

Novel Diagnostic Tool for p47^{phox}-Deficient Chronic Granulomatous Disease Patient and Carrier Detection

Dominik Wrona,^{1,2,3} Ulrich Siler,^{1,2,3,4} and Janine Reichenbach^{1,2,3,4}

¹Division of Immunology, University Children's Hospital Zurich, 8032 Zurich, Switzerland; ²Children's Research Center, University Children's Hospital Zurich, 8032 Zurich, Switzerland; ³Associated Group Institute for Regenerative Medicine, University of Zurich, 8952 Schlieren-Zurich, Switzerland

Chronic granulomatous disease (CGD) is a primary immunodeficiency caused by mutations of the phagocytic nicotinamide adenine dinucleotide phosphate (NADPH) oxidase. Autosomal recessive p47^{phox}-deficient CGD (p47^{phox} CGD) is the second most frequent form of the disease in western countries, and more than 94% of patients have a disease-causing dinucleotide deletion (Δ GT) in the neutrophil cytosolic factor 1 (*NCF1*) gene. The Δ GT mutation is most likely transferred onto the *NCF1* from one of its two pseudogenes co-localized on the same chromosome. The presence of *NCF1* pseudogenes in healthy individuals makes the genetic diagnostics of Δ GT p47^{phox} CGD challenging, as it requires the distinction between Δ GT in *NCF1* and in the two pseudogenes. We have developed a diagnostic tool for the identification of p47^{phox} CGD based on PCR co-amplification of *NCF1* and its pseudogenes, followed by band intensity quantification of restriction fragment length polymorphism products. The single-day, reliable p47^{phox} CGD diagnostics allow for robust discrimination of homozygous Δ GT p47^{phox} CGD patients from heterozygous carriers and healthy individuals, as well as for monitoring gene therapy efficacy.

INTRODUCTION

Chronic granulomatous disease (CGD) is a primary immunodeficiency of phagocytes, leading to recurrent severe bacterial and fungal infections due to impaired reactive oxygen species (ROS) production by the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex.¹ Disease-causing mutations are found in all NADPH oxidase subunits gp91^{phox}, p47^{phox}, p67^{phox}, p22^{phox}, and p40^{phox}. In western countries, p47^{phox} deficiency (p47^{phox} CGD) is the second most frequent form of CGD. p47^{phox} CGD is genetically exceptional, as 97% of patients share the same mutation, a dinucleotide deletion (Δ GT) within the GTGT sequence in exon 2 of the neutrophil cytosolic factor 1 (*NCF1*) gene.² On chromosome 7, the *NCF1* gene is accompanied by two pseudogenes (*NCF1B* and *NCF1C*) with 99.5% sequence homology to *NCF1* (Figures 1A–1C).

The Δ GT mutation is also present in both pseudogenes, in healthy individuals, in CGD patients, and in carriers. *NCF1* Δ GT results from unequal cross-over events between *NCF1* and one of its pseudogenes

during DNA replication or repair, leading to partial pseudogene sequence transfer, including Δ GT, onto the *NCF1* gene.^{3–5} Genetic diagnostics of Δ GT p47^{phox} CGD are challenging, as it requires the distinction between Δ GT in *NCF1* and in the pseudogenes. Currently, diagnosis of Δ GT p47^{phox} CGD relies on the gene-scan method,⁶ which is based on the comparison of fluorescence intensities of short co-amplified sequences of *NCF1* and its pseudogenes, which differ in size by the 2-nt of the Δ GT mutation. Other diagnostic methods are allele-specific hybridization⁵ and determination of the Δ GT:GTGT ratio by the TaqMan copy number variation (CNV) assay.⁷

RESULTS AND DISCUSSION

We have developed a novel diagnostic tool for the identification of Δ GT p47^{phox} CGD based on PCR co-amplification of *NCF1* and its pseudogenes, followed by band intensity quantification of restriction fragment length polymorphism products (Figures 1B–1G). This 1-day method determines the *NCF1* gene CNV by quantification of GTGT content in the *NCF1* gene and pseudogene loci, and thus it detects the presence or absence of the Δ GT mutation within *NCF1* gene and pseudogene alleles. It can be established in any molecular biology laboratory, and it allows for the robust discrimination of homozygous Δ GT p47^{phox} CGD patients from heterozygous carriers and healthy individuals for rapid diagnostic purposes, as well as for the monitoring of *NCF1* genome-editing-based gene therapy.⁸

For quantification of the GTGT content, corresponding *NCF1*, *NCF1B*, and *NCF1C* sequences were co-amplified in one PCR reaction, digested, and visualized by PAGE or agarose electrophoresis (Figures 2A and 2B; detailed characterization of bands in polyacrylamide gel in Figures S2–S6). PCR-restriction fragment length polymorphism (RFLP) analysis was performed for 60 healthy individuals, 10 conventionally diagnosed Δ GT p47^{phox} CGD patients, and 8 conventionally diagnosed Δ GT p47^{phox} CGD carriers (see Table S1).

Received 28 August 2018; accepted 4 February 2019;
<https://doi.org/10.1016/j.omtm.2019.02.001>.

⁴These authors contributed equally to this work.

