Identification of novel mutations in hemochromatosis genes by targeted next generation sequencing in Italian patients with unexplained iron overload

Sadaf Badar, Fabiana Busti, Alberto Ferrarini, Luciano Xumerle, Paolo Bozzini, Paola Capelli, Roberto Pozzi-Mucelli, Natascia Camponistrini, Giovanna De Matteis, Sergio Marin Vargas, Alejandro Giorgetti, Massimo Delledonne, Oliviero Olivieri, and Domenico Girelli.

Hereditary hemochromatosis (HH), one of the commonest genetic disorders in Caucasians, is mainly associated to homozygosity for the C282Y mutation in the HFE gene, which is highly prevalent (allele frequency up to near 10% in Northern Europe) and easily detectable through a widely available “first level” molecular test. However, in certain geographical regions like the Mediterranean area, up to 30% of patients with a HH phenotype have a negative or non-diagnostic (i.e. simple heterozygosity) test, because of a known heterogeneity involving at least four other genes (HAMP, HJV, TFR2, and SLC40A1). Mutations in such genes are generally rare/private, making the diagnosis of atypical HH essentially a matter of exclusion in clinical practice (from here the term of “non-HFE” HH), unless cumbersome traditional sequencing is applied. We developed a Next Generation Sequencing (NGS)-based test targeting the five HH genes, and applied it to patients with clinically relevant iron overload (IO) and a non-diagnostic first level genetic test. We identified several mutations, some of which were novel (i.e. HFE W163X, HAMP R59X, and TFR2 D555N) and allowed molecular reclassification of “non-HFE” HH clinical diagnosis, particularly in some highly selected IO patients without concurring acquired risk factors. This NGS-based “second level” genetic test may represent a useful tool for molecular diagnosis of HH in patients in whom HH phenotype remains unexplained after the search of common HFE mutations.


Introduction

Hereditary hemochromatosis (HH) is a potentially lethal group of genetic disorders due to insufficient production or response to the iron regulatory hormone hepcidin, leading to progressive accumulation of toxic amounts of iron in vital organs like liver, heart, pancreas, and endocrine glands [1]. Early diagnosis of HH is mandatory, since a prompt treatment before the development of end-stage organ damage can partially or totally reverse the clinical phenotype, and can be associated to a near normal life expectancy [2]. HH is among the commonest genetic diseases in Caucasian populations, being typically associated with homozygosity for the C282Y mutation in the HFE gene, which takes part in complex hepcidin regulation (for comprehensive reviews see [3,4]). This mutation, indeed, has an allele frequency up to near 10% in certain Northern European countries, like in Ireland [5]. On the other hand, only a minor fraction of C282Y homozygotes develop the disease, whose clinical penetrance is remarkably influenced by acquired and genetic risk factors. The main acquired cofactors, that acts by further suppressing hepcidin, are represented by alcohol abuse [6] and HCV infection [7]. The copresence of common features of the so-called metabolic syndrome [8] is also relevant in clinical practice, since they can contribute to further increase of serum ferritin levels, and are sometimes associated to a true iron overload [9,10]. The genetic modifiers remain still largely unknown, although recent studies have suggested a role for GNPAT [11], not confirmed by others [12,13]. Beyond the C282Y, another common HFE polymorphism, the H63D, can contribute to the development of biochemical and, sometimes, clinically relevant signs of iron overload (IO) when in compound heterozygosity with the C282Y mutation, particularly in subjects with one or more of the above mentioned cofactors [5]. A simple and widely available “first level” genetic test evaluating the presence of the C282Y and H63D variants is therefore a useful tool for clinicians facing with patients with suspected HH, although it should always be prescribed according to diagnostic needs.
algorithms included in guidelines [14,15]. On the other hand, soon after the discovery of the HFE gene in 1996 [16], it has become clear that HH is a genetically heterogeneous disorder, particularly in the Mediterranean area where up to near 30% of patients have a negative or “non-diagnostic” (i.e., C282Y and H63D simple heterozygosity) first level genetic test [17]. Rare mutations in the HFE gene itself, and/or in additional four genes encoding hepcidin (HAMP) [18], its receptor ferroportin (SLC40A1) [19,20], or critical modulators of the iron-sensing machinery controlling hepcidin like type 2 transferrin receptor (TFR2) [21,22], and hemochromatins (HJV) [23,24], have also been associated to HH. Such conditions are often collectively designated as “non-HFE” HH [25,26]. Of note, mutations in the above-mentioned genes are typically “private,” i.e., limited to just one family and/or few individuals, requiring DNA sequencing for establishing a possible molecular diagnosis. In clinical practice, “non-HFE” HH is still essentially a diagnosis of exclusion, which is made after a negative first level genetic test in subjects with clinical, biochemical, instrumental, and pathological findings suggesting primary iron overload. Only relatively few cases have been molecularly characterized at referral centers through traditional sequencing, which was hampered by time-consuming and costs, until recently. Recent breakthrough in the development of novel genomic technologies for high-throughput DNA sequencing, collectively known as “next generation sequencing” (NGS), now have the potential of revolutionizing the molecular diagnosis of inherited diseases [27]. This holds the promise of a significant impact on clinical care [28], with unprecedented accuracy, rapidity, and at constantly declining costs [29]. Here we describe the development of a NGS-based “second level” genetic test that simultaneously analyzes all the five hemochromatosis genes, whose performance was evaluated in patients attending our referral clinic because of the presence of relevant biochemical signs of HH and nondiagnostic first level genetic test, suggesting a possible atypical HH.

