

HIGH EFFICIENCY GENE TRANSFER AND EXPRESSION IN MAMMALIAN CELLS BY A MULTIPLE TRANSFECTION PROCEDURE OF MAR SEQUENCES

FIELD OF THE INVENTION

5 The present invention relates to purified and isolated DNA sequences having protein production increasing activity and more specifically to the use of matrix attachment regions (MARs) for increasing protein production activity in a eukaryotic cell. Also disclosed is a method for the identification of said active regions, in particular MAR
10 nucleotide sequences, and the use of these characterized active MAR sequences in a new multiple transfection method.

BACKGROUND OF THE INVENTION

15 Nowadays, the model of loop domain organization of eukaryotic chromosomes is well accepted (Boulikas T, "Nature of DNA sequences at the attachment regions of genes to the nuclear matrix", *J. Cell Biochem.*, 52:14-22, 1993). According to this model chromatin is organized in loops that span 50-100 kb attached to the nuclear matrix, a proteinaceous network made up of RNPs and other nonhistone proteins (Bode J,
20 Stengert-Iber M, Kay V, Schalke T and Dietz-Pfeilstetter A, *Crit. Rev. Euk. Gene Exp.*, 6:115-138, 1996).

The DNA regions attached to the nuclear matrix are termed SAR or MAR for respectively scaffold (during metaphase) or matrix (interphase) attachment regions
25 (Hart C and Laemmli U (1998), "Facilitation of chromatin dynamics by SARs" *Curr Opin Genet Dev* 8, 519-525.)

As such, these regions may define boundaries of independent chromatin domains, such that only the encompassing cis-regulatory elements control the expression of the
30 genes within the domain.

However, their ability to fully shield a chromosomal locus from nearby chromatin elements, and thus confer position-independent gene expression, has not been seen in stably transfected cells (Poljak L, Seum C, Mattioni T and Laemmli U. (1994) "SARs stimulate but do not confer position independent gene expression", *Nucleic Acids Res*
35 22, 4386-4394). On the other hand, MAR (or S/MAR) sequences have been shown to interact with enhancers to increase local chromatin accessibility (Jenuwein T, Forrester W, Fernandez-Herrero L, Laible G, Dull M, and Grosschedl R. (1997) "Extension of chromatin accessibility by nuclear matrix attachment regions" *Nature* 385, 269-272).
40 Specifically, MAR elements can enhance expression of heterologous genes in cell culture lines (Kalos M and Fournier R (1995) "Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain" *Mol Cell Biol* 15,198-207), transgenic mice (Castilla J, Pintado B, Sola, I, Sanchez-Morgado J, and Enjuanes L (1998) "Engineering passive immunity in
45 transgenic mice secreting virus-neutralizing antibodies in milk" *Nat Biotechnol* 16, 349-354) and plants (Allen G, Hall GJ, Michalowski S, Newman W, Spiker S, Weissinger A, and Thompson W (1996), "High-level transgene expression in plant cells: effects of a strong scaffold attachment region from tobacco" *Plant Cell* 8, 899-913). The utility of MAR sequences for developing improved vectors for gene therapy is also recognized
50 (Agarwal M, Austin T, Morel F, Chen J, Bohnlein E, and Plavec I (1998), "Scaffold attachment region-mediated enhancement of retroviral vector expression in primary T

cells" *J Virol* 72, 3720-3728).

Recently, it has been shown that chromatin-structure modifying sequences including MARs, as exemplified by the chicken lysozyme 5' MAR is able to significantly enhance reporter expression in pools of stable Chinese Hamster Ovary (CHO) cells (Zahn-Zabal M, et al., "Development of stable cell lines for production or regulated expression using matrix attachment regions" *J Biotechnol*, 2001, 87(1): p. 29-42). This property was used to increase the proportion of high-producing clones, thus reducing the number of clones that need to be screened. These benefits have been observed both for constructs with MARs flanking the transgene expression cassette, as well as when constructs are co-transfected with the MAR on a separate plasmid. However, expression levels upon co-transfection with MARs were not as high as those observed for a construct in which two MARs delimit the transgene expression unit. A third and preferable process was shown to be the transfection of transgenes with MARs both linked to the transgene and on a separate plasmid (Girod et al., submitted for publication). However, one persisting limitation of this technique is the quantity of DNA that can be transfected per cell. Many multiples transfection protocols have been developed in order to achieve a high transfection efficiency to characterize the function of genes of interest. The protocol applied by Yamamoto et al, 1999 ("High efficiency gene transfer by multiple transfection protocol", *Histochem. J.* 31(4), 241-243) leads to a transfection efficiency of about 80 % after 5 transfections events, whereas the conventional transfection protocol only achieved a rate of <40%. While this technique may be useful when one wishes to increase the proportion of expressing cells, it does not lead to cells with a higher intrinsic productivity. Therefore, it cannot be used to generate high producer monoclonal cell lines. Hence, the previously described technique has two major drawbacks:

- i) this technique does not generate a homogenous population of transfected cells, since it cannot favour the integration of further gene copy, nor does it direct the transgenes to favorable chromosomal loci,
- ii) the use of the same selectable marker in multiple transfection events does not permit the selection of doubly or triply transfected cells.

In patent application WO02/074969, the utility of MARs for the development of stable eukaryotic cell lines has also been demonstrated. However, this application does not disclose neither any conserved homology for MAR DNA element nor any technique for predicting the ability for a DNA sequence to be a MAR sequence.

In fact no clear-cut MAR consensus sequence has been found (Boulikas T, "Nature of DNA sequences at the attachment regions of genes to the nuclear matrix", *J. Cell Biochem.*, 52:14-22, 1993) but evolutionarily, the structure of these sequences seem to be functionally conserved in eukaryotic genomes, since animal MARs can bind to plant nuclear scaffolds and vice versa (Mielke C, Kohwi Y, Kohwi-Shigematsu T and Bode J, "Hierarchical binding of DNA fragments derived from scaffold-attached regions: correlation of properties in vitro and function in vivo", *Biochemistry*, 29:7475-7485, 1990).

The identification of MARs by biochemical studies is a long and unpredictable process; various results can be obtained depending on the assay (Razin SV, "Functional architecture of chromosomal DNA domains", *Crit Rev Eukaryot Gene Expr.*, 6:247-269, 1996). Considering the huge number of expected MARs in a eukaryotic genome and the amount of sequences issued from genome projects, a tool able to filter potential MARs in order to perform targeted experiments would be greatly useful.