Correspondence: Janine Reichenbach, Co-Head, Division of Immunology, University Children's Hospital Zurich, Steinwiesstr. 75, 8032 Zurich, Switzerland.
E-mail: janine.reichenbach@kispi.uzh.ch



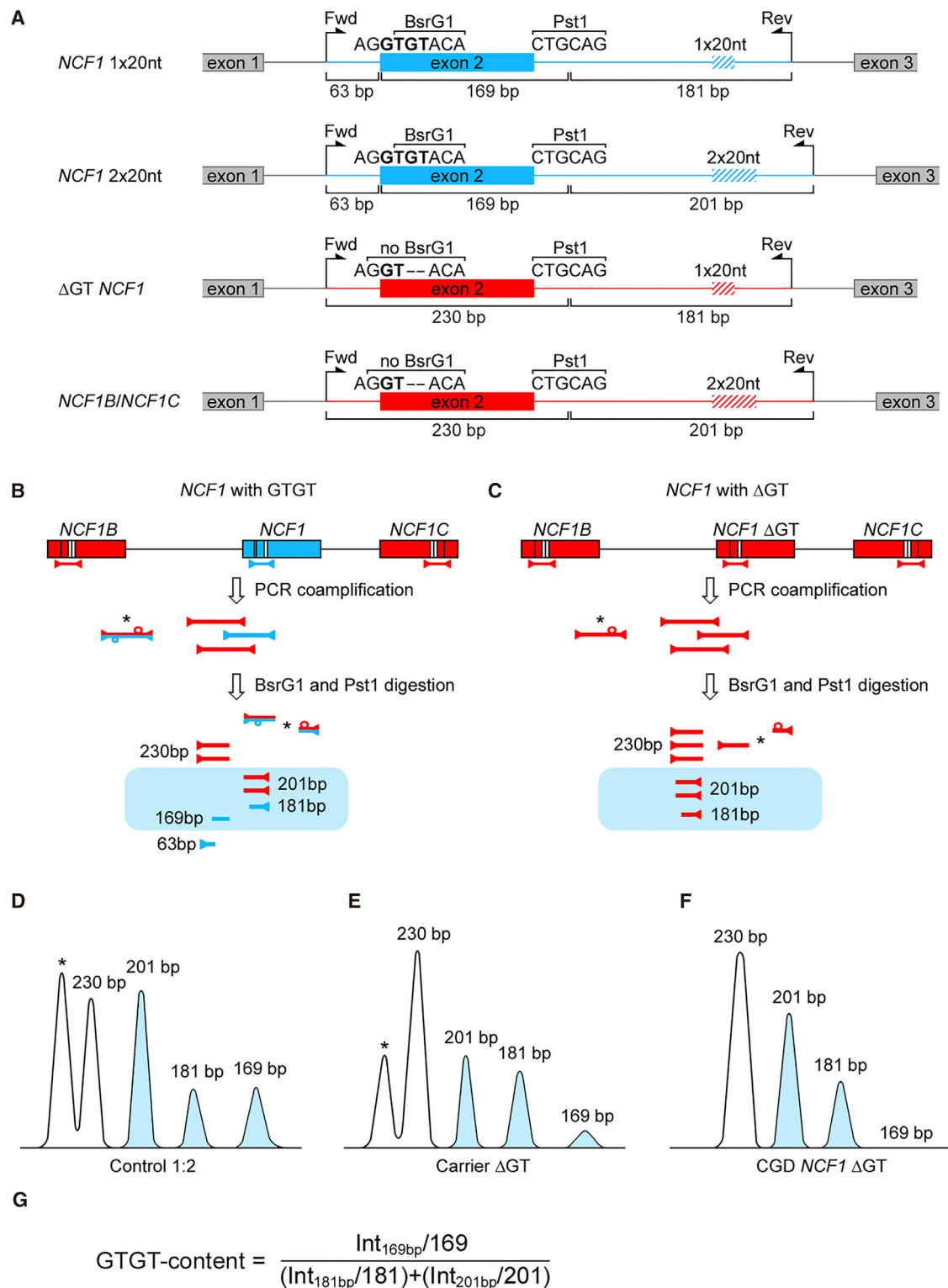


Figure 1. PCR-RFLP Analysis of *NCF1* Loci

(A) Co-amplified fragments of the *NCF1* gene and pseudogenes. Positions of the GT-dinucleotide, 20-nt repeat, BsrG1 and Pst1 restriction sites, and primer-binding sites (forward and reverse) are shown. (B) In healthy individuals and (C) in patients with the Δ GT mutation in *NCF1*, the PCR co-amplification of *NCF1* (correct, blue; mutated, red)

(legend continued on next page)

The co-amplification PCR products spanned the *NCF1* gene and pseudogene GTGT locus within exon 2, as well as the surrounding intronic regions containing one or two repeats of a 20-nt sequence (1×20 nt or 2×20 nt; Figure 1A). Figure 1 displays possible DNA sequence variations of these loci, their configuration on chromosome 7, corresponding PCR co-amplification products, as well as PCR-RFLP results. CGD patients with homozygous Δ GT mutation in *NCF1* can be identified by electrophoresis based on the absence of the 169-bp band (Figures 1C, 1F, and 2A, and 2B). The intensity of the 169-bp band was substantially weaker in heterozygous carriers. The absence of the 181-bp band (Figure 2, control 3) was observed in individuals who had two copies of the 20-nt repeat in intron 2 in all *NCF1* gene and pseudogene alleles, a genotype that may be observed in healthy individuals. Calculation of the GTGT content (Figure 1G) allowed for the differentiation between *NCF1* Δ GT mutation carriers and healthy individuals. Representative PCR-RFLP samples developed by PAGE and agarose electrophoresis (Figures 2A–2C) compare controls (healthy individuals with GTGT in two *NCF1* alleles and different 20-nt intronic repeat numbers), X-CGD and autosomal recessive CGD *NCF2* (gp91^{phox} and p67^{phox} deficiency, respectively), two *NCF1* Δ GT carriers (Carrier Δ GT #1 and Carrier Δ GT #2), an induced pluripotent stem cell (iPSC) *NCF1* Δ GT cell line,⁹ a human acute myeloid leukemia cell line PLB-985 (wild-type), and a PLB-985 *NCF1* Δ GT cell line.⁸

The median GTGT content in 55 healthy individuals who carried two *NCF1* gene alleles with the GTGT sequence and the Δ GT mutation in all four pseudogene alleles (Control ratio 1:2; Figure S1A) was 0.29 in the polyacrylamide gel and 0.20 in the agarose gel. With *NCF1* allele PCR co-amplification, followed by single-molecule real-time (SMRT) sequencing (Table S1), we identified five healthy individuals with two functional *NCF1* alleles with GTGT sequence plus one of four pseudogene alleles with GTGT sequence (Control ratio 1:1; Figure S1B). In these individuals, the determined GTGT content values were 0.49 and 0.29 in polyacrylamide and in agarose gel, respectively. The observed results in agarose gels were lower than theoretical values expected by the genetic background, which may be explained by a partial loss of signal intensity in the thick agarose gels. GTGT content values established for 8 Δ GT p47^{phox} CGD carriers, in which one *NCF1* allele carries the Δ GT mutation, was 0.10 in polyacrylamide gel (Figure 2D) and 0.08 in agarose gel (Figure 2E; data for individual samples and statistics in Table S1).