Methods

Patients. The newly developed second level, NGS-based, genetic test for HH was applied to a total of 106 subjects. For this study we selected DNA samples from patients attending our Regional Referral Centre for Iron Disorders at the Azienda Ospedaliera Universitaria Integrata of Verona (http://www.gimferverona.org). All patients (n = 34; mean age 51.3 years, CIs: 47.5–55.2 years; 82.3% males) had relevant biochemical signs of iron overload (i.e., serum ferritin levels between 30 and 200 or 300, C282Y and H63D simple heterozygosity) [19,20], or critical modulators of the iron-sensing machinery controlling hepcidin like type 2 transferrin receptor (TFR2) [21,22], and hemochromatins (HJV) [23,24], have also been associated to HH. Such conditions are often collectively designated as “non-HFE” HH [25,26]. Of note, mutations in the above-mentioned genes are typically “private,” i.e., limited to just one family and/or few individuals, requiring DNA sequencing for establishing a possible molecular diagnosis. In clinical practice, “non-HFE” HH is still essentially a diagnosis of exclusion, which is made after a negative first level genetic test in subjects with clinical, biochemical, instrumental, and pathological findings suggesting primary iron overload. Only relatively few cases have been molecularly characterized at referral centers through traditional sequencing, which was hampered by time-consuming and costs, until recently. Recent breakthrough in the development of novel genomic technologies for high-throughput DNA sequencing, collectively known as “next generation sequencing” (NGS), now have the potential of revolutionizing the molecular diagnosis of inherited diseases [27]. This holds the promise of a significant impact on clinical care [28], with unprecedented accuracy, rapidity, and at constantly declining costs [29]. Here we describe the development of a NGS-based “second level” genetic test that simultaneously analyzes all the five hemochromatosis genes, whose performance was evaluated in patients attending our referral clinic because of the presence of relevant biochemical signs of HH and nondiagnostic first level genetic test, suggesting a possible atypical HH.