Currently two different predictive tools for MARs are available via the Internet.

The first one, MAR-Finder (<http://futuresoft.org/MarFinder>; Singh GB, Kramer JA and Krawetz SA, "Mathematical model to predict regions of chromatin attachment to the nuclear matrix", *Nucleic Acid Research*, 25:1419-1425, 1997) is based on set of

patterns identified within several MARs and a statistical analysis of the co-occurrence of these patterns. MAR-Finder predictions are dependent of the sequence context, meaning that predicted MARs depend on the context of the submitted sequence. The other predictive software, SMARTest (<http://www.genomatix.de>; Frisch M, Frech K, Klingenhoff A, Cartharius K, Liebich I and Werner T, "In silico prediction of scaffold/matrix attachment regions in large genomic sequences", *Genome Research*, 12:349-354, 2001), use weight-matrices derived from experimentally identified MARs. SMARTest is said to be suitable to perform large-scale analyses. But actually aside its relative poor specificity, the amount of hypothetical MARs rapidly gets huge when doing large scale analyses with it, and in having no way to increase its specificity to restrain the number of hypothetical MARs, SMARTest becomes almost useless to screen for potent MARs from large DNA sequences.

Some other softwares, not available via the Internet, also exists; they are based as well on the frequency of MAR motifs (MRS criterion; Van Drunen CM et al., "A bipartite sequence element associated with matrix/scaffold attachment regions", *Nucleic Acids Res*, 27:2924-2930, 1999), (ChrClass; Glazko GV et al., "Comparative study and prediction of DNA fragments associated with various elements of the nuclear matrix", *Biochim. Biophys. Acta*, 1517:351-356, 2001) or based on the identification of sites of stress-induced DNA duplex (SIDD; Benham C and al., "Stress-induced duplex DNA destabilization in scaffold/matrix attachment regions", *J. Mol. Biol.*, 274:181-196, 1997). However, their suitability to analyze complete genome sequences remains unknown, and whether these tools may allow the identification of protein production-increasing sequences has not been reported.

Furthermore, due to the relatively poor specificity of these softwares (Frisch M, Frech K, Klingenhoff A, Cartharius K, Liebich I and Werner T, "In silico prediction of scaffold/matrix attachment regions in large genomic sequences", *Genome Research*, 12:349-354, 2001), the amount of hypothetical MARs identified in genomes rapidly gets unmanageable when doing large scale analyses, especially if most of these have no or poor activity in practice. Thus, having no way to increase prediction specificity to restrain the number of hypothetical MARs, many of the available programs become almost useless to identify potent genetic elements in view of efficiently increasing recombinant protein production.

Since all the above available predictive methods have some drawbacks that prevent large-scale analyses of genomes to identify reliably novel and potent MARs, the object of this invention is to 1) understand the functional features of MARs that allow improved recombinant protein expression; 2) get a new Bioinformatic tool compiling MAR structural features as a prediction of function, in order to 3) perform large scale analyses of genomes to identify novel and more potent MARs, and, finally 4) to demonstrate improved efficiency to increase the production of recombinant proteins from eukaryotic cells or organisms when using the newly identified MAR sequences.

SUMMARY OF THE INVENTION

This object has been achieved by providing an improved and reliable method for the identification of DNA sequences having protein production increasing activity, in

particular MAR nucleotide sequences, and the use of these characterized active MAR sequences in a new multiple transfection method to increase the production of recombinant proteins in eukaryotic cells.

5

BRIEF DESCRIPTION OF THE FIGURES

10

Fig. 1 shows the distribution plots of MARs and non-MARs sequences. Histograms are density plots (relative frequency divided by the bin width) relative to the score of the observed parameter. The density histogram for human MARs in the SMART DB database is shown in black, while the density histogram for the human chromosome 22 are in grey.

15

Fig. 2 shows Scatterplots of the four different criteria used by SMAR Scan® and the AT-content with human MARs from SMART DB.

20

Fig. 3 shows the distribution plots of MAR sequences by organism. MAR sequences from SMART DB of other organisms were retrieved and analyzed. The MAR sequences density distributions for the mouse, the chicken, the sorghum bicolor and the human are plotted jointly.

25

Fig. 4 shows SMAR Scan® predictions on human chromosome 22 and on shuffled chromosome 22. Top plot : Average number of hits obtained by SMAR Scan® with five: rubbled, scrambled, shuffled within nonoverlapping windows of 10 bp, order 1 Markov chains model and with the native chromosome 22. Bottom plot: Average number of MARs predicted by SMAR Scan® in five: rubbled, scrambled, shuffled within non-overlapping windows of 10 bp, order 1 Markov chains model and with the native chromosome 22.

30

Fig. 5 shows the dissection of the ability of the chicken lysozyme gene 5'-MAR to stimulate transgene expression in CHO-DG44 cells. Fragments B, K and F show the highest ability to stimulate transgene expression. The indicated relative strength of the elements was based on the number of high-expressor cells.

35

Fig. 6 shows the effect of serial-deletions of the 5'-end (upper part) and the 3'-end (lower part) of the 5'-MAR on the loss of ability to stimulate transgene expression. The transition from increased to decreased activity coincide with B-, K- and F-fragments.

40

Fig. 7 shows that portions of the F fragment significantly stimulate transgene expression. The F fragment regions indicated by the light grey arrow were multimerized, inserted in pGEGFP Control and transfected in CHO cells. The element that displays the highest activity is located in the central part of the element and corresponds to fragment FIII (black bar labelled minimal MAR). In addition, an enhancer activity is located in the 3'-flanking part of the FIII fragment (dark grey bar labelled MAR enhancer).

45

50

Fig. 8 shows a map of locations for various DNA sequence motifs within the *cLysMAR*. Fig. 8 (B) represents a Map of locations for various DNA sequence motifs within the *cLysMAR*. Vertical lines represent the position of the computer-predicted sites or sequence motifs along the 3034 base pairs of the *cLysMAR* and its active regions, as presented in Fig. 5. The putative transcription factor sites, (MEF2 05, Oct-1, USF-02, GATA, NFAT) for activators and (CDP, SATB1, CTCF, ARBP/MeCP2) for repressors of transcription, were identified using MatInspector (Genomatix), and CpG islands were identified with CPGPLOT. Motifs previously associated with MAR elements are labelled

in black and include CpG dinucleotides and CpG islands, unwinding motifs (AATATATT and AATATT), poly As and Ts, poly Gs and Cs, *Drosophila* topoisomerase II binding sites (GTNWAYATTNATTNATNNR) which had identity to the 6 bp core and High mobility group I (HMG-I/Y) protein binding sites. Other structural motifs include

5 nucleosome-binding and nucleosome disfavoured sites and a motif thought to relieve the superhelical strain of DNA. Fig. 8(A) represents the comparison of the ability of portions of the cLysMAR to activate transcription with MAR prediction score profiles with MarFinder. The top diagram shows the MAR fragment activity as in Fig. 5, while the middle and bottom curves show MARFinder-predicted potential for MAR activity and

10 for bent DNA structures respectively.

