CORRESPONDENCE

Pitfalls of automated comparative sequence analysis as a single platform for routine clinical testing for NF1

Mattocks *et al*[1] have used direct DNA sequencing and comparative sequence analysis to study patients with neurofibromatosis type 1 (NF1) and claim this study ''achieved the highest recorded mutation detection rate using a single technique for this gene.'' As a key point, the paper states that they studied 91 subjects fulfilling the NIH NF1 diagnostic criteria and achieved a mutation detection rate of 89% using automated comparative sequence analysis. They continue by saying ''This detection rate is the highest for a single technique and is therefore appropriate for routine clinical practice.''

When developing genetic tests, especially for large and complex genes such as *NF1*, a large cohort of patients needs to be studied in a comprehensive way in order to fully understand the spectrum of mutations present in that gene. From our experience, it is of utmost importance to analyse the *complete* gene for the presence of all possible alterations that may result in a premature stop codon at the mRNA level.[2] A significant fraction of the mutations in the *NF1* gene cause aberrant splicing, and many of them are caused by alterations outside the canonically conserved AG/GT acceptor and donor sequences and even reside deep in the large introns.[2–4] Also, various exonic mutations mimicking nonsense, missense, and even silent mutations at the genomic level have been described that are splicing mutations and exert their effect by creating a novel splice donor or acceptor or affect the function of an exonic splicing enhancer (ESE) or exonic splicing silencer (ESS).[2 5] Although we have now studied over 600 patients fulfilling the NIH criteria using multiple complementary techniques, we are still challenged and surprised by the diversity of mutations leading to this disorder.

There have, unfortunately, been some examples in published reports where an alteration is claimed to be a pathogenic mutation, and where later on this statement needs to be revoked as the alteration is proven to be an innocent polymorphism.[6] For thousands of hereditary disorders for which the genes have been cloned, patients await the availability of a *reliable* and *sensitive* diagnostic test and clinical molecular genetic laboratories worldwide rely on published reports to help distinguish a polymorphism or rare benign variant from a deleterious mutation.

This distinction is of the utmost importance and has major ethical implications with respect to the genetic counseling of patients seeking diagnostic, pre-/oligosymptomatic, and prenatal testing.

There is a need for a reliable and sensitive genetic test for the *NF1* gene, to help resolve diagnostic dilemmas in patients not fulfilling the NIH diagnostic criteria, especially young children but also atypical patients, to determine the affection status of family members of an affected person and to carry out prenatal or preimplantation diagnosis, if desired. NF1 is a progressive disorder and many features increase in frequency with age. Café au lait spots are often the first signs of NF1 and may already be present at birth, increasing in number during the first years of life. Only about half the patients with sporadic NF1 fulfill the NIH diagnostic criteria by one year of age and 5% still will not fulfil these criteria by the age of eight.[7] Waiting for more symptoms to appear with time in order to ascertain the diagnosis on a clinical basis can be very stressful for families. Earlier diagnosis of NF1 allows one to offer genetic counselling to parents and relatives earlier, as well as to initiate interventions for learning or developmental problems sooner. Earlier diagnosis will become even more important once more therapeutic options are available. A direct genetic test may help to establish the diagnosis earlier, especially in sporadic patients, but only when the testing has a high sensitivity—that is, finds the mutation in (almost) all patients who will eventually fulfill the NIH criteria (low false negative results) *and*, equally important, does not confuse a benign variant with a pathogenic mutation (no false positive results).

We have identified multiple sequence changes in the paper by Mattocks *et al* that are misclassified and hence need rectification to avoid potential misdiagnoses based on the latter information.

Whereas the title of the paper states that automated comparative sequence analysis identifies *mutations* in 89% of NF1 patients, table 2 describes these sequence alterations as ''*potentially significant*'' sequence alterations. There is also a table 3 summarising *polymorphisms* found in the study, which adds to the confusion and further suggests that alterations in table 2 are pathogenic mutations, to which the title also alludes. We think we need to make a clear distinction between a deleterious mutation, an unclassified variant, a rare benign variant, and a polymorphism.