Target capture and sequencing. Capture of exonic regions of the 5 HH genes (HFE, HFE2, TFR2, HAMP, and SLC40A1) was performed starting from 225 ng of genomic DNA using a custom design HaloPlex® Target Enrichment 1–500 kb Kit (Agilent, Santa Clara, CA) for 55–72 genetic variations. Of note, this approach has recently been recommended for high-coverage capturing of regions spanning 100–500 kb, particularly when relatively large batches have to be examined [31]. Library quality was determined using the Agilent High Sensitivity DNA kit on the Agilent 2100 bioanalyzer, and the quantity was measured using Qubit dsDNA assay kit with Qubit fluorimeter (Life Technologies, Carlsbad, CA) and by quantitative PCR using the KAPA Library Quantification kit (Kapa Biosystems). Libraries were pooled at equimolar concentrations and sequenced with TrueSeq PE Cluster Kit v3 and TrueSeq SBS Kit v3 (illumina, san Diego, CA) on illumina HiSeq 1000 sequencer (illumina, san Diego, CA) generating 100-bp paired-end reads at average target coverage of 1000X for captured regions.

Data analysis. Illumina adapter sequences were removed from sequence reads using Sickle software using default parameters (v. 0.994; https://github.com/vsbuffalo/sickle software). Reads were mapped against human genome version 19 (hg19) using BWA (Burrows Wheeler Aligner, version 0.7.10). Mappings and portion of mappings outside expected amplicons were removed using custom scripts. Duplicate PCR reads were removed using Picard, base quality scores were recalibrated using Genome Analysis Toolkit (GATK, version 3), and alignments near putative insertion/deletions were refined with gatk. Variants were called using Unified Genotyper module from gatk. Variant calls that failed to pass the following filters were eliminated from the call set: (i) MQ ≥ 48 & (MQS >0.5), (ii) DP > 5, (iii) QUAL > 30.0, (iv) QD < 2.0. Coding and splice variants were annotated using the SNP & Variation Suite software (version 8; golden Helix, Bozeman, MT). The potential functional impact of identified variants was assessed using three different predictive algorithms, i.e. Sorting Intolerant From Tolerant (sift), polymorphism Phenotyping version 2 (PolyPhen2 HumaVar), and MutationTaster (MT). The variants of interest were confirmed by traditional Sanger sequencing using a Cq G4000 capillary sequencer and Genome Lab DTCG Quick Start kit (beckman Coulter, Indianapolis, IN). When possible, some of the genetic variants that appeared most interesting were further studied by in silico modeling, as reported [32]. Methodological details and ad hoc references are reported in Supporting Information [53–62].

Results

Genetic variants and literature review

Supporting Information Table SI summarizes the variants found through NGS that survived rigorous filtering, in addition to the known HH associated HFE mutations (C282Y, H63D, S65C), and the known mutations previously found in methodological controls. For sake of completeness, we report also the variants found only in controls, which could be useful for comparison with future reports and for databases compilation [33,34]. Among the listed variants, five were apparently new, i.e. not reported in the following databases:

dbSNPs (http://www.ncbi.nlm.nih.gov/SNP/),
dbSNPs (http://www.ncbi.nlm.nih.gov/SNP/), 1,000 genome (http://www.1000genomes.org/category/sequence),
OMIM (http://www.ncbi.nlm.nih.gov/omim), and ESP Databases (http://evs.gs.washington.edu/evs/). We also checked a very recent paper reporting a catalogue of all hemochromatosis associated variants [35]. Two out of the new mutations, i.e., the H274fs in SLC40A1 and M45fs in TFR2 were found in control subjects. The SLC40A1 H274fs was found in a 43-year-old man with serum ferritin of 199 µg L−1. Although no prediction could be inferred by the three predictive algorithms (Supporting Information Table SI), given that ferroportin disease is the only dominant non-HFE HH subtype, this variant is likely non pathogenic. The same applies to the TFR2 M54fs, i.e., no useful information from the predictive algorithms, but since type 3 HH is autosomal recessive we cannot draw any conclusion about its possible pathogenetic role.

The three other novel private variants found only in patients from group 1, which could be considered very likely pathogenic, are described in detail below along with individuals’ clinical data (Table I).