The results of the PCR-RFLP analysis were confirmed by SMRT sequencing: undigested pools of barcoded co-amplification PCR products (Figures 1A and 1B) were sequenced, and the frequencies of GTGT signature-containing reads were calculated (Figure 2F;

Table S1). GTGT sequence was identified in 27% of reads from healthy individuals with two GTGT-carrying *NCF1* alleles (Control ratio 1:2), in 37% of reads from healthy individuals with GTGT within two *NCF1* alleles and one of four pseudogene alleles (Control ratio 1:1), and in 16% of reads from *NCF1* Δ GT carriers. Percentage differences in the GTGT signature-containing reads were statistically significant, and they allowed for the discrimination of healthy individuals from *NCF1* Δ GT carriers and CGD patients.

In conclusion, we propose a package of complementary methods to be used for single-day reliable p47^{phox} CGD diagnostics, based on PCR-RFLP, giving comparable results to SMRT sequencing. Furthermore, the PCR-RFLP diagnostic method represents an attractive alternative to the existing methods used in CGD diagnostics in terms of appliance requirements and costs per tested sample (Table 1). In addition to diagnostics, both methods can be also effectively applied for the assessment of correction of the Δ GT mutation upon genome-editing-based gene therapy.

MATERIALS AND METHODS

DNA Isolation and PCR Amplification

Sample processing was covered by ethical vote KEK ZH 2015/0135, BASEC-Nr. PB_2016-02202. Genomic DNA from healthy individuals, diagnosed Δ GT p47^{phox} CGD patients, and their family members was isolated using DNeasy Blood & Tissue Kit (QIAGEN, Hombrechtikon, Switzerland). The 411- to 433-bp fragments of *NCF1*, *NCF1B*, and *NCF1C* genes were PCR co-amplified using published PCR primers⁶ (Microsynth, Balgach, Switzerland). Phusion High-Fidelity DNA Polymerase and deoxyribonucleotides (dNTPs) were from Thermo Fisher Scientific (Reinach, Switzerland). PCR reaction included GC 10 \times buffer, dNTPs (200 μ M each), primers (240 nM each), 0.04 U/ μ L Phusion High-Fidelity DNA Polymerase, and 2.5 ng/ μ L DNA. Initial 3-min denaturation (95°C) was followed by 36–40 cycles of denaturation (95°C, 30 s), annealing (65°C, 30 s), and elongation (72°C, 15 s) and final elongation (72°C, 3 min).

Determination of the GTGT Content by RFLP

The PCR co-amplification products of *NCF1*, *NCF1B*, and *NCF1C* were digested with BsrG1 and Pst1 (New England Biolabs, Frankfurt am Main, Germany) (37°C, 180 min), followed by enzyme inactivation (80°C, 20 min) (Figures 1A–1C). The digestion fragments were developed in a 7.5% polyacrylamide (ratio 29:1) gel or a 5% (w/v) agarose gel stained with GelRed Nucleic Acid Gel Stain (Biotum, Fremont, CA, USA) and visualized using Gel Logic 100 Imaging System (Kodak, Eysins, Switzerland). Band intensities were quantified with ImageJ software.¹⁰ For determination of the GTGT content

and its pseudogenes (red) results in a mixture of PCR products with a defined stoichiometry. In the majority of individuals, co-amplified PCR products differ by 2-nt of the GT-dinucleotide locus, and in 20-nt of the intronic 20-nt repeat sequence. A significant fraction of the mixture comprises cross-hybridized PCR products derived from *NCF1* and the pseudogenes (marked with an asterisk) (Figures S2–S6). BsrG1 and Pst1 restriction digestion leads to the appearance of up to seven different restriction fragments in healthy individuals (A) and up to five fragments in patients with Δ GT deletion in *NCF1* (B) (Figure 2A). (D–F) Typical densitometry images of digested fragments in a polyacrylamide gel of (D) a healthy individual with *NCF1* gene to pseudogene ratio 1:2 (Control 1:2), (E) a carrier of the Δ GT mutation with *NCF1* gene to pseudogene ratio 1:5 (Carrier Δ GT), and (F) a Δ GT p47^{phox} CGD patient (CGD *NCF1* Δ GT). (G) Size-normalized band intensities of 169-, 181-, and 201-bp fragments (B–F, blue) are used for calculation of the GTGT content and for identification of Δ GT p47^{phox} CGD patients and Δ GT mutation carriers.