Fig. 9 shows the correlation of DNA physico-chemical properties with MAR activity. Fig. 9(A), represents the DNA melting temperature, double helix bending, major groove depth and minor groove width profiles of the 5'-MAR and were determined using the

15 algorithms of Levitsky et al (Levitsky VG, Ponomarenko MP, Ponomarenko JV, Frolov AS, Kolchanov NA "Nucleosomal DNA property database", *Bioinformatics*, 15; 582592, 1999). The most active B, K and F fragments depicted at the top are as shown as in Figure 1. Fig. 9(B), represents the enlargement of the data presented in panel A to display the F fragment map aligned with the tracings corresponding to the melting

20 temperature (top curve) and DNA bending (bottom curve). The position of the most active FIB fragment and protein binding site for specific transcription factors are as indicated.

Fig. 10 shows the distribution of putative transcription factor binding sites within the 5'-cLysMAR. Large arrows indicate the position of the CUE elements as identified with

25 SMAR Scan®.

Fig. 11 shows the scheme of assembly of various portions of the MAR. The indicated portions of the cLysMAR were amplified by PCR, introducing BglII-BamHI linker elements at each extremity, and assembled to generate the depicted composite

30 elements. For instance, the top construct consists of the assembly of all CUE and flanking sequences at their original location except that BglII-BamHI linker sequences separate each element.

35 Fig. 12 represents the plasmid maps.

Fig. 13 shows the effect of re-transfecting primary transfectants on GFP expression. Cells (CHO-DG44) were co-transfected with pSV40EGFP (left tube) or pMAR-SV40EGFP (central tube) and pSVneo as resistance plasmid. Cells transfected with

40 pMAR-SV40EGFP were re-transfected 24 hours later with the same plasmid and a different selection plasmid, pSVpuro (right tube). After two weeks selection, the phenotype of the stably transfected cell population was analysed by FACS.

Fig. 14 shows the effect of multiple load of MAR-containing plasmid. The pMAR-SV40EGFP/ pMAR-SV40EGFP secondary transfectants were used in a third cycle of transfection at the end of the selection process. The tertiary transfection was

45 accomplished with pMAR or pMAR-SV40EGFP to give tertiary transfectants. After 24 hours, cells were transfected again with either plasmid, resulting in the quaternary transfectants (see Table 4).

Fig. 15 shows comparative performance of SMAR prediction algorithms exemplified by region WP18A10A7. (A) SMAR Scan® analysis was performed with default settings. (B) SIDD analysis (top curve and left-hand side scale), and the attachment of several

50

DNA fragments to the nuclear matrix in vitro (bar-graph, right-hand side scale) was taken from Goetze et al (Goetze S, Gluch A, Benham C, Bode J, "Computational and in vitro analysis of destabilized DNA regions in the interferon gene cluster: potential of predicting functional gene domains." *Biochemistry*, 42:154-166, 2003).

Fig. 16 represents the results of a gene therapy-like protocol using MARs. The group of mice injected by MAR-network, induced from the beginning of the experiment, display a better induction of the hematocrit in comparison of mice injected by original network without MAR. After 2 months, hematocrits in "MAR-containing group" is still at values higher (65%) than normal hematocrit levels (45-55%).

Fig. 17 represents the scatterplot for the 1757 S/MAR sequences of the AT (top) and TA (bottom) dinucleotide percentages versus the predicted DNA bending as computed by SMAR Scan®.

Fig. 18 represents the dinucleotide percentage distribution plots over the 1757 non-S/MARs sequences.

Fig.19 shows the effect of various S/MAR elements on the production of recombinant green fluorescent protein (GFP). Populations of CHO cells transfected with a GFP expression vector containing or a MAR element, as indicated, were analyzed by a fluorescence-activated cell sorter (FACS®), and typical profiles are shown. The profiles display the cell number counts as a function of the GFP fluorescence levels.

Fig. 20 depicts the effect of the induction of hematocrit in mice injected by MAR-network.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a purified and isolated DNA sequence having protein production increasing activity characterized in that said DNA sequence comprises at least one bent DNA element, and at least one binding site for a DNA binding protein.

Certain sequences of DNA are known to form a relatively "static curve", where the DNA follows a particular 3-dimensional path. Thus, instead of just being in the normal B-DNA conformation ("straight"), the piece of DNA can form a flat, planar curve also defined as bent DNA (Marini, *et al.*, 1982 "Bent helical structure in kinetoplast DNA", *Proc. Natl. Acad. Sci. USA*, 79: 7664-7664).

Surprisingly, Applicants have shown that the bent DNA element of a purified and isolated DNA sequence having protein production increasing activity of the present invention usually contains at least 10% of dinucleotide TA, and/or at least 12% of dinucleotide AT on a stretch of 100 contiguous base pairs. Preferably, the bent DNA element contains at least 33% of dinucleotide TA, and/or at least 33% of dinucleotide AT on a stretch of 100 contiguous base pairs. These data have been obtained by the method described further.

According to the present invention, the purified and isolated DNA sequence usually comprises a MAR nucleotide sequence selected from the group comprising the sequences SEQ ID Nos 1 to 27 or a cLysMAR element or a fragment thereof. Preferably, the purified and isolated DNA sequence is a MAR nucleotide sequence

selected from the group comprising the sequences SEQ ID Nos 1 to 27, more preferably the sequences SEQ ID Nos 24 to 27.

Encompassed by the present invention are as well complementary sequences of the above-mentioned sequences SEQ ID Nos 1 to 27 and the cLysMAR element or fragment, which can be produced by using PCR or other means.

An "element" is a conserved nucleotide sequences that bears common functional properties (i.e. binding sites for transcription factors) or structural (i.e. bent DNA sequence) features.

A part of sequences SEQ ID Nos 1 to 27 and the cLysMAR element or fragment refers to sequences sharing at least 70% nucleotides in length with the respective sequence of the SEQ ID Nos 1 to 27. These sequences can be used as long as they exhibit the same properties as the native sequence from which they derive. Preferably these sequences share more than 80%, in particular more than 90% nucleotides in length with the respective sequence of the SEQ ID Nos 1 to 27.