A sixth patient from group 1 with history of consanguinity was found homozygous for the TFR2 N2411 mutation. This variant has
TABLE I. Clinical and Biochemical Features of Group 1 Patients With Iron Overload, Minimal or No Acquired Cofactors, and Non-diagnostic 1st Level Genetic Test

<table>
<thead>
<tr>
<th>Pt. ID</th>
<th>Age/sex</th>
<th>Ist level genetic test</th>
<th>TS, Ferritin</th>
<th>LIC&lt;sup&gt;a&lt;/sup&gt; (MR)</th>
<th>Liver biopsy</th>
<th>NGS</th>
<th>Reference</th>
<th>Clinical-molecular interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>#01</td>
<td>70/M</td>
<td>Wild-type&lt;sup&gt;b&lt;/sup&gt;</td>
<td>80%, 1,450 µg l&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>–</td>
<td>no (HH “pre-cirrhotic”)</td>
<td>–</td>
<td>–</td>
<td>Remains unexplained</td>
</tr>
<tr>
<td>#02</td>
<td>39/M</td>
<td>H63D+/−</td>
<td>61%, &gt;1,000 µg l&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>–</td>
<td>yes (HH “pre-cirrhotic”)</td>
<td>–</td>
<td>Novel</td>
<td>Atypical Type 1 HH (HFE-related)</td>
</tr>
<tr>
<td>#03</td>
<td>48/M</td>
<td>H63D+/−</td>
<td>73%, &gt;1,000 µg l&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>300</td>
<td>No</td>
<td>–</td>
<td>Novel</td>
<td>Digenic HH (HFE/HAMP)?</td>
</tr>
<tr>
<td>#04</td>
<td>62/F</td>
<td>Wild-type&lt;sup&gt;b&lt;/sup&gt;</td>
<td>60%, 1,786 µg l&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>270</td>
<td>yes (siderosis 3+, perportal gradient)</td>
<td>–</td>
<td>–</td>
<td>Remains unexplained</td>
</tr>
<tr>
<td>#05</td>
<td>58/M</td>
<td>H63D+/−</td>
<td>n.a., 1,089 µg l&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>160</td>
<td>No</td>
<td>–</td>
<td>Novel</td>
<td>Digenic HH (HFE/TFR2)?</td>
</tr>
<tr>
<td>#06</td>
<td>35/F</td>
<td>Wild-type&lt;sup&gt;b&lt;/sup&gt;</td>
<td>100%, n.a.</td>
<td>–</td>
<td>yes (HH “pre-cirrhotic”)</td>
<td>–</td>
<td>–</td>
<td>Remains unexplained</td>
</tr>
<tr>
<td>#07</td>
<td>66/M</td>
<td>Wild-type&lt;sup&gt;b&lt;/sup&gt;</td>
<td>48%, 2,352 µg l&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>250</td>
<td>yes (siderosis 2–3+, perportal gradient)</td>
<td>–</td>
<td>–</td>
<td>Type 3 HH</td>
</tr>
<tr>
<td>#08&lt;sup&gt;c&lt;/sup&gt;</td>
<td>47/M</td>
<td>H63D+/−</td>
<td>95%, 6,242 µg l&lt;sup&gt;-1&lt;/sup&gt;</td>
<td>–</td>
<td>no (iron removed &gt;10 g)</td>
<td>SLC40A1</td>
<td>–</td>
<td>Type 4 HH</td>
</tr>
</tbody>
</table>

<sup>a</sup> µM g<sup>−1</sup>.<br>
<sup>b</sup> = for both C282Y and H63D HFE variants.<br>
<sup>c</sup> This peculiar patient has been described in detail elsewhere (Ferbo et al., submitted).

Figure 1. Functional consequences of the HAMP R59X mutation. The sequence of hepcidin pro-pro-protein is shown (UniProtKB entry P81172). Arginine 59 is located upstream of the region coding the mature hepcidin-25 peptide, which derives from subsequent cleavage of a precursor (pre-pro-hepcidin). This stop codon mutation prevent the possibility of producing the mature peptide. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

been previously reported as pathogenic by others [36], since it was found in a young male whose HH was clinically diagnosed at the age of 17 years. Our data are also consistent with a highly penetrant mutation, since our patient was a female whose HH was clinically diagnosed during the fertile period (at the age of 35 years) because of sustained increase of transaminases, 100 percent transferrin saturation, and a liver biopsy showing iron overload and advanced fibrosis (Table I, patient ID = #06). In this patient we also found a SLC40A1 variant (rs11568344, Supporting Information Table SI), which was already reported as a benign polymorphism by others [37].