Table 2 contains four *silent* nucleotide changes: Q282Q, C680C, K1724K, and R1808R. The authors predict that these changes lead to a truncated peptide, but no experimental evidence is given to prove this. Two of these sequence changes have been observed by us and others and are definite rare benign variants. c.5172G→A (K1724K) was first described by Peters *et al*[8] as a polymorphism with allele frequencies of 0.99 for c.5172G and 0.01 for c.5172A. Peters *et al* report on a patient who carries this polymorphism in exon 28, as well as a frameshift mutation in exon 28. Fahsold *et al*[9] also describe a patient with the c.5172G→A sequence change as well as a pathogenic frameshift mutation in exon 37: c.6789del4. One of us (LM) has observed this sequence change in 2/570 NF1 patients in whom a clearly pathogenic mutation—that is, c.3216delC and c.1756_1759delACTA—was found as well. Thus this sequence variant represents an infrequent benign variant that should not be confused with a bona fide pathogenic mutation. Mattocks *et al* demonstrated that the silent nucleotide alteration K1724K was also found in the affected mother and thus segregated with the disease. However, this example clearly shows that segregation with the disease is not sufficient to provide final evidence of whether a variation is pathogenic. An effect on splicing must be shown before a silent change can be classified as pathogenic.

A second silent sequence change Q282Q (c.846G→A) predicted by Mattocks *et al* to result in a truncated protein has been reported by Luca *et al*[10] to be a polymorphism with a frequency of 2%. One of us (LM) also found Q282Q in two of 190 control samples and hence this alteration has also to be considered as a rare benign variant. By protein truncation testing and direct cDNA sequencing as described previously,[2] we did not observe the production of a truncated peptide nor any effect on splicing owing to Q282Q.

In the light of these obvious misclassifications, the prediction of the truncating effect of the two other silent sequence changes (c.2040C→T C680C; c.5427G→A R1808R) needs to be considered with great caution, especially as these changes fulfil none of the classic criteria for pathogenic mutations: neither of these silent changes has been reported previously, they have not been demonstrated to occur de novo in sporadic patients, neither were they shown to segregate with the disease in a given family, and most importantly they were not proven to affect splicing. Hence, these changes cannot be considered pathogenic unless data are produced showing that these changes have an effect on the correct splicing of the *NF1* gene. Through the study of over 600 unrelated NF1 patients fulfilling NIH criteria, we identified 29 patients carrying a pathogenic truncating mutation as well as a silent mutation, the latter without an observed effect on splicing (Messiaen *et al*, unpublished results).

Apart from the silent sequence changes, the classification of *NF1* missense mutations is also particularly challenging. Table 2 contains at least one missense alteration (D176E) which was reported previously to be a polymorphism.[9] One of us (LM) also identified D176E in one NF1 patient carrying another clearly pathogenic alteration and in one of 190 normal control samples, confirming that D176E is indeed a rare benign variant. We do not understand why the authors list a patient carrying this sequence change in table 2, as in their table 3 they state that this alteration is a polymorphism also found in unaffected individuals. Similarly, they list a patient carrying the missense alteration c.2617C→T (R873C) in table 2, while at the same time state in table 3 that this also is a benign variant they found in a patient who carried a clearly pathogenic mutation c.1-14_7del21bp.

Nevertheless, the authors mention *both* alterations in table 2 and these data, as well as the formerly mentioned misclassified alterations, are taken into account to come to the conclusion that the technique has an 89% detection rate.

Y489C (c.1466A→G) is one of the most frequent recurrent mutations in NF1 patients and was the first well understood *splice* mutation that could be misclassified as a missense mutation if only genomic DNA was studied.[11] This mutation results in the creation of a perfect novel splice donor that is used by the splicing machinery instead of the wild-type exon 10b donor, and leads to skipping of the last 62 nucleotides of exon 10b. Y489C has since been reported in many papers on *NF1* mutations.[9 12–14] Hence we do not understand why Mattocks *et al* describe this mutation, which they found in three NF1 patients, as a missense mutation that had not previously been reported.

G629R (c.1885G→A) cannot be considered a purely missense mutation either: Ars *et al* reported a splice effect in five patients.[3] We observed this splice effect in three unrelated

patients as well (Messiaen LM *et al*, unpublished results): the observed splicing error is readily understood by the creation of a novel splice acceptor site by this mutation, leading to skipping of the first 41 bp of exon 12b.

