Methylation of the imprinted *GNAS1* gene in cell-free plasma DNA: equal steady-state quantities of methylated and unmethylated DNA in plasma

William M. Puszyk, Kamaljit Chatha, Selina Elsenheimer, Francesco Crea, and

Robert W. Old

Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK

Address for correspondence:

Robert W. Old,

Department of Biological Sciences,

University of Warwick,

Coventry

CV4 7AL,

 UK

Telephone   +44 2476 523528

Fax   +44 2476 523701

Email   r.w.old@warwick.ac.uk

Abstract

Background: Genomic DNA sequences in cell-free plasma are biomarkers of cancer prognosis, where characteristic changes in methylation of tumour suppressor or oncogene DNA regions are indicative of changes in gene activity. Also, cell-free fetal DNA can be distinguished, by its methylation status, from the maternal DNA in the plasma of pregnant women, hence providing DNA biomarkers for the proposed minimally-invasive diagnosis of fetal aneuploidies, including Down's syndrome. However, the production and clearance of cell-free DNA from plasma in relation to its methylation status, are poorly understood processes.

Methods: We studied the methylation status of DNA derived from the imprinted *GNAS1* locus, in cell-free plasma DNA of healthy adults. Heterozygotes were identified that carried the SNP rs1800905 in the imprinted region. The parent-of-origin-dependent DNA methylation was analysed by bisulfite conversion, followed by cloning and sequencing.

Results: Genomic DNA molecules derived from both the methylated, maternal, allele and the unmethylated, paternal, allele were found in plasma. Methylated and unmethylated DNA molecules were present in equal numbers.

Conclusions: Our data indicate that the methylation status of a DNA sequence has no effect on its steady state concentration in the cell-free DNA component of plasma, in healthy adults.

1. Introduction

Cell-free human genomic DNA occurs in the plasma of healthy individuals [1]. The DNA molecules are heterogeneous in size, with the majority ranging from 100 bp to 2 kb, and their numbers vary between individuals, 50-2000 diploid genome equivalents/ml blood being typical for a particular DNA sequence. The origins and clearance of the cell-free plasma (cfp) DNA are poorly understood [2]. However, it has been concluded from studies of sex-mismatched bone marrow transplant patients, that most of the cfp DNA originates from the hematopoietic lineage [3]. Proposed mechanisms for the production of cfp DNA include 'accidental' lysis of nucleated blood cells, apoptosis, and a derivation from nuclei ejected from maturing erythrocytes [2]. A significant contribution from vascular endothelial cells has been excluded [4].

In pathological conditions and pregnancy, cfp DNA can clearly originate elsewhere. Thus, there has been much effort in analysing DNA markers in patients with solid tumours [5,6]. In these situations the quantity and methylation status of DNA regions derived from oncogenes or tumour suppressor genes provide biomarkers for disease progression and prognosis [1]. The changes in DNA methylation status indicate changes in gene expression of the relevant biomarker genes; for example, hypermethylated *APC* gene DNA in plasma may be a useful biomarker of aggressive disease in esophageal adenocarcinoma patients [7]. In pregnancy, a proportion (typically 3-6%) of the cell-free DNA in maternal plasma is of fetal genotype, originating largely from the placenta [8,9]. Cell-free fetal DNA can be detected from as early as 5 weeks after conception [10], and provides a basis for fetal chromosomal

gender determination (using *SRY* or *DYS14* markers [11]), and for detection of an *RhD+* fetus carried by a *RhD-* mother [12]. In those examples fetal DNA is distinguished from the maternal 'background' by its paternal genotype. In other, proposed, applications particular cell-free fetal DNA biomarker sequences in maternal plasma are to be distinguished on the basis of their characteristic methylation status. This is the rationale for minimally-invasive genetic diagnoses of fetal aneuploidies (such as Down's syndrome), using a peripheral blood sample from the mother [13].