The TFR2 rs41295942 variant (resulting in mutation R581H) was previously described by early reports using traditional sequencing, with contrasting or biased results [38–40]. It was also specifically investigated as a potential modifier of the C282Y genotype in a recent report [41], but no effect was found. In our study, this variant was found in homozygosity in a control subject (a 47-year-old male with serum ferritin of 167 µg L<sup>−1</sup>). This finding appears to definitively rule out a possible pathogenic role of this TFR2 variant. The HJV A197G variant (rs7540883) that we found in heterozygosity in a control subject was also previously reported by others [37], but without a definite conclusion on its possible pathogenetic role. It was only “possibly damaging” according PolyPhen2, and benign according to SIFT and MT, which seems more likely. Other variants already annotated/reported [42,43] and classified as benign/polymorphisms are reported in Supporting Information Table SI.

Novel private mutations with high probability of being pathogenic

Three interesting novel variants were found in patients from group 1 (Table I). Patient #03 carried a stop-codon mutation (R59X) in the HAMP gene. Of note, the Arginine at position 59 is located immediately upstream the sequence containing the hepcidin-25 mature peptide (Fig. 1). Thus, this mutation results in the synthesis of a pre-pro-hepcidin that cannot be processed into the functional peptide. This patient was diagnosed at the age of 48 years because of altered iron parameters (serum ferritin >1,000 µg L<sup>−1</sup>, transferrin saturation >70%) and substantial iron accumulation at MRI, but he was clinically asymptomatic and refused liver biopsy (Table I). At the first level genetic test he was only heterozygous for the HFE H63D variant. Thus, after NGS the patient could be reclassified as an example of HH due to digenic inheritance, i.e. the presence of two or more variants in distinct genes involved in the same pathway (hepcidin sensing/production) sufficient to give rise to the clinical disease [44].

Patient #02 was also a simple H63D heterozygous at first level genetic test, which clearly did not explain a highly penetrant HH phenotype with advanced fibrosis at liver biopsy performed at the age of 39 years (Table I). NGS revealed a new stop-codon mutation in the HFE gene (W163X), whose modeling is reported in Fig. 2. According to in silico analysis, the mutation determines a dramatic change of protein structure, with the truncated protein being predicted to lack some amino acids critical for the surface interaction between HFE and transferrin receptor 1 (TFR1). It could affect also other interactions, e.g. BMP type 1 receptor ALK3). Such effects would eventually prevent the signaling that gives rise to hepcidin production. The final functional effect of this mutation can be considered similar to that of the C282Y mutation, which has been reported to prevent the correct membrane localization of HFE protein [45]. Thus, the molecular diagnosis in this patient previously classified as likely affected by “non-HFE” HH, was indeed “atypical” type 1 (or HFE-related) HH due to compound heterozygosity for a novel and a known mutation (W163X/H63D).
Finally, regarding the performance of our second level genetic test, all the mutations previously described in the “methodological” group, including HJV N196K [30], TFR2 AVAQ 594-597 deletion [22], and SLC40A1 L233P/I152F [19], were confirmed by NGS. Moreover, all the new potentially causative variants detected by NGS were confirmed by Sanger sequencing (Supporting Information Fig. S2).

