Noncoding RNA and DNA as Biomarkers: Toward an Epigenetic Fetal Barcode for Use in Maternal Plasma

The human genome contains a large layer of hidden biological information that is not accessible by proteomic or metabolic methods (1, 2). This information does not involve the typical (end)products of gene expression such as proteins. Instead, it involves genes that are transcribed but not translated (noncoding RNA), and DNA sequences that are neither transcribed nor translated (noncoding DNA) (1, 2). When this information, despite being noncoding in nature, is unique for the fetus or at least different from that in maternal blood cells (1), it can be used as a biomarker during pregnancy for the purpose of noninvasive prenatal diagnostics such as detection of Down syndrome. This application, along with a systematic approach to target this category of mostly unexplored information, is elegantly shown by Chim and coworkers in the current issue of Clinical Chemistry (2).

Why is the reporting of these findings so timely? According to the Fantom cDNA3 database (2005 release), the percentage of biological information presented by noncoding RNA is at least one-third of the more than 100,000 transcripts expressed in humans (3). Noncoding RNAs include the rapidly expanding family of small (21–35 nucleotides in length), noncoding microRNAs (miRNA) (4, 5), including small interfering RNA (siRNA), repeat-associated RNA (rasiRNA) (6), piwi interacting RNA (piRNA) (7), and mirtrons (8). The second category consists of noncoding DNA carrying differential DNA markers (2). When these signatures involve (chemical) DNA modifications, such as methylation of CpG dinucleotides, in the absence of sequence variations, they are called epigenetic. Epigenetic signatures that differ from those of fetal placenta cells and maternal blood cells can be used for noninvasive prenatal diagnostics (1). As for noncoding RNA, the type, number, and nature of noncoding DNA sequences has gained increased interest, with strong reappraisal of their biological significance.

The paper by Stephen Chim and coworkers in the current issue of Clinical Chemistry (2) is an excellent demonstration of how the hidden and largely unexplored class of biological information provided by noncoding DNA and carried by placental DNA can be retrieved. Retrieval was done systematically and reliably with near completeness for the chromosome region of interest (2). By exploring methylated DNA markers, which differ from those of placenta and maternal blood cells, only those biomarkers were targeted that allow reliable discrimination between fetal target DNA and maternal background DNA when used for prenatal diagnostics. By focusing on markers on the long arm of chromosome 21, which are directly related to and therefore informative for the disease of interest (trisomy 21, Down syndrome) (9, 10), and analyzing all suitable markers present (114 CpG islands), Chim et al. identified multiple chromosome 21-specific DNA markers. By a combined approach, executed thoroughly, a large and complete set (n = 22) of differentially methylated fetal- and chromosome 21q-specific DNA markers was identified. Two of these markers, both unmethylated in the placenta, U-PDE9A and U-CGI137, and selected for their greatest difference in median methylation index, were validated in maternal plasma samples. For the latter gene, the method of analysis was successfully adapted to overcome the low number of differentially methylated CpG sites. The results of this investigation clearly demonstrate that markers of this nature, in which noncoding sequences represent and complement the fetal genetic barcode, can be detected reliably and specifically in the maternal plasma (2).

The consequences of these findings are manifold. Although plasma is usually the primary specimen for clinical testing and contains the largest version of the human proteome (11), maternal plasma has not been useful in this way with respect to the completeness of the biological information from the fetus. At least 19.3% of additional data, not accessible by proteomic or metabolic methods, is available by molecular targeting for this class of hidden information (2). Epigenetic differences between the placenta and maternal blood cells were found on 22 of 114 CpG islands studied on chromosome 21q (2). The same gain in information can be expected for other chromosomes and for noncoding RNA, including microRNA. Recent studies show that of 345 miRNAs analyzed, 275 (80%) are expressed in the human placenta, and 53 (15%) show preferential or exclusive expression in the placenta (12). Therefore, the percentage of additional biological information on fetal well-being that becomes available by targeting noncoding RNA and noncoding DNA likely ranges between 20% and 30%. By analogy with the exploration of chromosome 21q (2), large-scale searches for epigenetic and related noncoding markers...
of other chromosomes will increase by 20%–30% the number and type of epigenetic biomarkers for the majority of pregnancy-associated and pregnancy-induced conditions. Such conditions include those for which there is a need for genetic testing (e.g., trisomy 13 and 18) and those associated with placental origin and placental dysfunction (such as preeclampsia). In preeclampsia, maternal transmission of the familial forms of preeclampsia appears related to epigenetic regulation of susceptibility genes (13), and thus disease-specific epigenetic signatures, besides being of diagnostic relevance, could be informative for etiological pathways involved in the condition.

The development of epigenetic tests does not mean that proteomic methods will become obsolete. For example, in addition to chemical modifications of CpG dinucleotides, epigenetic modifications of DNA also consist of specific protein sets bound to noncoding DNA, a characteristic that is the rule rather than an exception. For transcription factors that function as master control switches, the number of bound target sequences can be several thousands (14). Recently, high-throughput assays, such as the ChipSeq assay, combined with solid phase amplification and sequencing have become available to map protein-DNA interactions comprehensively across the mammalian genomes (14). By this approach, cell-specific interactome signatures can be identified in a cost-effective manner. Transcription factors important for placental function (GCM1) or involved in placental dysfunction (STOX1) are excellent candidates for this approach (13, 15) and should be considered in the context of prenatal diagnostics.

In conclusion, if the systematic approach followed by Chim and coworkers is applied in a genome-wide manner, or preceded by chromatin immunoprecipitation as well as extended to chromosome-specific, cell-specific or genome-wide explorations of noncoding RNA, this approach will certainly identify novel networks of noncoding DNA and RNA sequences. Such information will not only increase the number of novel fetal biomarkers for molecular diagnostics in pregnancy, but also provide mechanistic insight regarding essential placental processes that regulate and control pregnancy. Systematic explorations as described in this issue of Clinical Chemistry (2) are highly recommended and suitable for this purpose.

Grant/funding Support: Supported in part by the SAFE network (Project Number: LSHB-CT-2004-503243).

Financial Disclosures: None declared.

Cees B. M. Oudejans
Department of Clinical Chemistry
VU University Medical Center
Amsterdam, the Netherlands
Address correspondence to the author at:
Department of Clinical Chemistry
VU University Medical Center
De Boelelaan 1117
1081HV Amsterdam, the Netherlands
Fax 31-20-444 3985
E-mail cbm.oudejans@vumc.nl

References


DOI: 10.1373/clinchem.2007.100123