It has been concluded from studies of cell-free fetal DNA in plasma that the DNA is cleared very rapidly, with a half-life of under an hour, after delivery of the fetus [14]. It remains unknown whether cfp DNA turns over with similar rapidity in healthy non-pregnant adults, although this seems likely. A known potential effect that may influence amounts of methylated and unmethylated cfp DNA in the circulation arises from interactions of DNA with cells of the immune system. Bacterial DNA with unmethylated CpG dinucleotides is immunostimulatory [15], and hypomethylated DNA from peripheral blood monocytes stimulates an autoimmune response in SLE [16]. The amount of cfp DNA is small (typically 50-2000 ge/ml blood) relative to the numbers of cells of the immune system, of various classes, present in blood. So it is conceivable that clearance of cfp DNA depends in part upon a methylation-dependent interaction with the immune system in healthy adults.

In view of the interest in the methylation of cfp DNA for providing biomarkers, we have investigated whether the methylation of the DNA influences its abundance in the plasma of healthy adults. As a critical test, we have examined the abundance of

genomic DNA molecules derived from the imprinted *GNAS1* locus, which encodes the alpha subunit of the stimulatory G protein, G$_s$.

2. Materials and Methods

2.1 Plasma and cellular DNA

Peripheral venous blood (10ml) was obtained from individual adult donors, with appropriate ethical approval. Blood was collected into EDTA and centrifuged at 1,600 x g for 10 min, and the plasma fraction was recentrifuged at 16,000 x g for 10 min. Plasma was stored frozen at -80°C until required. DNA was extracted using a QIAmp Blood Maxi kit (Qiagen, UK) using the 'Purification of DNA from whole blood' spin protocol, according to the manufacturer's instructions. Peripheral blood leukocyte genomic DNA, isolated from buffy coat preparations from individual donors, was purchased from BioChain (Hayward, California USA).

2.2 Bisulfite conversion

Genomic leukocyte DNA (10μg) was bisulfite-converted using a protocol described previously [17]. For cell-free plasma DNA the bisulfite protocol was modified to allow for the small quanitity of plasma DNA obtained. The plasma DNA was mixed with carrier DNA (2.5 μg of bacterial plasmid DNA) before bisulfite treatment, which was carried out as above, but at 95°C for 5 minutes followed by 55°C for 3.5 hours.

2.3 PCR amplification, cloning and sequencing

To amplify the *GNAS1* region for digestion with *Fnu*DII, the following primers were used 5' ctt cca aaa agg gac cca tc 3' and 5' agc ctc ctc tct tcc cac tc 3'. To amplify the *GNAS1* region after bisulfite conversion, bisulfite-specific primer pairs

(Bsps), were used. The Bsp primers were designed using MethPrimer software [18]. The primers were 5' gat ttt ttt tgt ttt tat gga ttt agg 3' and 5' cta acc aac taa acc tcc tct ctt c 3'. No CpG dinucleotides are included in Bsp sequences, to allow the amplification of both methylated and unmethylated sequences. PCR amplifications were performed as follows: 95°C for 3 min, (95°C for 30 s, 55°C annealing for 30 s, 72°C for 60 s) for 35 cycles, with a final step at 72°C for 10 min. PCRs were carried out in 25µl, including each Bsp primer at 0.4 µM, 2 units of Platinum Taq DNA polymerase (Invitrogen), and 4.0 µl of bisulfite-treated DNA, and standard PCR components. PCR products were purified using QIAquick PCR Purification Kit (Qiagen, UK), eluted in 50 µl of H_2O. For DNA sequence analysis, the amplified DNA was cloned into the pGEM-T Easy (Promega) vector.

3. Results

The *GNAS1* locus (Fig. 1) on chromosome 20 is imprinted, displaying parent-of-origin-dependent DNA methylation that regulates the complex transcriptional activity of this locus [19]. The region of DNA that comprises the XL promoter, first intron and first exon has been shown to be methylated in the maternal allele, in a number of adult tissues. The paternal allele is unmethylated in this DNA region. This region of differential methylation also includes a common single nucleotide polymorphism, rs1800905, which is an A/G difference [20].

