at frequencies high enough to be observed by chance in the same individual.⁹ Both NF1 and CMT are quite variable from one family to another and from one person to another within a family. Moreover, early approaches to analyzing NF1 variability proceeded with the tacit assumption that the explanation for the intrafamilial variation likely also accounted for the interfamilial variations.

The article by Moles et al.⁷ is relevant to NF1 in particular and genetic/genomic pathogenesis in general, precisely because this variation assumption has proved to be wrong. Specifically, NF1 results not only from many types of inherited and spontaneous intragenic mutations but also from relatively frequent somatic mosaicism and at least three different types of microdeletion syndromes. NF1 microdeletion results in at least four different types of disorders: zygotic intragenic single-nucleotide variation mutation inherited from either parent or occurring spontaneously in meiosis (90% paternal); zygotic whole-gene microdeletion inherited from either parent or occurring spontaneously in meiosis (90% maternal); postzygotic mitotic (mosaic) intragenic single-nucleotide variation mutations; and postzygotic mitotic (mosaic) whole-gene microdeletion that can result in segmental NF1. The ongoing accumulation of massive knowledge about the clinical, cell biological, and genomic elements of NF1, especially data associated with the microdeletion forms, when combined with the new and developing data from the patients and families with “duplications” of these same genomic sequences, is

¹The Neurofibromatosis Institute, La Crescenta, California, USA; ²Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA; ³Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA. Correspondence: Vincent M. Riccardi (Riccardi@medconsumer.com)

a virtual gold mine for the clinical and research genetic communities. Much can be learned from the genetic arithmetic of these subtractions and additions.

For example, if NF1 is a Rasopathy, are the patients with NF1 microduplication considered to have a Rasopathy disorder? What about Ras activity in cells derived from these patients, particularly cells of neural crest origin, such as melanocytes and Schwann cells? If the *NF1* gene is a “tumor suppressor,” will the patients with gene duplication be detectably more resistant to oncogenesis? As has been the case for studies of patients with *PMP22* duplication causing Charcot-Marie-Tooth disease type 1A (CMT1A; OMIM no. 118220) and deletion resulting in hereditary neuropathy with liability to pressure palsies (HNPP; OMIM no. 162500), innumerable other questions and considerations are easily brought to the fore as the patients with NF1 microdeletion and those with NF1 microduplication are compared.

The work of Moles et al.⁷ demonstrates that, when we define diseases on the basis of genomic etiology, the ultimate phenotypic spectrum is greater than that anticipated from the sum of the original cases’ clinical findings. As we ascertain cases based on genotype (increasingly common in a genomic era), the expansion and redefinition of phenotypes will likely be a salient feature of clinical genetics. The spectrum will likely include an apparent absence of findings even when the particular duplication variant is present. There appears to be much more genetic variation, including minimal penetrance, than we can presently account for. For example, the specific configuration of a duplication may matter such that sometimes $1 + 2 = 3$, an abnormal phenotype, and sometimes $1 + 2 = 2$, an ostensibly normal phenotype.

Furthermore, such variation of expression and potentially reduced penetrance confound proof of causality. Moles et al.⁷ describe seven individuals with a rare duplication CNV and a clinical problem that brought them to their physician. The physician’s clinical judgment prompted diagnostic testing. Nevertheless, Moles et al.⁷ show no significant difference between the frequency of the *NF1* duplication in the patient population versus the control “normal” population ($P = 0.26$). Does this mean that the duplication variant does not cause disease? How do we interpret such statistics? It is likely that genome-wide association studies focusing on common variants and segregation of marker genotypes in populations do not apply when rare variants and de novo gene mutations contribute in a substantial way to a (disease) trait not under selection pressure?¹⁰ The data emerging from genome-wide association studies, including array comparative genomic hybridization and exome sequencing, suggest that both rare variants and new mutations contribute significantly to medically actionable problems. These observations, embedded in the concept of “clan genomics,”¹⁰ suggest a tremendous clinical utility for genome-wide assays: the most important thing for the patients and their clinical care may be the recognition of rare variants (additions or subtractions) that have arisen among their immediate relatives or the probands themselves.¹⁰

