

Corina Anastasaki, PhD  
Lu Q. Le, MD, PhD  
Robert A. Kesterson, PhD  
David H. Gutmann, MD,  
PhD

*Neurol Genet*  
2017;3:e169; doi: 10.1212/  
NXG.000000000000169

## UPDATED NOMENCLATURE FOR HUMAN AND MOUSE NEUROFIBROMATOSIS TYPE 1 GENES

**OPEN**

Neurofibromatosis type 1 (NF1; OMIM 162200) is one of the most common neurogenetic conditions, affecting 1 in 3,000 people worldwide. Characterized by a propensity to develop nervous system tumors, learning and behavioral deficits, and pigmentary abnormalities, NF1 is caused by a germline sequence alteration in the *NFI* gene (OMIM: 613113; chromosome 17q11.2). In addition, somatic *NFI* sequence changes have been reported in numerous other cancers,<sup>1</sup> extending their importance to malignancies occurring in the general population.

The *NFI* gene, comprising 57 exons, was initially identified, and its complete coding sequence was assembled over 25 years ago.<sup>2</sup> In the ensuing decades, 4 alternatively spliced exons (9a, 10a-2, 23a, and 48a) were discovered and added to a large number of exons with temporary numerical assignments (e.g., 10a2 or 23.2). This has led to considerable confusion and inconsistencies in clinical and scientific publications, which also has critical implications for *NFI* DNA sequence alteration reporting as part of patient management, genotype-phenotype correlations, and the interpretation of small animal models engineered with patient-specific sequence changes. To provide a unified annotation system for the *NFI* gene, the mouse and human genes were aligned and assembled using published sequences for sequential numbering and comparison.

The full-length cDNA sequences of *NFI* (8457 nucleotides; figure e-1 at [Neurology.org/ng](http://Neurology.org/ng)) and *Nf1* (8463 nucleotides; figure e-2) genes were initially aligned using Nucleotide Basic Local Alignment Search Tool (BLAST) (NIH) optimized for megablast, revealing 92% sequence identity with 0.08% mismatches (figure e-2). Excluding the alternatively spliced exons, the ubiquitously expressed 57 exons were consecutively numbered (figure 1 and tables e1–e3). As such, the full-length human *NFI* transcript, NF1-002 (ENST00000356175.7), contains 57 exons and encodes 2818 amino acids, while the corresponding full-length mouse *Nf1* transcript, Nf1-003 (ENSMUST00000108251.8), comprises 57 exons and encodes 2,820 amino acids, with 2 additional amino acids encoded by exon 17. Alignment of

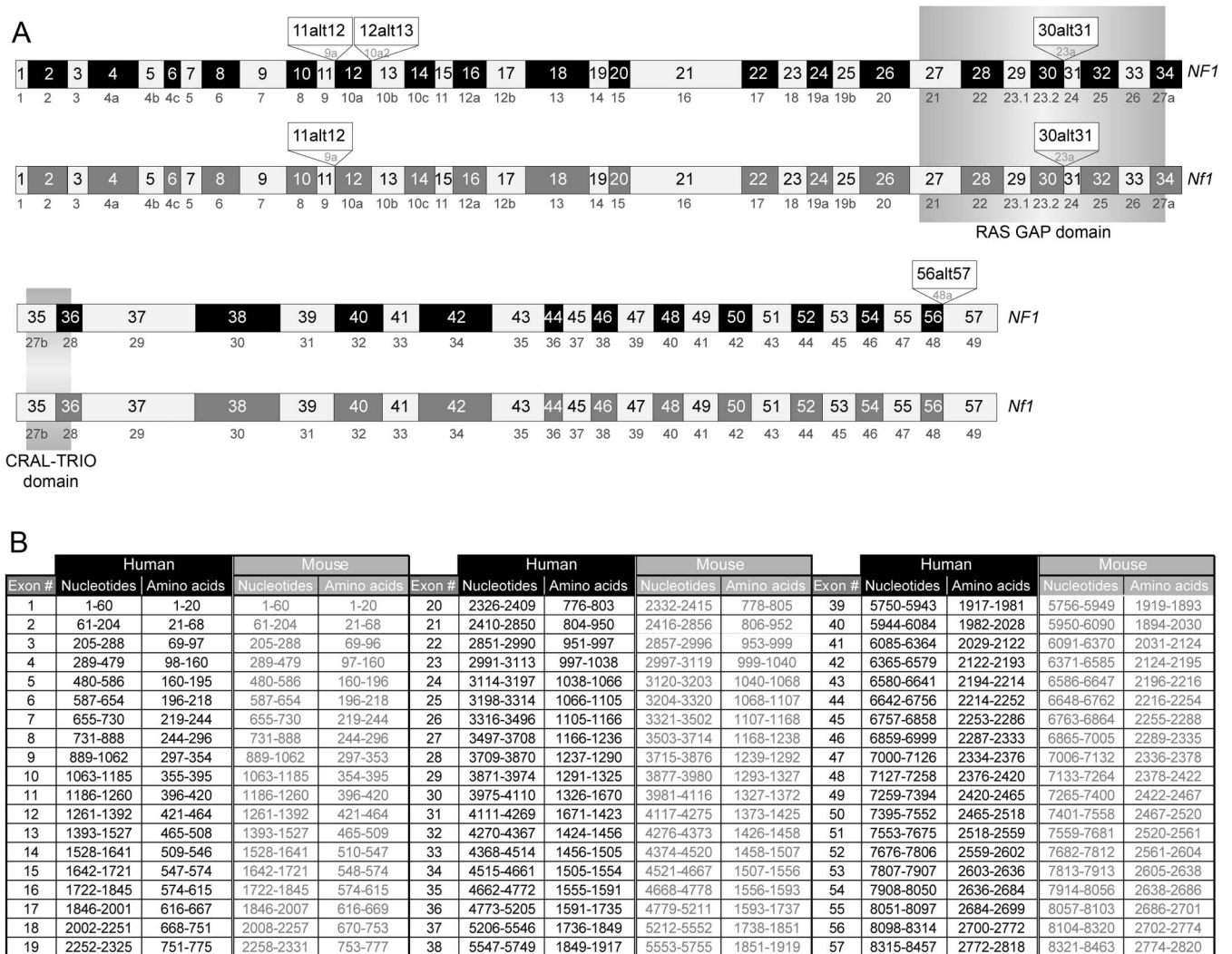
the predicted amino acid sequences of the human (RefSeq NM\_000267; UniProt P21359) and mouse (UniProt Q04690) full-length transcripts using protein BLAST revealed 98% amino acid sequence identity (figure e-3).