Recalculating the number of *putative* missense mutations after subtraction of the above mentioned misclassifications, table 2 of the Mattocks *et al* paper still contains 12 different missense mutations or small deletions of one or two amino acids that affect 15 patients fulfilling the NIH diagnostic criteria—that is, H31R, L145P, E337V, C324R, L532P, S574R, L844P, R1276G, R1276Q, ΔE1438, ΔIY1658-9, and ΔNF2366-7. Hence as many as 16.5% of the patients fulfilling the NIH diagnostic criteria (15 of 91) harbour putative missense mutations and small deletions of one or more amino acids. This number appears quite high in comparison with previous reports with high mutation detection rates.[2 3] Thus it may very well be that a portion of the novel missense mutations found in the study by Mattocks *et al* are splicing mutations. The authors are aware of this possibility and have developed tools such as a functional splicing assay using a minigene system to test for the effect on splicing. We do not understand why they did not *apply* these tools in this study to achieve a conclusion on the effect of silent and missense mutations. Furthermore, some of the novel missense mutations—such as D176E, R873C, and A2058D—may turn out to be non-pathogenic rare sequence variants. Their finding of different missense mutations (R873C and A2058D) in two NF1 patients carrying another clearly pathogenic mutation further underscores this possibility.

In the absence of functional assays, rigorous criteria must be applied before a novel missense alteration in the *NF1* gene can be classified as the disease causing mutation in order to avoid diagnostic errors.

The following criteria are proposed and applied when clinical testing is offered:

- Absence of any other possible deleterious mutation after analysis of the whole coding region. Analysis must include screening for a total gene deletion, smaller deletions (one to multiple exons deletions), and splice mutations including deep intronic mutations affecting splicing. This is not achieved when only genomic DNA is studied as described in the paper by Mattocks *et al*. If RNA based mutation analysis reveals an effect on splicing, the ''missense'' mutation can be considered deleterious.
- Absence of the sequence alteration in a large number of unrelated control samples. This is a necessary but insufficient criterion. Indeed, we still find novel benign variants on the wild-type *NF1* gene, inherited from the unaffected parent, even after analysing >600 patients.
- Support from evolutionary conservation in *Mus musculus*, *Rattus norvegicus*, *Takifugu rubripes*, and *Drosophila melanogaster* of the amino acid under consideration as well as support from algorithms such as the ones developed by Miller and Kumar.[15]
- Finally and importantly, clinical and molecular genetic assessment of the family. In case the patient is a sporadic case, the missense mutation needs to be proven to be a de novo event, and clinical evaluation of both parents needs to show absence of the disorder in either. In case the patient has a positive family history, the missense mutation needs to be proven to segregate with the disorder in the family by analysis of one affected relative.

Taken together, the two main findings summarised in the title of the paper by Mattocks *et al* do not withstand a critical review of the data provided. We believe it is not justified to draw conclusions on detection rates of the assay presented here unless the pathogenity of the novel silent *and* missense alterations contained in table 2 have been proved in some way. After pointing out a number of obvious misclassifications in the list, it is fair to state that the detection rate is certainly lower than calculated by the authors.

Furthermore, owing to these misclassifications and the lack of evidence that a fraction of the remaining missense alterations do not affect splicing it is not justified to use these data to confirm or refute whether a mutation cluster in exons 11–17 points to the existence of a novel functional domain.

We disagree that this single technique as applied here is appropriate for clinical practice and advocate comprehensive analysis of the complete coding region before a missense or silent alteration is considered to be a pathogenic alteration. In the absence of any other possible pathogenic alteration, missense mutations must be evaluated according to the criteria discussed above. Silent mutations cannot be considered pathogenic unless *proof* is provided that they are altering function—for example, correct splicing. We are hopeful and optimistic that mutation detection in the *NF1* gene will become more easily accomplished, faster, and cheaper once more reliable data become available and technologies develop further. Then, direct cycle sequencing or resequencing arrays will allow detection in a fast and efficient manner of a fair number of pathogenic lesions. However, genomic DNA sequencing assays alone will not allow one to decide on the pathogenicity of missense or silent alterations not previously described.

This reply is meant to stimulate vigilance in the community to avoid potential diagnostic errors.