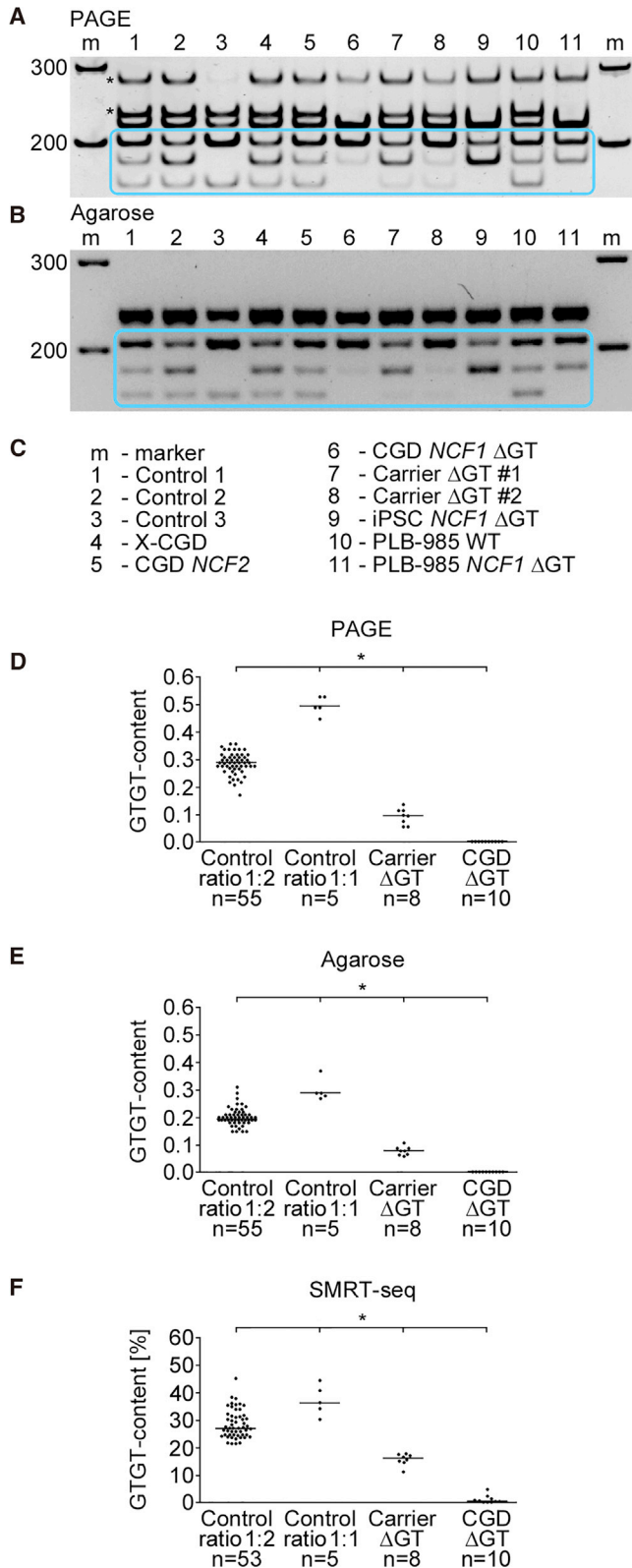


Figure 2. GTGT Content Determination by PCR-RFLP

(A) PAGE and (B) agarose gel electrophoresis of BsrG1- and PstI-digested PCR co-amplification products of the *NCF1* loci (Figures 1B and 1C; full gel images available in Figure S7). The band of size 63-bp is not displayed in the gels. Bands of 201-, 181-, and 169-bp (blue box) were used to determine the GTGT content (Figure 1G). (C) List of samples presented in (A) and (B). Controls 1–3, GTGT sequence in both *NCF1* gene alleles; Control 1, a single 20-nt intronic repeat (1×20 nt) in two and a double 20-nt repeat (2×20 nt) in four *NCF1* alleles; Control 2, 1×20 nt in three of six *NCF1* alleles whereas 2×20 nt in three remaining *NCF1* alleles; Control 3, all six *NCF1* alleles contain the 2×20 nt; X-CGD and CGD *NCF2*, gp91^{phox}- and p67^{phox}-deficient CGD, respectively; carrier ΔGT, heterozygous ΔGT mutation in *NCF1*; iPSC *NCF1* ΔGT, induced pluripotent stem cell line with a homozygous ΔGT mutation in *NCF1* (see Jiang et al.⁹); human acute myeloid leukemia cell line PLB-985 (wild-type); and a cellular model of ΔGT p47^{phox} CGD, PLB-985 *NCF1* ΔGT (see Wrona et al.⁹). (D and E) The GTGT content in polyacrylamide (D) and agarose gels (E). (F) The GTGT content determined by SMRT sequencing. Control ratio 1:2 (Figure S1A), the GTGT sequence in *NCF1* and the ΔGT in *NCF1B* and *NCF1C*; Control ratio 1:1 (Figure S1B), the GTGT sequence in *NCF1* and in one of four *NCF1* pseudogene alleles. (D–F) Horizontal lines represent median values. *p < 0.01, n, number of samples.

(Figure 1G), RFLP band intensities of 169-, 181-, and 201-bp BsrG1/PstI digestion products (Figures 1D–1F, 2A, and 2B) were size normalized by dividing band intensities by their length (number of base pairs). The size-normalized band intensity of the 169-bp band was divided by the sum of normalized band intensities of the 181- and 201-bp bands (Figure 1G).

SMRT Sequencing

PCR co-amplification products of the *NCF1* gene and its pseudogenes were produced using individually barcoded Fwd1 primers, utilizing PCR conditions described above. PCR products were purified using the QIAquick Gel Purification Kit (QIAGEN). 20 ng gel-purified bar-coded PCR products of individual subjects were pooled and analyzed with SMRT sequencing¹¹ by Functional Genomics Center Zurich, ETH/University of Zurich, Switzerland, as described.⁸

Statistical Analysis

The Kruskal-Wallis tests with post hoc Dunn’s multiple comparison tests were performed using IBM SPSS Statistics version 23.0 (IBM, Armonk, NY, USA).

Table 1. Comparison of Methods Used for ΔGT p47^{phox} CGD Diagnostics

| Method | Time (Days) | Primer Labeling | Equipment | Cost |
|---|-------------|--|---|------|
| PCR-RFLP (new method described in this article) | 1* | no* | PAGE or agarose electrophoresis system* | * |
| Gene scan | 1* | yes (fluorescently labeled)** | DNA sequencer*** | *** |
| Allele-specific hybridization | 2** | yes (32P oligonucleotides)** | autoradiography equipment* | ** |
| TaqMan CNV | 1* | yes (fluorescently and MGB-labeled probe)*** | qPCR instrument**** | *** |

Asterisks indicate time and resource requirements (*lowest, ***highest). MGB, minor groove binder.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and two tables and can be found with this article online at <https://doi.org/10.1016/j.omtm.2019.02.001>.

AUTHOR CONTRIBUTIONS

D.W. conducted experiments. D.W., U.S., and J.R. designed experiments, analyzed data, made figures, and wrote the manuscript.