The present invention also includes variants of the aforementioned sequences SEQ ID Nos 1 to 27 and the cLysMAR element or fragment, that is nucleotide sequences that vary from the reference sequence by conservative nucleotide substitutions, whereby one or more nucleotides are substituted by another with same characteristics.

The sequences SEQ ID Nos 1 to 23 have been identified by scanning human chromosome 1 and 2 using SMAR Scan®, showing that the identification of novel MAR sequences is feasible using the tools reported thereafter whereas SEQ ID No 24 to 27 have been identified by scanning the complete human genome using the combined SMAR Scan® method.

In a first step, the complete chromosome 1 and 2 were screened to identify bent DNA element as region corresponding to the highest bent, major groove depth, minor groove width and lowest melting temperature as shown in figure 3. In a second step, this collection of sequence was scanned for binding sites of regulatory proteins such as SATB1, GATA, etc. as shown in the figure 8B) yielding sequences SEQ ID 1-23. Furthermore, sequences 21-23 were further shown to be located next to known gene from the Human Genome Data Base.

With regard to SEQ ID No 24 to 27 these sequences have been yielded by scanning the human genome according to the combined method and were selected as examples among 1757 MAR elements so detected.

Molecular chimera of MAR sequences are also considered in the present invention. By molecular chimera is intended a nucleotide sequence that may include a functional portion of a MAR element and that will be obtained by molecular biology methods known by those skilled in the art.

Particular combinations of MAR elements or fragments or sub-portions thereof are also considered in the present invention. These fragments can be prepared by a variety of methods known in the art. These methods include, but are not limited to, digestion with restriction enzymes and recovery of the fragments, chemical synthesis or polymerase chain reactions (PCR).

Therefore, particular combinations of elements or fragments of the sequences SEQ ID

Nos 1 to 27 and cLysMAR elements or fragments are also envisioned in the present invention, depending on the functional results to be obtained. Elements of the cLysMAR are e.g. the B, K and F regions as described in WO 02/074969, the disclosure of which is hereby incorporated herein by reference, in its entirety. The preferred elements of the cLysMAR used in the present invention are the B, K and F regions. Only one element might be used or multiple copies of the same or distinct elements (multimerized elements) might be used (see Fig. 8 A)).

By fragment is intended a portion of the respective nucleotide sequence. Fragments of a MAR nucleotide sequence may retain biological activity and hence bind to purified nuclear matrices and/or alter the expression patterns of coding sequences operably linked to a promoter. Fragments of a MAR nucleotide sequence may range from at least about 100 to 1000 bp, preferably from about 200 to 700 bp, more preferably from about 300 to 500 bp nucleotides. Also envisioned are any combinations of fragments, which have the same number of nucleotides present in a synthetic MAR sequence consisting of natural MAR element and/or fragments. The fragments are preferably assembled by linker sequences. Preferred linkers are BglII-BamHI linker.

"Protein production increasing activity" refers to an activity of the purified and isolated DNA sequence defined as follows: after having been introduced under suitable conditions into a eukaryotic host cell, the sequence is capable of increasing protein production levels in cell culture as compared to a culture of cell transfected without said DNA sequence. Usually the increase is 1.5 to 10 fold, preferably 4 to 10 fold. This corresponds to a production rate or a specific cellular productivity of at least 10 pg per cell per day (see Example 11 and Fig.13).

As used herein, the following definitions are supplied in order to facilitate the understanding of this invention.

"Chromatin" is the protein and nucleic acid material constituting the chromosomes of a eukaryotic cell, and refers to DNA, RNA and associated proteins.

A "chromatin element" means a nucleic acid sequence on a chromosome having the property to modify the chromatine structure when integrated into that chromosome.

"Cis" refers to the placement of two or more elements (such as chromatin elements) on the same nucleic acid molecule (such as the same vector, plasmid or chromosome).

"Trans" refers to the placement of two or more elements (such as chromatin elements) on two or more different nucleic acid molecules (such as on two vectors or two chromosomes).

Chromatin modifying elements that are potentially capable of overcoming position effects, and hence are of interest for the development of stable cell lines, include boundary elements (BEs), matrix attachment regions (MARs), locus control regions (LCRs), and universal chromatin opening elements (UCOE).

Boundary elements ("BEs"), or insulator elements, define boundaries in chromatin in many cases (Bell A and Felsenfeld G. 1999; "Stopped at the border: boundaries and insulators, *Curr Opin Genet Dev* 9, 191-198) and may play a role in defining a transcriptional domain in vivo. BEs lack intrinsic promoter/enhancer activity, but rather are thought to protect genes from the transcriptional influence of regulatory elements in the surrounding chromatin. The enhancer-block assay is commonly used to identify

insulator elements. In this assay, the chromatin element is placed between an enhancer and a promoter, and enhancer-activated transcription is measured. Boundary elements have been shown to be able to protect stably transfected reporter genes against position effects in *Drosophila*, yeast and in mammalian cells. They have also been shown to increase the proportion of transgenic mice with inducible transgene expression.

Locus control regions ("LCRs") are cis-regulatory elements required for the initial chromatin activation of a locus and subsequent gene transcription in their native locations (Grosveld, F. 1999, "Activation by locus control regions?" *Curr Opin Genet Dev* 9, 152-157). The activating function of LCRs also allows the expression of a coupled transgene in the appropriate tissue in transgenic mice, irrespective of the site of integration in the host genome. While LCRs generally confer tissue-specific levels of expression on linked genes, efficient expression in nearly all tissues in transgenic mice has been reported for a truncated human T-cell receptor LCR and a rat LAP LCR. The most extensively characterized LCR is that of the globin locus. Its use in vectors for the gene therapy of sickle cell disease and (3-thalassemias is currently being evaluated.