The SNP provides an opportunity to distinguish between alleles from different parents in heterozygous individuals. This A/G SNP creates a restriction site polymorphism (Fig.1). for the restriction enzyme *Fnu*DII whose target site is CGCG. This was

exploited to detect individuals heterozygous for the SNP, using amplicons of the relevant region obtained by PCR amplification of cfp DNA obtained from a number of individual healthy adult plasma donors.

The ability to detect SNP heterozygotes in this way (data not shown) immediately indicated that both paternal and maternal alleles were represented in cfp DNA. Making the assumption that parent-of-origin differential methylation is maintained in this adult DNA source, this finding suggested that both methylated and unmethylated molecules were represented in cfp DNA in readily detectable amounts.

To enable us to interrogate the methylation status of several CpG sites in the *GNAS1* imprinted region, and to directly examine the representation of methylated and unmethylated DNA molecules in cfp DNA in detail, we employed bisulfite conversion. Treatment of DNA with bisulfite converts unmethylated cytosine residues to uracil residues. Following PCR amplification by a DNA polymerase (such as Taq DNA polymerase) that is tolerant of uracil in DNA, unmethylated cytosine residues appear as thymine in the amplified product. Methylated cytosine residues are resistant to conversion by bisulfite [17].

Following bisulfite conversion of cfp DNA from SNP heterozygous donors, part of the differentially methylated region was amplified using bisulfite-specific primers. In addition, bisulfite converted cfp DNA from some homozygous donors was similarly processed, as well as leukocyte DNA to represent genomic DNA isolated directly from a cellular source. The region targeted by the bisulfite-specific primers contains four or five CpG dinucleotides, depending upon whether it is derived from the A or G

SNP allele respectively (Fig. 1). For each plasma donor, amplicons were cloned in a plasmid vector, and numbers of independent recombinant plasmids were DNA-sequenced. The presence of C residues in the DNA sequences indicated methylated C residues at each of the four or five CpG sites (Fig. 2).

The results for leukocyte cellular DNA obtained from two heterozygous donors demonstrated clear parent-of-origin-dependent methylation. For the leukocyte cellular DNA from donor 6, the A allele is hypomethylated, with the G allele hypermethylated; and *vice versa* for donor 7 (Fig. 2). Although overall the imprinting is distinct between the two alleles, as expected from previous work on methylation of this locus in other tissues there is a certain amount of variation in the methylation status of the two alleles, with a small proportion of methylated CpG sites in the presumed paternal allele and unmethylated sites in the presumed maternal allele. Combining data from both individuals, 19 maternal and 16 paternal alleles were recovered, which is consistent with the expectation that there is no bias in our experimental system for the bisulfite conversion, PCR, cloning and sequencing of the two alleles, i.e. that both alleles are recovered equally.

For the plasma DNA from three individual heterozygous donors, 1-3, the results also showed distinct imprinting, with clear parent-of-origin-dependent methylation. In each case the A allele was the hypomethylated allele, presumed to be paternal. Combining data for all three heterozygotes, 27 maternal and 24 paternal alleles were recovered, consistent with the hypothesis that there is equal representation of both alleles in cfp DNA. For such data, it could be argued that dividing the data into maternal (hypermethylated) and paternal (hypomethylated) classes on the basis of the

SNP ignores partially-methylated molecules, where presence of even a single methylated CpG site might be significant in determining abundance in cfp DNA. However, the variations from complete hyper- or hypomethylation in cfpDNA are similar to those seen in the cellular DNA, indicating that the population of molecules in cfp DNA does in fact reflect cellular DNA.