CONFLICTS OF INTEREST

The described diagnostic tool has been submitted for intellectual property (IP) filing.

ACKNOWLEDGMENTS

This study was supported by the CGD Society (grant CGDS16/01) and Hochspezialisierte Medizin Schwerpunkt Immunologie (HSM-2-Immunologie). D.W. received a research grant from the University of Zurich (Forschungskredit, grant FK-17-041). J.R. is supported by the Uniscientia Foundation and the Clinical Research Priority Program ImmuGene of the University of Zurich. We thank the Functional Genomics Center Zurich for performing SMRT sequencing and data analysis.

REFERENCES

1. Segal, B.H., Leto, T.L., Gallin, J.I., Malech, H.L., and Holland, S.M. (2000). Genetic, biochemical, and clinical features of chronic granulomatous disease. *Medicine (Baltimore)* 79, 170–200.
2. Noack, D., Rae, J., Cross, A.R., Ellis, B.A., Newburger, P.E., Curnutte, J.T., and Heyworth, P.G. (2001). Autosomal recessive chronic granulomatous disease caused by defects in NCF-1, the gene encoding the phagocyte p47-phox: mutations not arising in the NCF-1 pseudogenes. *Blood* 97, 305–311.
3. Hayrapetyan, A., Dencher, P.C.D., van Leeuwen, K., de Boer, M., and Roos, D. (2013). Different unequal cross-over events between NCF1 and its pseudogenes in autosomal p47(phox)-deficient chronic granulomatous disease. *Biochim. Biophys. Acta* 1832, 1662–1672.
4. Roesler, J., Curnutte, J.T., Rae, J., Barrett, D., Patino, P., Chanock, S.J., and Goerlach, A. (2000). Recombination events between the p47-phox gene and its highly homologous pseudogenes are the main cause of autosomal recessive chronic granulomatous disease. *Blood* 95, 2150–2156.
5. Vázquez, N., Lehrnbecher, T., Chen, R., Christensen, B.L., Gallin, J.I., Malech, H., Holland, S., Zhu, S., and Chanock, S.J. (2001). Mutational analysis of patients with p47-phox-deficient chronic granulomatous disease: The significance of recombination events between the p47-phox gene (NCF1) and its highly homologous pseudogenes. *Exp. Hematol.* 29, 234–243.
6. Dekker, J., de Boer, M., and Roos, D. (2001). Gene-scan method for the recognition of carriers and patients with p47(phox)-deficient autosomal recessive chronic granulomatous disease. *Exp. Hematol.* 29, 1319–1325.
7. Heyworth, P.G., Noack, D., and Cross, A.R. (2002). Identification of a novel NCF-1 (p47-phox) pseudogene not containing the signature GT deletion: significance for A47 degrees chronic granulomatous disease carrier detection. *Blood* 100, 1845–1851.
8. Wrona, D., Siler, U., and Reichenbach, J. (2017). CRISPR/Cas9-generated p47^{phox}-deficient cell line for Chronic Granulomatous Disease gene therapy vector development. *Sci. Rep.* 7, 44187.
9. Jiang, Y., Cowley, S.A., Siler, U., Melguizo, D., Tilgner, K., Browne, C., Dewilton, A., Przyborski, S., Saretzki, G., James, W.S., et al. (2012). Derivation and functional analysis of patient-specific induced pluripotent stem cells as an in vitro model of chronic granulomatous disease. *Stem Cells* 30, 599–611.
10. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675.
11. Levene, M.J., Korlach, J., Turner, S.W., Foquet, M., Craighead, H.G., and Webb, W.W. (2003). Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* 299, 682–686.

OMTM, Volume 13

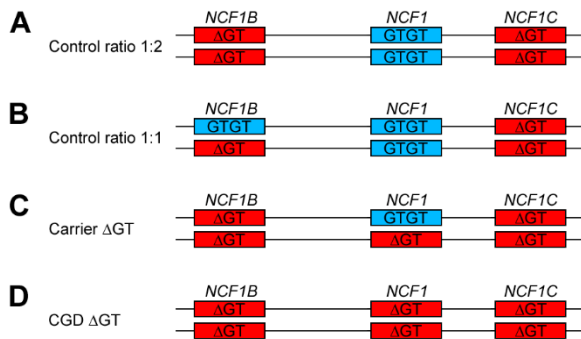
Supplemental Information

Novel Diagnostic Tool

**for p47^{phox}-Deficient Chronic Granulomatous
Disease Patient and Carrier Detection**

Dominik Wrona, Ulrich Siler, and Janine Reichenbach

Supplemental Information
Supplemental Figures
Figure S1



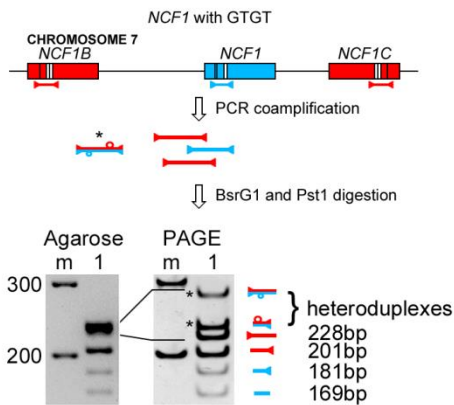
The scheme of the genetic configuration of GTGT and ΔGT sequences located in *NCF1* and pseudogene alleles on chromosome 7

(A) A majority of healthy individuals exhibit the GTGT sequence in both alleles of the *NCF1* gene, and the ΔGT mutation in all four alleles of the *NCF1* pseudogenes (*NCF1B* and *NCF1C*), resulting in a ratio of GTGT to ΔGT carrying *NCF1* alleles of 1:2. (B) A minority of healthy individuals exhibit three GTGT carrying *NCF1* alleles, two located on both *NCF1* gene alleles and one present on one of the pseudogenes, while the ΔGT mutation is shared by the remaining three *NCF1* pseudogene alleles. Hence, in these healthy individuals the ratio of GTGT to ΔGT carrying *NCF1* alleles is 1:1. (C) In carriers of the ΔGT mutation, GTGT is present in only one of the *NCF1* gene alleles, while (D) in ΔGT p47^{phox} CGD patients, all *NCF1* loci carry the ΔGT mutation, resulting in a ratio of 1:5 and 0:6, respectively.