"MARs", according to a well-accepted model, may mediate the anchorage of specific DNA sequence to the nuclear matrix, generating chromatin loop domains that extend outwards from the heterochromatin cores. While MARs do not contain any obvious consensus or recognizable sequence, their most consistent feature appears to be an overall high A/T content, and C bases predominating on one strand (Bode J, Schlake T, RiosRamirez M, Mielke C, Stengart M, Kay V and KlehrWirth D, "Scaffold/matrix-attached regions: structural properties creating transcriptionally active loci", *Structural and Functional Organization of the Nuclear Matrix: International Review of Cytology*, 162A:389453, 1995). These regions have a propensity to form bent secondary structures that may be prone to strand separation. They are often referred to as base-unpairing regions (BURs), and they contain a core-unwinding element (CUE) that might represent the nucleation point of strand separation (Benham C and al., Stress induced duplex DNA destabilization in scaffold/matrix attachment regions, *J. Mol. Biol.*, 274:181-196, 1997). Several simple AT-rich sequence motifs have often been found within MAR sequences, but for the most part, their functional importance and potential mode of action remain unclear. These include the A-box (AATAAAYAAA), the T-box (TTWTWTTWTT), DNA unwinding motifs (AATATATT, AATATT), SATB1 binding sites (H-box, A/T/C25) and consensus Topoisomerase II sites for vertebrates (RNYNNCNNGYNGKTNyny) or *Drosophila* (GTNWAYATTNATNNR).

Ubiquitous chromatin opening elements ("UCOE", also known as "ubiquitously-acting chromatin opening elements") have been reported in WO 00/05393.

An "enhancer" is a nucleotide sequence that acts to potentiate the transcription of genes independent of the identity of the gene, the position of the sequence in relation to the gene, or the orientation of the sequence. The vectors of the present invention optionally include enhancers.

A "gene" is a deoxyribonucleotide (DNA) sequence coding for a given mature protein. As used herein, the term "gene" shall not include untranslated flanking regions such as RNA transcription initiation signals, polyadenylation addition sites, promoters or enhancers.

A "product gene" is a gene that encodes a protein product having desirable characteristics such as diagnostic or therapeutic utility. A product gene includes, e. g.,

structural genes and regulatory genes.

5 A "structural gene" refers to a gene that encodes a structural protein. Examples of structural genes include but are not limited to, cytoskeletal proteins, extracellular matrix proteins, enzymes, nuclear pore proteins and nuclear scaffold proteins, ion channels and transporters, contractile proteins, and chaperones. Preferred structural genes encode for antibodies or antibody fragments.

10 A "regulatory gene" refers to a gene that encodes a regulatory protein. Examples of regulatory proteins include, but are not limited to, transcription factors, hormones, growth factors, cytokines, signal transduction molecules, oncogenes, proto-oncogenes, transmembrane receptors, and protein kinases.

15 "Orientation" refers to the order of nucleotides in a given DNA sequence. For example, an inverted orientation of a DNA sequence is one in which the 5' to 3' order of the sequence in relation to another sequence is reversed when compared to a point of reference in the DNA from which the sequence was obtained. Such reference points can include the direction of transcription of other specified DNA sequences in the source DNA and/or the origin of replication of replicable vectors containing the
20 sequence.

"Eukaryotic cell" refers to any mammalian or non-mammalian cell from a eukaryotic organism. By way of non-limiting example, any eukaryotic cell that is capable of being maintained under cell culture conditions and subsequently transfected would be
25 included in this invention. Especially preferable cell types include, e. g., stem cells, embryonic stem cells, Chinese hamster ovary cells (CHO), COS, BHK21, NIH3T3, HeLa, C2C12, cancer cells, and primary differentiated or undifferentiated cells. Other suitable host cells are known to those skilled in the art.

30 The terms "host cell" and "recombinant host cell" are used interchangeably herein to indicate a eukaryotic cell into which one or more vectors of the invention have been introduced. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or
35 environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

40 The terms "introducing a purified DNA into a eukaryotic host cell" or "transfection" denote any process wherein an extracellular DNA, with or without accompanying material, enters a host cell. The term "cell transfected" or "transfected cell" means the cell into which the extracellular DNA has been introduced and thus harbours the extracellular DNA. The DNA might be introduced into the cell so that the nucleic acid is replicable either as a chromosomal integrant or as an extra chromosomal element.

45 "Promoter" as used herein refers to a nucleic acid sequence that regulates expression of a gene.

50 "Co-transfection" means the process of transfecting a eukaryotic cell with more than one exogenous gene, or vector, or plasmid, foreign to the cell, one of which may confer a selectable phenotype on the cell.

The purified and isolated DNA sequence having protein production increasing activity also comprises, besides one or more bent DNA element, at least one binding site for a DNA binding protein.

- 5 Usually the DNA binding protein is a transcription factor. Examples of transcription factors are the group comprising the polyQpolyP domain proteins.
Another example of a transcription factor is a transcription factor selected from the group comprising SATB1, NMP4, MEF2, S8, DLX1, FREAC7, BRN2, GATA 1/3, TATA, Bright, MSX, AP1, C/EBP, CREBP1, FOX, Freac7, HFH1, HNF3alpha, Nkx25,
10 POU3F2, Pit1, TTF1, XFD1, AR, C/EBPgamma, Cdc5, FOXD3, HFH3, HNF3 beta, MRF2, Oct1, POU6F1, SRF, V\$MTATA_B, XFD2, Bach2, CDP CR3, Cdx2, FOXJ2, HFL, HP1, Myc, PBX, Pax3, TEF, VBP, XFD3, Brn2, COMP1, Evl, FOXP3, GATA4, HFN1, Lhx3, NKX3A, POU1F1, Pax6, TFIIA or a combination of two or more of these transcription factors are preferred. Most preferred are SATB1, NMP4, MEF2 and
15 polyQpolyP domain proteins.

SATB1, NMP4 and MEF2, for example, are known to regulate the development and/or tissue-specific gene expression in mammals. These transcription factors have the capacity to alter DNA geometry, and reciprocally, binding to DNA as an allosteric ligand
20 modifies their structure. Recently, SATB1 was found to form a cage-like structure circumscribing heterochromatin (Cai S, Han HJ, and Kohwi-Shigematsu T, "Tissue-specific nuclear architecture and gene expression regulated by SATB1" *Nat Genet*, 2003. 34(1): p. 42-51).

25 Yet another object of the present invention is to provide a purified and isolated cLysMAR element and/or fragment, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

30 More preferably, the cLysMAR element and/or fragment are consisting of at least one nucleotide sequence selected from the B, K and F regions.

A further object of the present invention is to provide a synthetic MAR sequence comprising natural MAR element and/or fragments assembled between linker
35 sequences.

Preferably, the synthetic MAR sequence comprises a cLysMAR element and/or fragment a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.
40 Also preferably, linker sequences are BglII-BamHI linker.