There is a need to consider a new genetic arithmetic. Even though there is a lack of statistical significance for association ($P = 0.26$), Moles et al.⁷ have convinced us of the causality of

this duplication CNV. However, with so many of their ostensibly “unaffected” patients bearing the genetic change, we can see that it is not merely a matter of adding or subtracting DNA segments (including traditional genes defined on the basis of the Central Dogma).¹¹ The first step forward is performing careful comparisons of the patients with *NF1* microdeletion and those with *NF1* microduplication, the latter of whom have a triplication of the “normal” *NF1* gene; extensive analyses must be performed using clinical, cell biology, and genomic data. The next step is to refine the language and symbols of manipulating these data. Is one gene on one chromosome plus two genes on the other chromosome a matter of duplication (a chromosomal or cytogenetic attribution) or a matter of triplication (a genomic attribution, respecting the total number of the genes in a cell)? It could be either, depending on the questions asked and the points to be made. The questions and answers are daunting, but we consider them to be central to assisting immediately patients with all types of NF1 and to make more sense of the human genome in general. For example, a homozygous duplication (gene dosage $N = 2 + 2 = 4$) may not be equivalent to a heterozygous triplication (gene dosage of $N = 3 + 1 = 4$), even though, by ordinary arithmetic, the summed gene dosage is the “same” for both scenarios ($N = 2 + 2 = 3 + 1 = 4$). Much as particle physicists have had to go beyond classical Newtonian physics, with recourse to counterintuitive number crunching of quantum mechanics and relativity, so too modern geneticists may need to use differently the numbers that are genetic data in a genomics world.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Kayes LM, Riccardi VM, Burke W, Bennett RL, Stephens K. Large *de novo* DNA deletion in a patient with sporadic neurofibromatosis 1, mental retardation, and dysmorphism. *J Med Genet* 1992;29:686–690.
2. Upadhyaya M, Cheryson A, Broadhead W, et al. A 90 kb DNA deletion associated with neurofibromatosis type 1. *J Med Genet* 1990;27:738–741.
3. Kehrer-Sawatzki H, Vogt J, Mußotter T, Kluwe L, Cooper DN, Mautner VF. Dissecting the clinical phenotype associated with mosaic type-2 NF1 microdeletions. *Neurogenetics* 2012;13:229–236.
4. Mautner VF, Kluwe L, Friedrich RE, et al. Clinical characterisation of 29 neurofibromatosis type-1 patients with molecularly ascertained 1.4 Mb type-1 NF1 deletions. *J Med Genet* 2010;47:623–630.
5. Messiaen L, Vogt J, Bengesser K, et al. Mosaic type-1 NF1 microdeletions as a cause of both generalized and segmental neurofibromatosis type-1 (NF1). *Hum Mutat* 2011;32:213–219.
6. Grisart B, Rack K, Vidrequin S, et al. NF1 microduplication first clinical report: association with mild mental retardation, early onset of baldness and dental enamel hypoplasia? *Eur J Hum Genet* 2008;16:305–311.
7. Moles KJ, Gowans GC, Gedela S, et al. NF1 microduplications: identification of seven nonrelated individuals provides further characterization of the phenotype. *Genet Med* 2012;14:508–514.
8. Lupski JR, de Oca-Luna RM, Slaugenhaupt S, et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell* 1991;66:219–232.
9. Lupski JR, Pentao L, Williams LL, Patel PI. Stable inheritance of the CMT1A DNA duplication in two patients with CMT1 and NF1. *Am J Med Genet* 1993;45:92–96.
10. Lupski JR, Belmont JW, Boerwinkle E, Gibbs RA. Clan genomics and the complex architecture of human disease. *Cell* 2011;147:32–43.
11. Riccardi, VM. Does defective neurofibromin alone explain NF1 pathogenesis? *J Jap Soc Recklinghausen Dis* (2012) (in press).