Introduction to Figures S2 to S6

The following figures describe the origin of heteroduplex PCR products that result from simultaneous PCR co-amplification of the *NCF1*, *NCF1B* and *NCF1C* sequences.

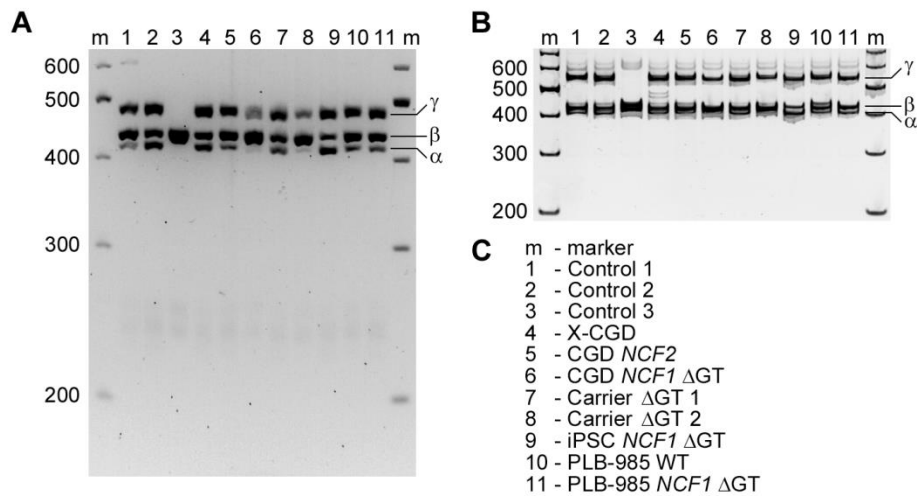
Supplemental Figure S2



PCR co-amplification and RFLP products of the *NCF1* gene and its pseudogenes developed in agarose and polyacrylamide gels.

The heteroduplex PCR products (marked with asterisks) form during the PCR reaction by annealing of single-stranded PCR products derived from the *NCF1* gene and its pseudogenes. These heteroduplexes migrate differently in polyacrylamide and agarose gels, as compared to the homoduplex PCR products, even after restriction digestion. PCR-RFLP product of size 63 bp is not shown in the gels.

Figure S3



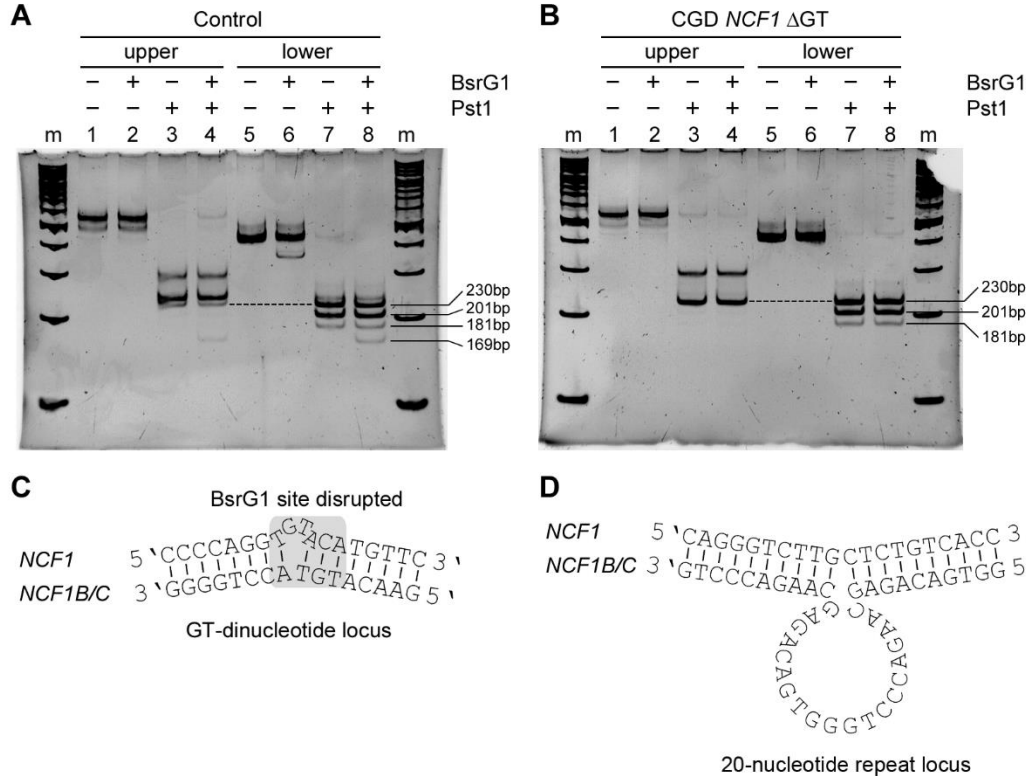
PCR products of co-amplified *NCF1*, *NCF1B*, and *NCF1C* sequences in agarose (A) and polyacrylamide (B) gels

Sequences around the *NCF1* exon 2 were PCR co-amplified utilizing one pair of primers, and representative PCR products from eleven individuals (C) were developed in a 5 % agarose (A) or a 7.5 % polyacrylamide (B) gel. Migration of bands “α” and “β” indicate PCR product sizes between 400 and 450 bp in both, the agarose gel (A) and the polyacrylamide gel (B). The band “γ” migrates differently: in the agarose gel it is located below the 500 bp band of the marker, whereas in the polyacrylamide gel it is located between 500 bp and 600 bp.