Next, we examined the previously published alternatively spliced exons: exon 9a contains a 30-nucleotide sequence encoding 10 amino acids and is predominantly expressed in postmitotic brain neurons,<sup>3</sup> whereas the muscle-specific exon 48a contains 54 nucleotides that encode 18 amino acids at the carboxyl terminus of the protein.<sup>4</sup> The 63-nucleotide exon 23a encodes 21 amino acids that are inserted within the RAS-GTPase-activating protein (GAP)-related domain of the *NFI* protein (neurofibromin), resulting in diminished RAS-GAP activity.<sup>4</sup> The most recently described alternatively spliced exon, exon 10a-2, contains 45 nucleotides and encodes 15 amino acids, is located at the amino terminus of neurofibromin, and this sequence has been hypothesized to direct intracellular membrane targeting by virtue of its predicted transmembrane domain.<sup>5</sup> Using published alternatively spliced exon sequences (figure e-1), BLAST searches were performed, and these alternatively spliced exons were renumbered and named according to their location (figure 1)—between exons 11–12 (11alt12; formerly 9a), exons 12–13 (12alt13; formerly 10a-2), exons 30–31 (30alt31; formerly 23a), and exons 56–57 (56alt57; formerly 48a). In this regard, we propose that the nucleotides (and the encoded amino acids) of these alternatively spliced exons be numbered 11alt12\_1–30 (1–10), 12alt13\_1–45 (1–15), 30alt31\_1–63 (1–21), and 56alt57\_1–54 (1–18). Possible new alternatively spliced exons should be similarly named based on their location.

11alt12 is 100% identical to both the cDNA and amino acid sequences of the predicted *Mus musculus* transcript variant X7 (mRNA XM\_006532442.3), while 30alt31 is 98.4% identical to the cDNA and 95.2% identical to the amino acid sequence (containing a single conservative lysine-to-arginine change) of *Mus musculus Nf1* (ENSMUST00000071325.8) (figures e-4 and e-5). Notably, the 12alt13 and 56alt57 human sequences were not found in either mouse transcriptome or genome databases following

Supplemental data at  
[Neurology.org/ng](http://Neurology.org/ng)

**Figure 1** Proposed nomenclature for human *NF1* and mouse *Nf1* genes



(A) Schematic diagram illustrating the 57 exons of the human (top panel) and mouse (bottom panel) full-length transcripts, as well as the positions of the alternatively spliced exons (11alt12, 12alt13, 30alt31, and 56alt57). The proposed numbering system is denoted in each exon, while the prior numbering system is indicated below each exon. The RAS-GTPase-activating protein and the CRAL-TRIO domains are highlighted in gray. (B) Table detailing the nucleotides and amino acids spanning each of the 57 exons in human *NF1* and mouse *Nf1* genes. CRAL-TRIO domain = cellular retinaldehyde-binding protein and TRIO guanine exchange factor structural domain.

megablast, discontinuous megablast, or BLASTN searches (figures e-4 and e-5). Based on the absence of 12alt13 and 56alt57 sequences in the mouse, further studies will be required to define their roles in neurofibromin function. It should be noted that neither 12alt13 nor 56alt57 has been reported in other reference model organisms (rat or zebrafish; data not shown).