An other aspect of the invention is to provide a method for identifying a MAR sequence using a Bioinformatic tool comprising the computing of values of one or more DNA sequence features corresponding to DNA bending, major groove depth and minor
45 groove width potentials and melting temperature. Preferably, the identification of one or more DNA sequence features further comprises a further DNA sequence feature corresponding to binding sites for DNA binding proteins, which is also computed with this method.

50 Preferably, profiles or weight-matrices of said bioinformatic tool are based on dinucleotide recognition.

The bioinformatic tool used for the present method is preferably, SMAR Scan®, which contains algorithms developed by Gene Express ([http://srs6.bionet.nsc.ru/srs6bin/cgi-bin/wgetz?-e+\[FEATURES-SiteID:'nR'\]](http://srs6.bionet.nsc.ru/srs6bin/cgi-bin/wgetz?-e+[FEATURES-SiteID:'nR'])) and based on Levitsky *et al.*, 1999. These algorithms recognise profiles, based on dinucleotides weight-matrices, to compute the theoretical values for conformational and physicochemical properties of DNA.

Preferably, SMAR Scan® uses the four theoretical criteria also designated as DNA sequence features corresponding to DNA bending, major groove depth and minor groove width potentials, melting temperature in all possible combination, using scanning windows of variable size (see Fig. 3). For each function used, a cut-off value has to be set. The program returns a hit every time the computed score of a given region is above the set cut-off value for all of the chosen criteria. Two data output modes are available to handle the hits, the first (called "profile-like") simply returns all hit positions on the query sequence and their corresponding values for the different criteria chosen. The second mode (called "contiguous hits") returns only the positions of several contiguous hits and their corresponding sequence. For this mode, the minimum number of contiguous hits is another cut-off value that can be set, again with a tunable window size. This second mode is the default mode of SMAR Scan®. Indeed, from a semantic point of view, a hit is considered as a core-unwinding element (CUE), and a cluster of CUEs accompanied by clusters of binding sites for relevant proteins is considered as a MAR. Thus, SMAR Scan® considers only several contiguous hits as a potential MAR.

To tune the default cut-off values for the four theoretical structural criteria, experimentally validated MARs from SMARt DB (<http://transfac.gbf.de/-SMARtDB>) were used. All the human MAR sequences from the database were retrieved and analyzed with SMAR Scan® using the "profile-like" mode with the four criteria and with no set cut-off value. This allowed the setting of each function for every position of the sequences. The distribution for each criterion was then computed according to these data (see Fig. 1 and 3).

The default cut-off values of SMAR Scan® for the bend, the major groove depth and the minor groove width were set at the average of the 75th quantile and the median. For the melting temperature, the default cut-off value should be set at the 75th quantile. The minimum length for the "contiguous-hits" mode should be set to 300 because it is assumed to be the minimum length of a MAR (see Fig. 8 and 9). However, one skilled in the art would be able to determine the cut-off values for the above-mentioned criteria for a given organism with minimal experimentation.

Preferably, DNA bending values are comprised between 3 to 5 ° (radial degree). Most preferably they are situated between 3.8 to 4.4 °, corresponding to the smallest peak of Fig. 1.

Preferably the major groove depth values are comprised between 8.9 to 9.3 Å (Angström) and minor groove width values between 5.2 to 5.8 Å. Most preferably the major groove depth values are comprised between 9.0 to 9.2 Å and minor groove width values between 5.4 to 5.7 Å.

Preferably the melting temperature is comprised between 55 to 75 ° C (Celsius degree). Most preferably, the melting temperature is comprised between 55 to 62 ° C.

The DNA binding protein of which values can be computed by the method is usually a transcription factor preferably a polyQpolyP domain or a transcription factor selected

from the group comprising SATB1, NMP4, MEF2, S8, DLX1, FREAC7, BRN2, GATA 1/3, TATA, Bright, MSX, AP1, C/EBP, CREBP1, FOX, Freac7, HFH1, HNF3alpha, Nkx25, POU3F2, Pit1, TTF1, XFD1, AR, C/EBPgamma, Cdc5, FOXD3, HFH3, HNF3 beta, MRF2, Oct1, POU6F1, SRF, V\$MTATA_B, XFD2, Bach2, CDP CR3, Cdx2, FOXJ2, HFL, HP1, Myc, PBX, Pax3, TEF, VBP, XFD3, Brn2, COMP1, Evl, FOXP3, GATA4, HFN1, Lhx3, NKX3A, POU1F1, Pax6, TFIIA or a combination of two or more of these transcription factors.

However, one skilled in the art would be able to determine other kinds of transcription factors in order to carry out the method according to the present invention.

In case SMAR Scan® is envisaged to perform, for example, large scale analysis, then, preferably, the above-mentioned method further comprises at least one filter predicting DNA binding sites for DNA transcription factors in order to reduce the computation.

The principle of this method combines SMAR Scan® to compute the structural features as described above and a filter, such as for example, the pfsearch, (from the pftools package as described in Bucher P, Karplus K, Moeri N, and Hofmann K, "A flexible search technique based on generalized profiles", *Computers and Chemistry*, 20:324, 1996) to predict the binding of some transcription factors.

Examples of filters comprise, but are not limited to, pfsearch, MatInspector, RMatch Professional and TRANSFAC Professional

This combined method uses the structural features of SMAR Scan® and the predicted binding of specific transcription factors of the filter that can be applied sequentially in any order to select MARs, therefore, depending on the filter is applied at the beginning or at the end of the method.

The first level selects sequences out of the primary input sequence and the second level, consisting in the filter, may be used to restrain among the selected sequences those which satisfy the criteria used by the filter.

In this combined method the filter detects clusters of DNA binding sites using profiles or weightmatrices from, for example, MatInspector (Quandt K, Frech K, Karas H, Wingender E, Werner T, "MatInd and MatInspector New fast and versatile tools for detection of consensus matches in nucleotide sequence data", *Nucleic Acids Research*, 23, 48784884, 1995.). The filter can also detect densities of clusters of DNA binding sites.

The combined method is actually a "wrapper" written in Perl for SMAR Scan® and, in case the pfsearch is used as a filter, from the pftools. The combined method performs a twolevel processing using at each level one of these tools (SMAR Scan® or filter) as a potential "filter", each filter being optional and possible to be used to compute the predicted features without doing any filtering.

If SMAR Scan® is used in the first level to filter subsequences, it has to be used with the "all the contiguous hits" mode in order to return sequences. If the pfsearch is used in the first level as first filter, it has to be used with only one profile and a distance in nucleotide needs to be provided. This distance is used to group together pfsearch hits that are located at a distance inferior to the distance provided in order to return sequences; The combined method launches pfsearch, parses its output and returns

sequences corresponding to pfsearch hits that are grouped together according to the distance provided. Then whatever the tool used in the first level, the length of the sub-sequences thus selected can be systematically extended at both ends according to a parameter called "hits extension".