Fig. 2 also presents data for the cfp DNA from two SNP homozygotes. As expected the data indicate two distinct classes of DNA molecules on the basis of their methylation. Here the parent-of-origin distinction cannot be made on the basis of the SNP (both donors are G homozygotes), but categorising the molecules simply on the basis of whether the majority of CpG sites are methylated or unmethylated, we conclude that 22 maternal and 21 paternal alleles were recovered from these samples.

4. Discussion

Analysing an imprinted locus in the cell-free component of blood has provided, for the first time, a critical test for the steady-state concentrations of methylated and unmethylated cfp DNA. This test has the advantage that it is focused on a single DNA sequence (or two allelic sequences that differ only by an SNP) and is therefore independent of assumptions about equal representations of completely unrelated sequences in plasma.

We have found that the population of cfp DNA contains equal numbers of both maternal and paternal alleles, with their characteristic different methylation profiles, which correspond with the methylation profiles found in cellular DNA sources. We

can therefore conclude that methylation has no effect upon the representation of the *GNAS1* DNA molecules in plasma. We suggest that this is likely to be generally true for cfp DNA from other genomic regions, with the possible exception of heterochromatinised regions of the genome, such as the inactive X-chromosome in females, or Alu repeat sequences which are highly methylated, associated with a closed chromatin configuration, and present in increased proportions in serum when compared to unique gene sequences [21]. Such methylated regions may persist as nucleosomes in the circulation, perhaps protecting the DNA from degradation [22].

Although the production and clearance of cfp DNA from the circuation are not well understood processes [2], it is likely that our finding of equal representation in the steady state can be seen as indicating equal rates of clearance of the two allelic forms from the circulation. Our data indicate that interaction with cells of the immune system [23], and other potential factors affecting clearance, do not in fact lead to differential clearance of methylated and unmethylated DNA from the circulation in healthy adults. A consequence of this conclusion is that the quantity of a specific methylated (or unmethylated) cell-free DNA biomarker in plasma is a reliable indication of the release of the biomarker by a tumour for cancer studies, or by the placenta for minimally-invasive prenatal diagnosis of fetal aneuploidy.

References

[1] Tsang JCH, Lo YMD. Circulating nucleic acids in plasma/serum. Pathology 2007;39:197-207.

[2] Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. Clin Chim Acta 2001;313:139-42

[3] Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. Clin Chem 2002;48:421-7.

[4] Jahr S, Hentze H, Englisch S, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res 2001;61:1659-65.

[5] Sorenson GD, Pribish DM, Valone FH, Memoli VA, Bzik DJ, Yao SL. Soluble normal and mutated DNA sequences from single-copy genes in human blood. Cancer Epidemiol Biomarkers Prev 1994;3:67-71.

[6] Vasioukhin V, Anker P, Maurice P, Lyautey J, Lederrey C, Stroun M. Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. Br J Haematol 1994;86:774-9.

[7] Kawakami K, Brabender J, Lord RV, et al. Hypermethylated APC DNA in plasma and prognosis of patients with esophageal adenocarcinoma. J Natl Cancer Inst 2000;92:1805-11.

[8] Bianchi D. Circulating fetal DNA:its origin and potential diagnostic potential – a review. Placenta 2004;25(Suppl A): S93-S101.

[9] Alberry M, Maddocks D, Jones M, et al. Free fetal DNA in maternal plasma in anembryonic pregnancies: confirmation that the origin is the trophoblast. Prenat Diag 2007;27:415-8.

[10] Birch L, English CA, O'Donoghue K, Barigye O, Fisk NM, Keer JT. Accurate and robust quantification of circulating fetal and total DNA in maternal plasm from 5 to 41 weeks of gestation. Clin Chem 2005; 51:312-20.

[11] Honda H, Miharu N, Ohashi Y et al. Fetal gender determination in early pregnancy through qualitative and quantitative analysis of fetal DNA in maternal serum. Hum Genet 2002;110:75-9.