Taken together, we assembled and updated the human *NF1* and mouse *Nf1* sequences to facilitate greater consistency in both clinical reporting and basic research studies on NF1. The clarification of the nomenclature for the *NF1* gene is a necessary step to enable standardization of exon, nucleotide, and amino acid numbering, which is particularly important for reporting patient-specific DNA sequence variants. In addition, implementation of this proposed

numbering system will provide better calibration of human genetic testing results with *Nf1* preclinical precision medicine models developed using genetically engineered mice<sup>6</sup> and induced pluripotent stem cells.<sup>7</sup>

*From the Department of Neurology (C.A., D.H.G.), Washington University in St. Louis, MO; Department of Dermatology (L.Q.L.), University of Texas Southwestern Medical Center, Dallas; and Department of Genetics (R.A.K.), The University of Alabama at Birmingham.*

*Author contributions: C.A. performed the analyses, assembled the figures and tables, and helped write the manuscript. L.Q.L. and R.A.K. were involved in the design of the study and edited the manuscript. D.H.G. wrote the manuscript with C.A. and oversaw the study at all stages.*

*Study funding: D.H.G. is supported by a Research Program Award from the National Institute of Neurological Disorders and Stroke (1-R35-NS097211-01). The authors also acknowledge support from The Giorgio Foundation.*

*Disclosure: C. Anastasaki and L.Q. Le report no disclosures. R.A. Kesterson has served on the editorial board of Frontiers in Genetics and has received research support from UAB Hepato/Renal Fibrocytic Diseases Core Center, UAB Comprehensive Cancer Center, UAB Center for Nutrition and Obesity Research, UAB Rheumatic Disease Core Center, UAB Diabetes Research Center, and the Giorgio Foundation. D.H. Gutmann has served on scientific advisory boards for the American Association for Cancer Research and the National Institute for Neurological Disorders and Stroke Advisory Council; has served on the editorial boards of Glia and Familial Cancer; holds patents for Identification of the NF1 gene and protein and Identification of NF1 protein as mTOR regulator; has received research support from the US Army Department of Defense, NCI Tumor Microenvironment Network, NIH, the Giorgio Foundation, and Children's Tumor Foundation; and receives royalty payments from the University of Michigan. Go to Neurology.org/ng for full disclosure forms. The Article Processing Charge was funded by the authors.*

*This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.*

*Received February 10, 2017. Accepted in final form April 20, 2017.*

*Correspondence to Dr. Gutmann: gutmann@wustl.edu*

1. Patil S, Chamberlain RS. Neoplasms associated with germline and somatic NF1 gene mutations. *Oncologist* 2012;17:101–116.
2. Marchuk DA, Saulino AM, Tavakkol R, et al. cDNA cloning of the type 1 neurofibromatosis gene: complete sequence of the NF1 gene product. *Genomics* 1991;1:931–940.
3. Geist RT, Gutmann DH. Expression of a developmentally-regulated neuron-specific isoform of the neurofibromatosis 1 (NF1) gene. *Neurosci Lett* 1996;211:85–88.
4. Gutmann DH, Andersen LB, Cole JL, Swaroop M, Collins FS. An alternatively-spliced mRNA in the carboxy terminus of the neurofibromatosis type 1 (NF1) gene is expressed in muscle. *Hum Mol Genet* 1993;2:989–992.
5. Kaufmann D, Müller R, Kenner O, et al. The N-terminal splice product NF1-10a-2 of the NF1 gene codes for a transmembrane segment. *Biochem Biophys Res Commun* 2002;294:496–503.
6. Li K, Turner AN, Chen M, et al. Mice with missense and nonsense NF1 mutations display divergent phenotypes compared with human neurofibromatosis type I. *Dis Model Mech* 2016;9:759–767.
7. Anastasaki C, Woo AS, Messiaen LM, Gutmann DH. Elucidating the impact of neurofibromatosis-1 germline mutations on neurofibromin function and dopamine-based learning. *Hum Mol Genet* 2015;24:3518–3528.