Gene promoter analysis in molecular diagnostics: do or don't?

Expert Rev. Mol. Diagn. 9(5), 403–405 (2009)

Karen MK de Vooght, PhD
Author for correspondence
Department of Clinical Chemistry and Haematology, Laboratory for Red Blood Cell Research, University Medical Center Utrecht, Postbus 85500, 3508 GA, Utrecht, The Netherlands
Tel.: +31 887 558 748
Fax: +31 887 555 418
k.devooght@umcutrecht.nl

Wouter W van Solinge, PhD
Department of Clinical Chemistry and Haematology, Laboratory for Red Blood Cell Research, University Medical Center Utrecht, Postbus 85500, 3508 GA, Utrecht, The Netherlands
Tel.: +31 887 559 594
Fax: +31 887 555 418
wsolinge@umcutrecht.nl

"Mutations in the promoter region of a gene can alter or abolish the binding ability of sequence motifs for the transcription factors that normally interact with them, thereby disrupting the normal process of gene activation and transcriptional initiation."

Approximately 1% of single basepair substitutions causing human genetic disease occur within gene promoter regions where they disrupt the normal processes of gene activation and transcriptional initiation, and usually decrease or increase the level of mRNA and, thus, protein [1]. Although promoter mutations are known to have functionally important consequences for gene expression, promoter analysis is not a regular part of molecular diagnostics. One reason is that the effect of promoter mutations can be very subtle. For instance, the majority of missense mutations cause a fairly easy to identify qualitative defect. By contrast, promoter mutations may cause small quantitative defects that may be hard to detect. Even if the promoter of an autosomal gene is completely downregulated as result of mutation, half of the normal amount of protein is present, which is often enough to prevent severe disease. Another reason for not performing promoter analysis is that it is complex and the assays that are needed to investigate the functional relationship between the mutation and disease are laborious and difficult to perform. As a result, thorough studies of promoter mutations are scarce and often confined to research laboratories. By contrast, in epidemiological studies in which large amounts of single-nucleotide polymorphisms (SNPs) are studied in relation to disease, several promoter SNPs have recently been shown to be associated with specific diseases. Unfortunately, appropriate functional analysis of these SNPs is often lacking and it can be questioned if they are themselves causing disease.

Gene promoter mutations

The promoter of a gene is a regulatory region of DNA located upstream of that specific gene, typically consisting of a core promoter (surrounding the transcriptional start site) and a proximal promoter. Both contain multiple sequence-specific binding sites for transcription factors. The structural gathering of these DNA-binding proteins directs basal transcription (core promoter) and enhances transcriptional activity (proximal promoter) [2]. Mutations in the promoter region of a gene can alter or abolish the binding ability of sequence motifs for the transcription factors that normally interact with them, thereby disrupting the normal process of gene activation and transcriptional initiation. There is hardly any literature regarding the incidence of promoter mutations in relation to disease. However, the incidence of reported promoter mutations in genes seems to depend on the extensiveness by which genes have been studied in the past and the severity of disease caused by the mutation [3]. Examples of diseases caused by promoter mutations are β-thalassemia [4], Bernard–Soulier syndrome [5], pyruvate kinase deficiency [6], familial hypercholesterolemia [101] and hemophilia [7] (for a review, see [8]). In general, polymorphic sequence variations are considered to be rather harmless, especially if located in noncoding sections of a gene. However, polymorphic variations that occur in the promoter may affect gene expression and may, thus, have the potential to be of phenotypic or even of pathological significance [1]. An increasing number of promoter polymorphisms has been characterized by

EXPERT REVIEWS

 10.1586/ERM.09.28 ISSN 1473-7159

functional studies. Some may well be of pathological significance (e.g., those in the plasminogen activator inhibitor type 1, TNF-α, Apo AI, lipoprotein lipase and IL-6 genes [1]).