In healthy controls 1 and 2 (Control 1 and 2; lanes 1 and 2 in A and B) restriction digestion of PCR products with BsrG1 and Pst1 generates a 201 bp fragment derived from the pseudogenes and a 181 bp fragment (see also: **Figure 1B and C**, as well as **Figure 2A and B**) derived from the *NCF1* gene. These fragments differ from one another only by the presence or absence of the second copy of the 20-nucleotide (20-nt) sequence repeat within intron 2.

Interestingly, PCR products of healthy control 3 (Control 3; lane 3 in A and B) do not include PCR products “α” and “γ”. Restriction digestion of the PCR products of this control with BsrG1 and Pst1 results in a band pattern where the 181 bp band is absent (see: **Figure 2A, B**). The lack of this band indicates the presence of a double 20-nt repeat sequence within the intron 2 of all six *NCF1* loci. This configuration of the *NCF1* intronic 20-nt repeat has not been linked to the CGD phenotype.

Figure S4



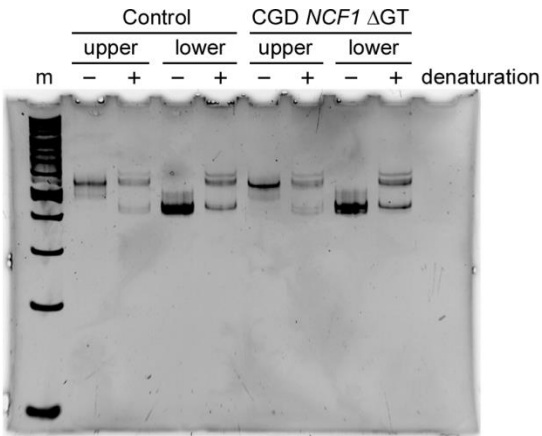
The restriction fragment length polymorphism analysis of upper (“γ”) and lower fractions (“α” and “β”) of PCR co-amplification products of the *NCF1* loci

PCR co-amplification products of the *NCF1* loci from a healthy control, and from a ΔGT p47^{phox} CGD patient were separated in an agarose gel (see: **Figure S3A**), followed by gel purification of upper and lower fractions (gel not shown). The lower fraction consisted of bands “α” and “β” (see: **Figure S3A**) and the upper fraction corresponded to the band “γ” (see: **Figure S3A**). To avoid changes in the composition of isolated cross-hybridized PCR heteroduplexes, the temperature has not exceeded 46 °C during purification. The upper (band “γ”) and lower (bands “α” and “β”) fractions of PCR co-amplification products were subjected to restriction digestion with BsrG1 and Pst1 restriction enzymes. Digestion was performed as described in Methods, without the final denaturation of DNA and without inactivation of the enzymes at 80 °C. The digestion products were visualized in a 7.5 % polyacrylamide gel. In healthy controls the lower fraction consisting of bands “α” and “β” (**Figure S4A and B, lanes 5 to 8**) can be digested by BsrG1 (**Figure S4A, lane 6**) and by Pst1 (**Figure S4A, lane 7**), whereas double digestion results in the expected band pattern for healthy individuals, with four bands in total (**Figure S4A, lane 8**). In ΔGT p47^{phox} CGD, however, the ΔGT abolishes the BsrG1 restriction site in the lower fraction of PCR products (**Figure S4B, lane 6**), resulting only in 3 bands after digestion with BsrG1 and Pst1 (**Figure S4B, lane 8**). As expected, the smallest 169 bp band, which is the result of BsrG1 digestion of the PCR co-amplification products from the healthy control (**Figure S4A, lanes 6 and 8**), is absent in digestion samples from the ΔGT p47^{phox} CGD patient (**Figure S4B, lane 8**). The 181 bp and 201 bp bands were produced solely by Pst1 digestion of the lower fraction of PCR products of both, the healthy control and the ΔGT p47^{phox} CGD patient (**Figure S4A and B, lanes 7 and 8**).

The upper fraction of the undigested PCR co-amplification products (see: **Figure S3A and B, band “γ”**) can be digested only by Pst1 for both samples, from the healthy individual (**Figure S4A, lanes 1 to 4**) and from ΔGT p47^{phox}-deficient CGD patient (**Figure S4B, lanes 1 to 4**). The resulting two bands are also present in the polyacrylamide gel for samples derived from healthy controls and carriers of the ΔGT mutation (**Figure 2A**, marked with asterisks). Of these two, the smaller digestion product from the ΔGT p47^{phox} CGD patient overlaps with the 230 bp band derived from digestion of the lower fraction with both BsrG1 and Pst1 enzymes (**Figure 2A lanes 6, 9 and 11, and Figure S4A and B**; a dashed line indicates the difference in the position of the smaller digestion products of the upper fraction in lanes 3, 4, 7 and 8).

The bending centers responsible for the altered curvature and electrophoresis motility of heteroduplex PCR products are presented in **Figure S4C and D**: the Δ GT mutation in one strand of the GTGT locus (**Figure S4C**) and the 20-nt repeat locus (**Figure S4D**). In healthy controls, the heteroduplexes contain one strand derived from the pseudogenes carrying the Δ GT mutation, and another from the *NCF1* gene with the GTGT sequence, in which the restriction site for BsrG1 is disrupted (**Figure S4C**). These molecules retain their curvature upon restriction digestion with Pst1, as they still migrate slower (**Figure S3A, lane 3 and 4**). Conversely, the sequence of the 5' Pst1 digested fragment for Δ GT p47^{phox} CGD patients is shared between the mutated *NCF1* and the pseudogenes. Because of Pst1 digestion, the 5' digestion fragment of the upper fraction loses its curvature and migrates with the same speed as the 230 bp fragment of the lower fraction (**Figure S3B, lane 3 and 4**). Additionally, the pattern of digested PCR products for healthy control 3, who exhibits two copies of the 20-nt intronic repeat in all *NCF1* alleles (**Figure 2A, lane 3**), does not display the upper of the two additional bands marked with asterisks in **Figure 2A, lane 3**. The lack of this band again can be explained by lost DNA curvature upon Pst1 digestion. As already mentioned, this healthy control presents two copies of the 20-nt repeat sequence in all *NCF1* loci, and Pst1 digestion releases the linear 3' fragment of the heteroduplex that migrates as the regular 201 bp restriction fragment in **Figure 2A, lane 3**.