[12] Van der Schoot CE, Soussan AA, Koelewijn J, Bonsel G, Paget-Christiaens LG, de Haas M. Non-invasive antenatal RHD typing. Transfus Clin Biol 2006;13:53-7.

[13] Chim SS, Tong YK, Chiu RW, et al. Detection of the placental epigenetic signature of the maspin gene in maternal plasma. Proc Natl Acad Sci USA 2005;102:14753-8.

[14] Lo YM, Zhang J, Leung TN, Lau TK, Chang AM, Hjelm NM. Rapid clearance of fetal DNA from maternal plasma. Am J Hum Genet 1999;64:218-24.

[15] Bauer M, Heeg K, Wagner H, Lipford GB. DNA activates human immune cells through a CpG-dependent manner. Immunology 1999;97:699-705.

[16] Januchowski R, Prokop J, Jagodiński PP. Role of epigenetic DNA alterations in the pathogenesis of systematic lupus erythematosus. J Appl Genet 2004;45:237-48.

[17] Grunau C, Clark SJ, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nuc Acids Res 2001;29:e65.

[18] Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002;11:1427-31.

[19] Eggermann T, Meyer E, Schönherr N, Flick F, Chauvitré H, Mavany M. Mutation analysis of *GNAS1* and overlapping transcripts in Silver-Russel syndrome patients. Mol Genet Metabolism 2007;90:224-6.

[20] Reference SNP Cluster Report: rs1800905

URL http://www.ncbi.nlm.nih.gov/SNP/ snp_ref.cgi?rs=1800905.

[21] Stroun M, Lyautey J, Lederrey C, Mulcahy HE, Anker P. Alu repeat sequences are present in increased proportions compared to a unique gene in plasma/serum DNA: evidence for a preferential release from viable cells? Ann N Y Acad Sci 2001;945:258-64.

[22] Deligezer U, Yaman F, Erten N, Dalay N. Frequent copresence of methylated DNA and fragmented nucleosomal DNA in plasma of lymphoma patients. Clin Chim Acta. 2003;335:89-94.

[23] Anker P, Stroun M. Immunological aspects of circulating DNA. Ann N Y Acad Sci. 2006;1075:34-9.

Figure legends

Fig. 1

(A) Schematic of the complex *GNAS1* locus on chromosome 20. The patterns of primary transcripts produced from the maternal and paternal alleles are indicated above and below the indicated exonic structure of the locus [19]. Curved arrows represent transcription initiation sites. Maternal-specific and paternal-specific regions of differential DNA methylation are indicated by asterisks. The rs1800509 SNP is indicated within the intron separating exons XL and 1A.

(B) Nucleotide sequence of the DNA region surrounding the rs1800509 SNP. The single nucleotide polymorphism is indicated by R, representing the purines A or G at this site. CpG dinucleotides are indicated in large bold type. The region contains four CpG sites in the A allele. The G allele has an additional CpG site.