Promoter mutation analysis

The analysis of disease-causing promoter mutations typically consists of *in silico* analysis of the promoter mutation, functional promoter assays, such as transient transfection assays, and DNA-transcription factor binding assays, such as electrophoretic mobility shift assays (EMSAs) and chromatin immunoprecipitation (ChIP) assays [3]. Overall, promoter assays are complex and time-consuming, making specific demands on laboratory equipment (i.e., cloning and cell culturing facilities). A prerequisite for lowering the hurdles to promoter mutation analysis is the development of dedicated and simple functional promoter assays and transcription factor binding assays. At this moment, laboratories need cell-culturing facilities to perform functional (transfection) assays. The development of a cell-free functional assay would increase the opportunities for general laboratories to characterize promoter mutations and sequence variations. Traditionally, EMSAs and supershift assays are performed using radiolabeled oligonucleotide probes. Owing to safety concerns associated with the proper storage, usage and disposal of radioactivity, and to make the assay more accessible to laboratories without an isotope laboratory or phosphoimager, nonradioactive EMSAs have been developed. In these assays, biotinylated probes, chemiluminescence detection of fluorescence probes are used [9]. A disadvantage, however, is that the sensitivity of these assays is often lower than that of radioactive EMSAs.

> "Overall, promoter assays are complex and time-consuming, making specific demands on laboratory equipment..."

To improve the use of EMSAs, more-sensitive nonradioactive probes or techniques have to be developed. The inclusion of ChIP assay data in promoter mutation studies is often a prerequisite for publication. The value of this assay lies in the opportunity to demonstrate the *in vivo* relevance of transcription factor binding. The main disadvantage of this assay is that it is technically challenging. The method requires high-quality antibodies capable of recognizing the fixed transcription factor and the optimization of chromatin shearing conditions can be difficult [10]. Commercial ChIP assay kits have become available recently, and optimization of these assays will make them more accessible for less-experienced laboratories. Therefore, not only dedicated assays will make promoter mutation analysis easier. To investigate whether a promoter mutation disrupts or creates a putative transcription factor binding site, and thereby to assess the change of a 'positive outcome' of an extensive characterization of a promoter mutation, better prediction programs for *in silico* analysis are needed. The improvement of these prediction programs strongly depends on the expansion of the data input (e.g., transcription factors, transcription factor binding sequences and mutations), improvement of the quality of data (solid and complete functional and binding assay results), improvement of species specificity and inclusion of data regarding tissue-specific transcription factors and expression. Knowledge of tissue-specific expression of genes appears to be crucial for proper promoter mutation analysis.

Pros & cons of gene promoter analysis in molecular diagnostics

There are several reasons for a reserved standpoint towards promoter mutation analysis, the main one being the opinion that promoter mutations have only a subtle effect on transcription, generally too mild to cause disease. Even when there seems to be an association between one specific promoter mutation and disease, (i.e., low pyruvate kinase activity and the -148C>T mutation in the erythroid-specific promoter of *PKLR*), it may be impossible to prove it using *in vitro* transfection studies [11]. In addition, promoter mutations that turn out to be nonfunctional are often not reported (publication bias). Furthermore, the assessment of the relevance of a promoter mutation is difficult because the location of the promoter is frequently not well defined, the effort needed to analyze these mutations is overly high and it is often not possible to report unambiguously on the clinical effect of the identified mutation. Advocates of extensive promoter mutation analysis refer to overviews of promoter mutations that have been proven to cause disease [8,12]. It is even conceivable that some promoter mutations have failed to be noticed thus far because promoter analysis is not a regular part of the molecular diagnostic line-up in suspected inherited diseases, and if included, often only a limited region of the promoter is analyzed.

Today, epidemiological studies, in which large numbers of SNPs are studied in relation to disease, generate considerable numbers of putative functional promoter SNPs. However, a causal link between these promoter polymorphisms and disease is often absent, since these studies generally lack functional promoter assays. This could lead to the false impression of a high number of promoter mutations having phenotypic consequences. On the other hand, if in this case a causal relationship would be proven by use of functional and binding assays, such disease-associated promoter sequence variations could have a big impact on the general awareness of the importance of promoter mutations.