The second and optional level can be used to filter out sequences (already filtered sequences or unfiltered input sequences) or to get the results of SMAR Scan® and/or pfsearch without doing any filtering on these sequences. If the second level of combined method is used to filter, for each criteria considered cutoff values (hit per nucleotide) need to be provided to filter out those sequences (see Fig. 20).

Another concern of the present invention is also to provide a method for identifying a MAR sequence comprising at least one filter detecting clusters of DNA binding sites using profiles or weightmatrices. Preferably, this method comprises two levels of filters and in this case, SMAR Scan® is totally absent from said method. Usually, the two levels consist in pfsearch.

Also embraced by the present invention is a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the combined method or the method comprising at least one filter.

Analysis by the combined method of the whole human genome yielded a total of 1757 putative MARs representing a total of 1 065 305 base paires. In order to reduce the number of results, a dinucleotide analysis was performed on these 1757 MARs, computing each of the 16 possible dinucleotide percentage for each sequence considering both strands in the 5' to 3' direction.

Surprisingly, Applicants have shown that all of the "super" MARs detected with the combined method contain at least 10% of dinucleotide TA on a stretch of 100 contiguous base pairs. Preferably, these sequences contain at least 33% of dinucleotide TA on a stretch of 100 contiguous base pairs.

Applicants have also shown that these same sequences further contain at least 12% of dinucleotide AT on a stretch of 100 contiguous base pairs. Preferably, they contain at least 33% of dinucleotide AT on a stretch of 100 contiguous base pairs.

An other aspect of the invention is to provide a purified and isolated MAR DNA sequence of any of the preceding described MARs, comprising a sequence selected from the sequences SEQ ID Nos 1 to 27, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

Preferably, said purified and isolated MAR DNA sequence comprises a sequence selected from the sequences SEQ ID Nos 24 to 27, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants. These sequences 24 to 27 correspond to those detected by the combined method and show a higher protein production increasing activity over sequences 1 to 23.

The present invention also encompasses the use of a purified and isolated DNA sequence comprising a first isolated matrix attachment region (MAR) nucleotide sequence which is a MAR nucleotide sequence selected from the group comprising

- a purified and isolated DNA sequence having protein production increasing activity,
 - a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the combined method or the method comprising at least one filter,
 - the sequences SEQ ID Nos 1 to 27,
 - a purified and isolated cLysMAR element and/or fragment,
 - a synthetic MAR sequence comprising natural MAR element and/or fragments assembled between linker sequences,
- a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants or a MAR nucleotide sequence of a cLysMAR element and/or fragment, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants for increasing protein production activity in a eukaryotic host cell.

Said purified and isolated DNA sequence usually further comprises one or more regulatory sequences, as known in the art e.g. a promoter and/or an enhancer, polyadenylation sites and splice junctions usually employed for the expression of the protein or may optionally encode a selectable marker. Preferably said purified and isolated DNA sequence comprises a promoter which is operably linked to a gene of interest.

The DNA sequences of this invention can be isolated according to standard PCR protocols and methods well known in the art.

Promoters which can be used provided that such promoters are compatible with the host cell are, for example, promoters obtained from the genomes of viruses such as polyoma virus, adenovirus (such as Adenovirus 2), papilloma virus (such as bovine papilloma virus), avian sarcoma virus, cytomegalovirus (such as murine or human cytomegalovirus immediate early promoter), a retrovirus, hepatitis-B virus, and Simian Virus 40 (such as SV 40 early and late promoters) or promoters obtained from heterologous mammalian promoters, such as the actin promoter or an immunoglobulin promoter or heat shock promoters. Such regulatory sequences direct constitutive expression.

Furthermore, the purified and isolated DNA sequence might further comprise regulatory sequences which are capable of directing expression of the nucleic acid preferentially in a particular cell type (e. g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, et al., 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and Baltimore, 1989. *EMBOJ.* 8: 729-733) and immunoglobulins (Banerji, et al., 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33:741-748), neuron-specific promoters (e. g., the neurofilament promoter; Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, et al., 1985. *Science* 230: 912-916), and mammary gland-specific promoters (e. g., milk whey promoter; U. S. Pat. No. 4,873,316 and European Application No. 264,166).

Developmentally-regulated promoters are also encompassed. Examples of such promoters include, e.g., the murine hox promoters (Kessel and Gruss, 1990. *Science* 249: 374-379) and thea-fetoprotein promoter (Campes and Tilghman, 1989. *Genes*

Dev. 3: 537-546).

Regulatable gene expression promoters are well known in the art, and include, by way of non-limiting example, any promoter that modulates expression of a gene encoding a desired protein by binding an exogenous molecule, such as the CRE/LOX system, the TET system, the doxycycline system, the NFkappaB/UV light system, the Leu3p/isopropylmalate system, and the GLVPc/GAL4 system (See e. g., Sauer, 1998, Methods 14 (4): 381-92 ; Lewandoski, 2001, Nat. Rev. Genet 2 (10): 743-55; Legrand-Poels et al., 1998, J. Photochem. Photobiol. B. 45: 18; Guo et al., 1996, FEBS Lett. 390 (2): 191-5; Wang et al., PNAS USA, 1999, 96 (15): 84838). However, one skilled in the art would be able to determine other kinds of promoters that are suitable in carrying out the present invention.

Enhancers can be optionally included in the purified DNA sequence of the invention then belonging to the regulatory sequence, e.g. the promoter.

The "gene of interest" or "transgene" preferably encodes a protein (structural or regulatory protein). As used herein "protein" refers generally to peptides and polypeptides having more than about ten amino acids. The proteins may be "homologous" to the host (i.e., endogenous to the host cell being utilized), or "heterologous," (i.e., foreign to the host cell being utilized), such as a human protein produced by yeast. The protein may be produced as an insoluble aggregate or as a soluble protein in the periplasmic space or cytoplasm of the cell, or in the extracellular medium. Examples of proteins include hormones such as growth hormone or erythropoietin (EPO), growth factors such as epidermal growth factor, analgesic substances like enkephalin, enzymes like chymotrypsin, receptors to hormones or growth factors, antibodies and include as well proteins usually used as a visualizing marker e.g. green fluorescent protein.