Fig 2. Methylation status of the four or five CpG sites surrounding rs1800509, in seven individuals. DNA was bisulfite converted, amplified by PCR, cloned in a plasmid vector, and individual plasmids sequenced. For each donor, 1-7, the methylation status of each CpG site in each sequenced DNA is indicated along a short horizontal line by an open (unmethylated) or closed (methylated) circle. The extra CpG site present in the G allele is indicated in parenthesis, with a gap at the corresponding position in the A allele. Samples of cell-free plasma DNA were obtained from donors 1-5. Leukocyte DNA was obtained from donors 6 and 7.
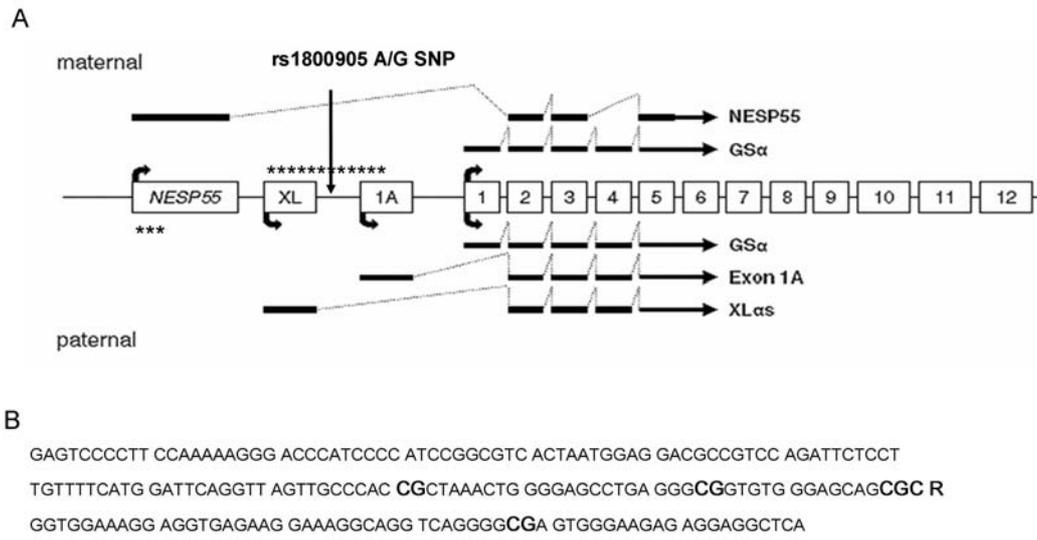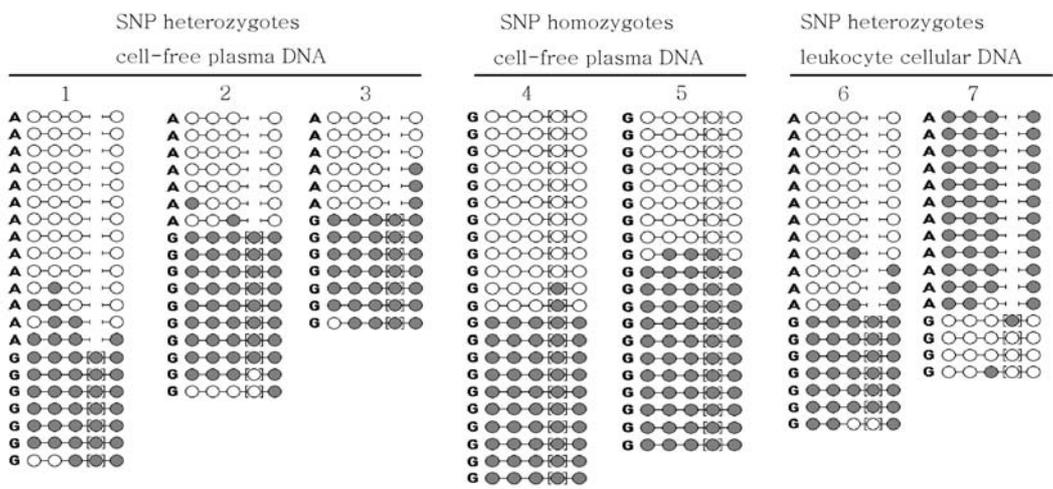
14

A

maternal

rs1800905 A/G SNP

NESP55
GSα

| NESP55 | XL | 1A | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |

***

GSα
Exon 1A
XLαs

paternal

B

GAGTCCCCTT CCAAAAAGGG ACCCATCCCC ATCCGGCGTC ACTAATGGAG GACGCCGTCC AGATTCTCCT
TGTTTTCATG GATTCAGGTT AGTTGCCCAC **CG**CTAAACTG GGGAGCCTGA GGG**CG**GTGTG GGAGCAG**CGC R**
GGTGGAAAGG AGGTGAGAAG GAAAGGCAGG TCAGGGG**CG**A GTGGGAAGAG AGGAGGCTCA

FIGURE 1

FIGURE 2