Local sequence determinants of two in-frame triplet deletion/duplication hotspots in the RHD/RHCE genes

Jian-Min Chen, David Cooper, Claude Férec

Human Genomics

LETTER TO THE EDITOR Open Access

Local sequence determinants of two in-frame triplet deletion/duplication hotspots in the *RHD/RHCE* genes

Jian-Min Chen[1*], David N Cooper[2] and Claude Férec[1]

Different types of human gene mutation can vary in size quite dramatically (e.g., single nucleotide substitutions vs. copy number variations), but what they all have in common is that their occurrence is often closely bound up with specific characteristics of the local DNA sequence environment [1]. Here, we highlight the importance of local sequence features that underlie the two in-frame triplet deletion/duplication hotspots in the *cis*-linked, highly homologous *RHD* and *RHCE* paralogs.

The first hotspot refers to an 8-bp sequence tract in exon 1 of the *RHD* and *RHCE* genes, in which three different variants were reported (Figure 1a) [2-4]. The first variant is a deletion of one of two juxtaposed CTC triplets in the *RHD* gene, which gives rise to an in-frame deletion of a single amino acid, Leu27 [2]. The second variant is identical to the first but occurred at the analogous location in the *RHCE* gene [3]. Henceforth, we shall employ the term 'deduplication' [5], which emphasizes the identity of the deleted sequence and the sequence immediately abutting the site of the deletion, to describe this particular type of microdeletion (<21 bp in length in accordance with Ball et al. [6]). Deduplication accounts for a significant proportion of disease-causing microdeletions; indeed, for microdeletion events of 2–5 bp, 38 % were found to be deduplications [6]. Replication slippage is currently regarded as the major mechanism underlying the generation of deduplications: the primer strand containing the newly synthesized first direct repeat dissociates from the template strand and then misaligns (slipping forward) at the second direct repeat; continued DNA synthesis then leads to the deletion of one of the two direct repeats. It should be pointed out that while direct repeats are a prerequisite for replication slippage, they are certainly not the sole determinant of this mutational mechanism. In this regard, we noted that the two CTC triplets together constitute a DNA polymerase arrest site (consensus sequence WGGAG, where W = A or T [7]) (Figure 1a). It is therefore tempting to speculate that a combination of these two sequence features could have served to make this short region particularly prone to replication slippage.

Recently, Pereira and colleagues reported the first in-frame triplet duplication in the *RHD* gene; this duplication affected the same short region as the aforementioned two deduplications in exon 1 (Figure 1a) [4]. As pointed out by the original authors, this duplication could have resulted from either a duplication of c.74-76TTC or c.75_77TCT. These authors emphasized the importance of a DNA motif (i.e., TTCTC that was identified by analogy to previously reported deletion-predisposing DNA motifs in the *RHD* gene [9]) in generating this duplication but did not provide a model to explain how this duplication could have been generated. Given that the sequence tract in question is prone to replication slippage, we surmised that this duplication might also be explicable in terms of such a mechanism. Indeed, as illustrated in Figure 1b, it can be readily explained by the model of serial replication slippage [8], invoking one step of forward slippage and one step of backward slippage.

The second hotspot refers to a 63-bp region of exon 5 in the *RHD* and *RHCE* genes, in which four in-frame triplet deletions (c.644_646delTCT [3], c.684_686delGAG, and c.705_707delAGA [9] in *RHD*; c.685_687delAGA [10] in *RHCE*) were reported (Figure 2). Several distinct DNA repeats or motifs (e.g., GAGAA and GAAGA) have previously been implicated in the generation of three of these four variants [9]. A comparative evaluation of the four variants led us to propose a consensus motif RAGAA (R = A or G) (Figure 2). Since only the c.644_646delTCT variant can be explained in terms of replication slippage, it may be that RAGAA is associated with a recombination-

* Correspondence: Jian-Min.Chen@univ-brest.fr
[1]Etablissement Français du Sang (EFS) – Bretagne and INSERM, U1078, Brest, France
Full list of author information is available at the end of the article