## Discussion

This study, along with a very similar one in Australian patients recently published by McDonald and colleagues [40], further illustrates the usefulness of an NGS-based approach test for the molecular diagnosis in patients with a provisional diagnosis of “non-HFE” HH. Indeed, in highly selected patients attending our referral center for Iron Disorders with an unexplained HH phenotype, because of either non-diagnostic 1st level genetic test or absence of acquired risk factors for nongenetic IO, we could obtain a plausible molecular explanation in five out of eight, finding three novel variants in HFE, HAMP, and TFR2. The same test also allowed us to detect a rare SCL40A1 mutation in one patient with a distinct phenotype, who has been described separately (Ferbo L, manuscript under revision). A limitation lies on the fact that NGS was exonic, so there still could be causative mutations in the analyzed genes within non-exonic sequences. Anyhow, similarly to our results, McDonald et al. [40], using a different NGS platform and an enlarged panel of genes possibly involved in a broad range of iron disorders (n = 39), were able to identify a likely causal mutation (in HJV, SCL40A1, HFE, or TFR2, but not in HAMP) in five out of eight patients with “atypical IO.” Of note, notwithstanding the broad panel used, even in their study some patients with a HH phenotype remained molecularly “unexplained” after extensive NGS. While, of course, no NGS method, whatever the platform used, can ensure 100% sensitivity in detecting causal mutations, it is tempting to conclude that some still unknown genes could be involved in determining HH-like clinical phenotypes in humans. These genes could be eventually detected by using other NGS-based techniques like whole genome/exome sequencing [27], however their application is presently limited not only by costs, but also by difficulties in interpretation and ethical problem in reporting incidental findings [46].

Until now, the diagnosis of atypical or “non-HFE” HH is largely a diagnosis of exclusion, and our knowledge on these conditions is mainly based on individual studied by traditional Sanger sequencing. Current recommendations suggest a stepwise approach for the molecular diagnosis of atypical systemic iron overload disorders, in which sequencing should be made initially according to clinical features (i.e., prioritizing HJV and HAMP in patients with early onset IO) [5,14]. While this is reasonable, it is time consuming and does not take appropriately into account the possible occurrence of digenic inheritance [47]. This new approach of simultaneous NGS-based sequencing of the five HH genes at a reasonable cost will clearly simplify the molecular diagnosis in a near future. Moreover, it will allow to expand/refine knowledge on the clinical features of the distinct HH subtypes. For example, based on the few cases reported until now, there is some controversy regarding whether or not type 3 (TFR2-related) HH has an earlier (or intermediate) onset as compared to type 1 [36]. Similarly, it is unclear whether or not type 4 HH (SLC40A1-related or ferroportin disease) is a relatively mild form of HH with low incidence of overt cirrhosis and HCC [20]. A widespread application of NGS-based molecular diagnosis is expected to substantially increase the accuracy of HH molecular diagnosis, allowing, preferably through the creation of ad hoc registries, a more precise phenotypic description of HH subtypes.

Noteworthily, we did not find any relevant mutation in patients from group 2, i.e., those with IO possibly out of proportion with...
The study examined the effectiveness of a NGS-based approach targeting a specific gene panel for genetic testing. It highlighted the importance of accurate clinical selection before ordering any genetic test, which should be done only in subjects with the highest pretest probability. In their experience, NGS-based methods clearly represent a major technical advance, interpretation of the clinical relevance of molecular data remains a difficult task, requiring close and repeated contacts among clinicians, molecular analysts, and bioinformatic experts. Nonetheless, our study further emphasizes the usefulness of a NGS-based approach targeting a specific gene panel for the molecular diagnosis of patients clinically classified as “atypical” or non-HFE HH after first level genetic testing.

Acknowledgments
This work was performed (in part) in the LURM (Laboratorio Universitario Ricerca Medica) Research Center, University of Verona.

Author Contributions
SB performed sequencing, reviewed literature/databases, and co-wrote the manuscript. DG designed the research, reviewed the study results and genotype/phenotype correlations, and co-wrote/edit the manuscript. FB, PB, RPM, PC, OO participated in clinical enrolling/work-up of HH patients. AF, LX, GDM, NC, and MD performed the molecular/biochemical/bioinformatic analyses, and edited the manuscript. SMV and AG performed in silico modeling of mutations.

References