> "Even when there seems to be an association between one specific promoter mutation and disease ... it may be impossible to prove it using *in vitro* transfection studies."

In our opinion, the only solution to the contradiction on the relevance of promoter mutations is to increase our knowledge by improving promoter mutation analysis. More insight into the carefully orchestrated symphony of transcriptional regulation [13] will expand our understanding of the role of transcriptional regulation in human disease. Therefore, we believe that routine laboratories should cooperate more with research groups involved in gene promoter research. This will work in both directions: routine laboratories can translate the findings of research groups into diagnostic tools, whereas groups involved in gene promoter research, depend on the identification of these mutations in patients, to improve knowledge on transcriptional regulation

of the gene of interest and the role of disturbed transcriptional regulation in disease [3]. Improving the data on the *in vivo* and *in vitro* effects of promoter mutations associated with disease will hopefully lead to a proper assessment of the incidence of disease causing promoter mutations with respect to other known genetic causes of disease, free of most bias or prejudice.

Conclusion

More and better promoter mutation analysis could lead to better *in silico* prediction programs that could prevent needless promoter mutation studies. We hope that, in the future, promoter prediction programs will be able to reliably assess the putative effect of a promoter mutation. In addition, fast and easy-to-perform functional and transcription factor-binding assays will make it possible to instantly report on the *in vivo* effect of a promoter mutation. This will increase the awareness and understanding of the genetic diseases in which we should consider the presence of promoter mutations. Hopefully, we will then also be able to answer the question of whether gene promoter analysis is an essential part of molecular diagnostics.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

1 Cooper DN. Human gene mutation in pathology and evolution. *J. Inherit Metab. Dis.* 25(3), 157–182 (2002).

2 Latchman DS. *Eukaryotic Transcription Factors (3rd Edition)*. Academic Press, London, UK (1998).

3 de Vooght KM, van Wijk R, van Solinge WW. Management of gene promoter mutations in molecular diagnostics. *Clin. Chem.* 55(4), 698–708 (2009).

4 Hardison RC, Chui DHK, Giardine B *et al.* HbVar: a relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. *Hum. Mutat.* 19(3), 225–233 (2002).

5 Ludlow LB, Schick BP, Budarf ML *et al.* Identification of a mutation in a GATA binding site of the platelet glycoprotein Ibβ promoter resulting in the Bernard–Soulier syndrome. *J. Biol. Chem.* 271(36), 22076–22080 (1996).

6 Van Wijk R, Van Solinge WW, Nerlov C *et al.* Disruption of a novel regulatory element in the erythroid-specific promoter of the human *PKLR* gene causes severe pyruvate kinase deficiency. *Blood* 101(4), 1596–1602 (2003).

7 Reijnen MJ, Sladek FM, Bertina RM *et al.* Disruption of a binding site for hepatocyte nuclear factor 4 results in hemophilia B Leyden. *Proc. Natl Acad. Sci. USA* 89(14), 6300–6303 (1992).

8 Maston GA, Evans SK, Green MR. Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.* 7, 29–59 (2006).

9 Ruscher K, Reuter M, Kupper D *et al.* A fluorescence based non-radioactive electrophoretic mobility shift assay. *J. Biotechnol.* 78(2), 163–170 (2000).

10 Orlando V, Strutt H, Paro R. Analysis of chromatin structure by *in vivo* formaldehyde cross-linking. *Methods* 11(2), 205–214 (1997).

11 de Vooght KM, van Wijk R, van Wesel AC *et al.* Characterization of the -148C>T promoter polymorphism in *PKLR*. *Haematologica* 93(9), 1407–1408 (2008).

12 Wray GA. The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* 8(3), 206–216 (2007).

13 Lander ES, Linton LM, Birren B *et al.* Initial sequencing and analysis of the human genome. *Nature* 409(6822), 860–921 (2001).

Website

101 Locus-specific sequence variation database www.ucl.ac.uk/ldlr/Current/