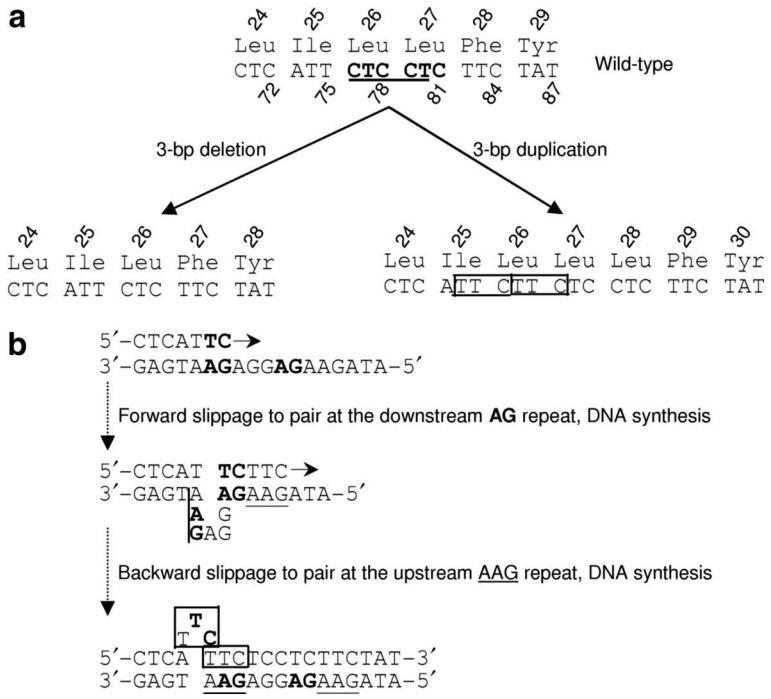

Figure 1 The three currently known in-frame triplet deletion/duplication variants in exon 1 of the *RHD*/*RHCE* genes. Their underlying mutational mechanisms are also shown. (**a**) Illustration of the identical deletion of one of two neighboring CTC repeats (*in bold*) which occurred at the corresponding positions of *RHD* [2] and *RHCE* [3] genes as well as the 3-bp in-frame duplication in the *RHD* gene [4], illustrated here as having arisen from the duplication of c.74_76 (*boxed*). The *RHD* and *RHCE* genes share 100% sequence identity in the region from c.70_87. The *underlined* CTCCT motif corresponds to the complementary strand of the DNA polymerase arrest site WGGAG. (**b**) Schema for how the 3-bp duplication could have been generated in accordance with the model of serial replication slippage (see Chen et al. [8] for details). Short direct repeats that could have mediated the two steps of replication slippage are highlighted in *bold* or are *underlined*. The *horizontal arrows* indicate the direction of DNA synthesis. The identical 3-bp deletions are explicable by a single step of forward slippage (not shown).

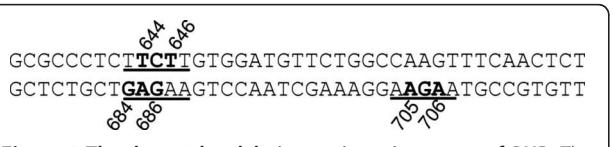

Figure 2 The three 3-bp deletion variants in exon 5 of *RHD*. The deletions are highlighted in *bold*. The *underlined* sequences refer to the consensus motif RAGAA or its complement. The c.685_687AGA deletion in RHCE [10] is not shown.

predisposing activity that is distinct from the DNA polymerase arrest site WGGAG. In other words, the different local sequence contexts in exons 1 and 5 of *RHD* and *RHCE* could predispose to subtly different mutational processes.

Author details
[1]Etablissement Français du Sang (EFS) – Bretagne and INSERM, U1078, Brest, France. [2]Institute of Medical Genetics, Cardiff University, Cardiff CF14 4XN, UK.

Received: 10 May 2012 Accepted: 14 May 2012
Published: 2 August 2012

References

1. Cooper DN, Bacolla A, Férec C, Vasquez KM, Kehrer-Sawatzki H, Chen JM: **On the sequence-directed nature of human gene mutation: the role of genomic architecture and the local DNA sequence environment in mediating gene mutations underlying human inherited disease.** *Hum Mutat* 2012, **32:**1075–1099.
2. Tilley L, Rodrigues MJ, Malde R, Marais I: **Two novel RhD variants caused by triplet deletions in the *RHD* gene.** *Transfus Med* 2007, **17:**32.
3. Döscher A, Vogt C, Bittner R, Gerdes I, Petershofen EK, Wagner FF: ***RHCE* alleles detected after weak and/or discrepant results in automated Rh blood grouping of blood donors in Northern Germany.** *Transfusion* 2009, **49:**1803–1811.
4. Pereira JC, Rodrigues MJ, Tilley L, Poole J, Chabert T, Ribeiro ML: **RhD variant caused by an in-frame triplet duplication in the *RHD* gene.** *Transfusion* 2011, **51:**570–573.
5. Kondrashov AS, Rogozin IB: **Context of deletions and insertions in human coding sequences.** *Hum Mutat* 2004, **23:**177–185.
6. Ball EV, Stenson PD, Abeysinghe SS, Krawczak M, Cooper DN, Chuzhanova NA: **Microdeletions and microinsertions causing human genetic disease: common mechanisms of mutagenesis and the role of local DNA sequence complexity.** *Hum Mutat* 2005, **26:**205–213.
7. Chuzhanova N, Chen JM, Bacolla A, Patrinos GP, Férec C, Wells RD, Cooper DN: **Gene conversion causing human inherited disease: evidence for involvement of non-B-DNA-forming sequences and recombination-promoting motifs in DNA breakage and repair.** *Hum Mutat* 2009, **30:**1189–1198.
8. Chen JM, Chuzhanova N, Stenson PD, Férec C, Cooper DN: **Complex gene rearrangements caused by serial replication slippage.** *Hum Mutat* 2005, **26:**125–134.
9. Flegel WA, Eicher NI, Doescher A, Hustinx H, Gowland P, Mansouri Taleghani B, Petershofen EK, Bauerfeind U, Ernst M, von Zabern I, Schrezenmeier H, Wagner FF: **In-frame triplet deletions in *RHD* alter the D antigen phenotype.** *Transfusion* 2006, **46:**2156–2161.

10. Chen YX, Peng J, Novaretti M, Reid ME, Huang CH: **Deletion of arginine codon 229 in the Rhce gene alters e and f but not c antigen expression.** *Transfusion* 2004, **44:**391–398.

doi:10.1186/1479-7364-6-8
Cite this article as: Chen *et al.*: Local sequence determinants of two in-frame triplet deletion/duplication hotspots in the *RHD/RHCE* genes. *Human Genomics* 2012 **6**:8.