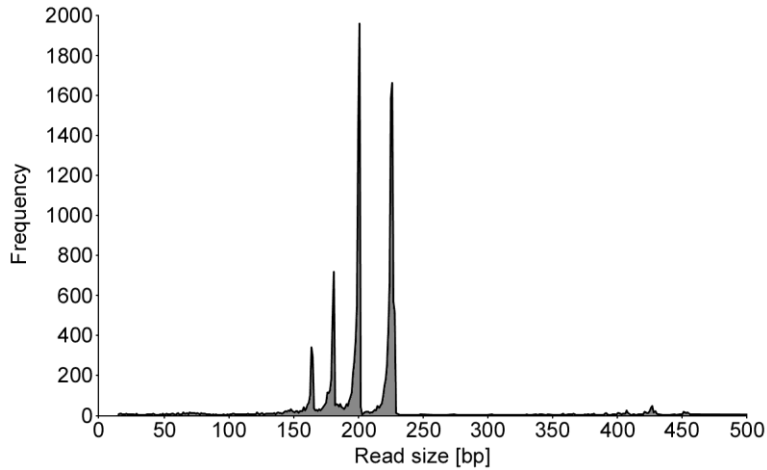
Figure S5



Denaturation and renaturation of the upper and lower fractions of the PCR co-amplification products

PCR co-amplification products of the *NCF1* loci from a healthy control and a Δ GT p47^{phox} CGD patient were separated in an agarose gel, the lower fraction consisting of bands “ α ” and “ β ” (see: **Figure S3**), and the upper fraction corresponding to band “ γ ” (see: **Figure S3**) were purified (gel not shown). During isolation the temperature has not exceeded 46 °C. Thereafter, isolated upper and lower fractions were denatured at 95 °C for 3 minutes and renatured at room temperature. The denatured and renatured samples were developed in a 7.5 % polyacrylamide gel along the unprocessed samples. Denaturation of all fractions led to reappearance of the PCR heteroduplexes, indicating that the band “ γ ” (**Figure S3**) is not an unspecific PCR product, but a product of cross-hybridization of PCR co-amplification products derived from different *NCF1* loci, the *NCF1* gene and the *NCF1* pseudogenes.

Figure S6

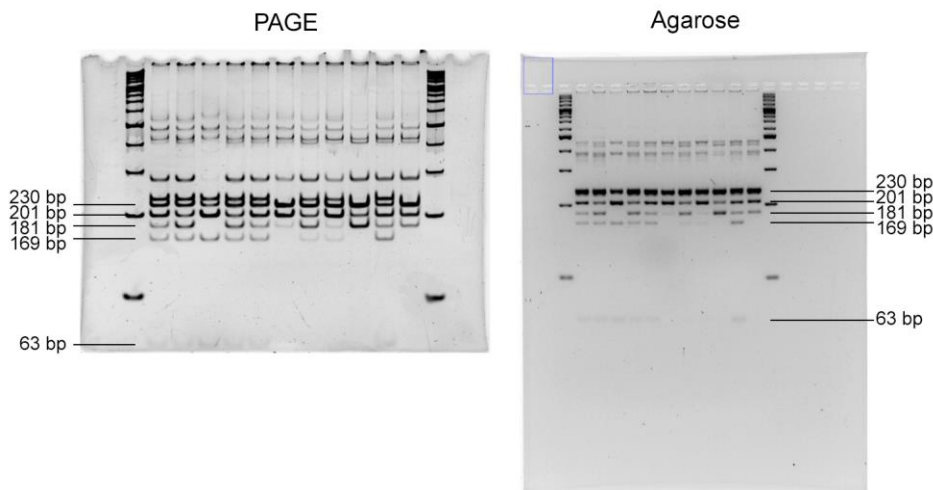


SMRT sequencing of digested PCR co-amplification

PCR co-amplification products of *NCF1* loci consisting of the upper fraction (band “ γ ”) and the lower fraction (bands “ α ” and “ β ”), as visualized in **Figure S3**, were digested with BsrG1 and Pst1 resulting in up to seven bands in a polyacrylamide gel (see: **Figure 1B**). All bands except for 63 bp were gel purified and analyzed by SMRT sequencing, as described in Methods. Size distribution of sequenced reads identified four distinct products that correspond to the lengths of expected digestion fragments of the PCR homoduplexes.

Altogether, the results presented in Supplemental Figures S2-S6 suggest that the upper fraction (**Figure S3A and B, band “ γ ”**) consists of cross-hybridized PCR heteroduplexes, where one strand of the DNA heteroduplex originates from the *NCF1* gene and another from one of its pseudogenes (see: **Figure S2**, marked with asterisks). These heteroduplexes are curved DNA molecules that contain two major bending centers: the GT-dinucleotide deletion in one strand in the GTGT locus (**Figure S4C**) and the 20-nt repeat locus (**Figure S4D**). Due to the curvature, the heteroduplexes migrate slower than expected by their size, as reported before for other curved DNA molecules.¹

Figure S7



Uncropped PAGE and agarose gels shown in **Figure 2A and B** of the main text

Supplemental Tables

Table S1

The list of Δ GT p47^{phox}-deficient CGD patients and carriers of the Δ GT mutation tested in the study.

The Table S1 is in the excel file: Supplemental Table S1.

Table S2

The GTGT-content calculated for individual tested in the study, CGD patients, carriers and controls, as determined by restriction length polymorphism visualized in polyacrylamide and in agarose gels, or determined by SMRT sequencing.

The Table S2 is in the excel file: Supplemental Table S2.

Supplemental References

1. Stellwagen, N. C. Electrophoresis of DNA in agarose gels, polyacrylamide gels and in free solution. *Electrophoresis* **30 Suppl 1**, S188-95 (2009). DOI: 10.1002/elps.200900052