L M Messiaen
Laboratory of Medical Genomics, 1530 3rd Avenue, S Kaul Building, Birmingham, Alabama 35249, USA

K Wimmer
Department of Human Genetics, Clinical Institute for Medical and Chemical Laboratory Diagnostics, Medical University of Vienna, Austria

Correspondence to: Professor Ludwine M Messiaen; lmessiaen@genetics.uab.edu

Competing interests: none declared

References

1 **Mattocks C**, Baralle D, Tarpey P, ffrench-Constant C, Bobrow M, Whittaker J. Automated comparative sequence analysis identifies mutations in 89% of NF1 patients and confirms a mutation cluster in exons 11–17 distinct from the GAP related domain. *J Med Genet*, 2004 Apr, **41**:e48.
2 **Messiaen LM**, Callens T, Mortier G, Beysen D, Vandenbroucke I, Van Roy N, Speleman F, Paepe AD. Exhaustive mutation analysis of the NF1 gene allows identification of 95% of mutations and reveals a high frequency of unusual splicing defects. *Hum Mutat* 2000;**15**:541–55.
3 **Ars E**, Serra E, Garcia J, Kruyer H, Gaona A, Lazaro C, Estivill X. Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum Mol Genet* 2000;**9**:237–47.
4 **Perrin G**, Morris MA, Antonarakis SE, Boltshauser E, Hutter P. Two novel mutations affecting mRNA splicing of the neurofibromatosis type 1 (NF1) gene. *Human Mutation* 1996;**7**:172–5.
5 **Zatkova A**, Messiaen L, Vandenbroucke I, Wieser R, Fonatsch C, Krainer AR, Wimmer K. Disruption of exonic splicing enhancer elements is the principal cause of exon skipping associated with seven nonsense or missense alleles of NF1. *Hum Mutat* 2004;**24**:491–501.
6 **Lambert J**, Naeyaert JM, De Paepe A, Van Coster R, Ferster A, Song M, Messiaen L. Arg-Cys substitution at codon 1246 of the human myosin Va gene is not associated with Griscelli syndrome. *J Invest Dermatol* 2000;**114**:731–3.
7 **DeBella K**, Szudek J, Friedman JM. Use of the National Institutes of Health criteria for diagnosis of neurofibromatosis 1 in children. *Pediatrics* 2000;**105**:608–14.
8 **Peters H**, Luder A, Harder A, Schuelke M, Tinschert S. Mutation screening of neurofibromatosis type 1 (NF1) exons 28 and 29 with single strand conformation polymorphism (SSCP): five novel mutations, one recurrent transition and two polymorphisms in a panel of 118 unrelated NF1 patients. Mutations in brief No 229, online. *Hum Mutat* 1999;**13**:258.
9 **Fahsold R**, Hoffmeyer S, Mischung C, Gille C, Ehlers C, Kucukceylan N, Abdel-Nour M, Gewies A, Peters H, Kaufmann D, Buske A, Tinschert S, Nurnberg P. Minor lesion mutational spectrum of the entire NF1 gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. *Am J Hum Genet* 2000;**66**:790–818.
10 **De Luca A**, Buccino A, Gianni D, Mangino M, Giustini S, Richetta A, Divona L, Calvieri S, Mingarelli R, Dallapiccola B. NF1 gene analysis based on DHPLC. *Hum Mutat* 2003;**21**:171–2.
11 **Messiaen L**, Callens T, Roux K, Mortier G, de Paepe A, Abramowicz M, Pericak-Vance M, Vance J, Wallace MR. Exon 10b of the NF1 gene represents a mutational hotspot and harbors a recurrent missense mutation Y489C associated with aberrant splicing. *Genet Med* 2004;**1**:248–53.
12 **Ars E**, Kruyer H, Morell M, Pros E, Serra E, Ravella A, Estivill X, Lazaro C. Recurrent mutations in the NF1 gene are common among neurofibromatosis type 1 patients. *J Med Genet* 2003;**40**:e82.
13 **Han SS**, Cooper DN, Upadhyaya MN. Evaluation of denaturing high performance liquid chromatography (DHPLC) for the mutational analysis of the neurofibromatosis type 1 (NF1) gene. *Hum Genet* 2001;**109**:487–97.
14 **Osborn MJ**, Upadhyaya M. Evaluation of the protein truncation test and mutation detection in the NF1 gene: mutational analysis of 15 known and 40 unknown mutations. *Hum Genet* 1999;**105**:327–32.
15 **Miller MP**, Kumar S. Understanding human disease mutations through the use of interspecific genetic variation. *Hum Mol Genet* 2001;**10**:2319–28.