Preferably the purified DNA sequence further comprises at least a second isolated matrix attachment region (MAR) nucleotide sequence selected from the group comprising

- a purified and isolated DNA sequence having protein production increasing activity,
- a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the combined method or the method comprising at least one filter,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
- a synthetic MAR sequence comprising natural MAR element and/or fragments assembled between linker sequences,

a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants. The isolated matrix attachment region (MAR) nucleotide sequence might be identical or different.

Alternatively, a first and a second identical MAR nucleotide sequence are used.

Preferably, the MAR nucleotide sequences are located at both the 5' and the 3' ends of the sequence containing the promoter and the gene of interest. But the invention also envisions the fact that said first and or at least second MAR nucleotide sequences are located on a sequence distinct from the one containing the promoter and the gene of interest.

Embraced by the scope of the present invention is also the purified and isolated DNA sequence comprising a first isolated matrix attachment region (MAR) nucleotide sequence which is a MAR nucleotide sequence selected from the group comprising

- a purified and isolated DNA sequence having protein production increasing activity,
- a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the combined method or the method comprising at least one filter,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
- a synthetic MAR sequence comprising natural MAR element and/or fragments assembled between linker sequences,

a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants that can be used for increasing protein production activity in a eukaryotic host cell by introducing the purified and isolated DNA sequence into a eukaryotic host cell according to well known protocols. Usually applied methods for introducing DNA into eukaryotic host cells applied are e.g. direct introduction of cloned DNA by microinjection or microparticle bombardment; electrotransfer ;use of viral vectors; encapsulation within a carrier system; and use of transfecting reagents such as calcium phosphate, diethylaminoethyl (DEAE) –dextran or commercial transfection systems like the Lipofect-AMINE 2000 (Invitrogen). Preferably, the transfection method used to introduce the purified DNA sequence into a eukaryotic host cell is the method for transfecting a eukaryotic cell as described below.

The purified and isolated DNA sequence can be used in the form of a circular vector. Preferably, the purified and isolated DNA sequence is used in the form of a linear DNA sequence as vector.

As used herein, "plasmid" and "vector" are used interchangeably, as the plasmid is the most commonly used vector form. However, the invention is intended to include such other forms of expression vectors, including, but not limited to, viral vectors (e. g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The present invention further encompasses a method for transfecting a eukaryotic host cell, said method comprising

- a) introducing into said eukaryotic host cell at least one purified DNA sequence comprising at least one DNA sequence of interest and/or at least one purified and isolated DNA sequence comprising a MAR nucleotide sequence or other chromatin modifying elements,
- b) subjecting within a defined time said transfected eukaryotic host cell to at least one additional transfection step with at least one purified DNA sequence comprising at least one DNA sequence of interest and/or with at least one purified and isolated DNA sequence comprising a MAR nucleotide sequence or other chromatin modifying elements
- c) selecting said transfected eukaryotic host cell.

Preferably at least two up to four transfecting steps are applied in step b).

In order to select the successful transfected cells, a gene that encodes a selectable marker (e. g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. The gene that encodes a selectable marker might be located

on the purified DNA sequence comprising at least one DNA sequence of interest and/or at least one purified and isolated DNA sequence consisting of a MAR nucleotide sequence or other chromatin modifying elements or might optionally be co-introduced in separate form e.g. on a plasmid. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. The amount of the drug can be adapted as desired in order to increase productivity

Usually, one or more selectable markers are used. Preferably, the selectable markers used in each distinct transfection steps are different. This allows selecting the transformed cells that are "multi-transformed" by using for example two different antibiotic selections.

Any eukaryotic host cell capable of protein production and lacking a cell wall can be used in the methods of the invention. Examples of useful mammalian host cell lines include human cells such as human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol 36, 59 (1977)), human cervical carcinoma cells (HELA, ATCC CCL 2), human lung cells (W138, ATCC CCL 75), human liver cells (Hep G2, HB 8065); rodent cells such as baby hamster kidney cells (BHK, ATCC CCL 10), Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77, 4216 (1980)), mouse sertoli cells (TM4, Mather, *Biol. Reprod* 23, 243-251 (1980)), mouse mammary tumor (MMT 060562, ATCC CCL51); and cells from other mammals such as monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); myeloma (e.g. NS0) /hybridoma cells.

Preferably, the selected transfected eukaryotic host cells are high protein producer cells with a production rate of at least 10 pg per cell per day.

Most preferred for uses herein are mammalian cells, more preferred are CHO cells.

The DNA sequence of interest of the purified and isolated DNA sequence is usually a gene of interest preferably encoding a protein operably linked to a promoter as described above. The purified and isolated DNA sequence comprising at least one DNA sequence of interest might comprise additionally to the DNA sequence of interest MAR nucleotide sequence or other chromatin modifying elements.

Purified and isolated DNA sequence comprising a MAR nucleotide sequence are for example selected from the group comprising the sequences SEQ ID Nos 1 to 27 and/or particular elements of the cLysMAR e.g. the B, K and F regions as well as fragment and elements and combinations thereof as described above. Other chromatin modifying elements are for example boundary elements (BEs), locus control regions (LCRs), and universal chromatin opening elements (UCOEs) (see Zahn-Zabal et al. already cited). An example of multiple transfections of host cells is shown in Example 12 (Table 3). The first transfecting step (primary transfection) is carried out with the gene of interest (SV40EGFP) alone, with a MAR nucleotide sequence (MAR) alone or with the gene of interest and a MAR nucleotide sequence (MAR-SV40EGFP). The second transfecting step (secondary transfection) is carried out with the gene of interest (SV40EGFP) alone, with a MAR nucleotide sequence (MAR) alone or with the gene of interest and a MAR nucleotide sequence (MAR-SV40EGFP), in all possible combinations resulting from the first transfecting step.

Preferably the eukaryotic host cell is transfected by:

- a) introducing a purified DNA sequence comprising one DNA sequence of interest and additionally a MAR nucleotide sequence,
- b) subjecting within a defined time said transfected eukaryotic host cell to at least one additional transfection step with the same purified DNA sequence comprising one DNA sequence of interest and additionally a MAR nucleotide sequence of step a).

Also preferably, the MAR nucleotide sequence of the of the purified and isolated DNA sequence is selected from the group comprising

- a purified and isolated DNA sequence having protein production increasing activity,
- a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the combined method or the method comprising at least one filter,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
- a synthetic MAR sequence comprising natural MAR element and/or fragments assembled between linker sequences,

a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

Surprisingly, a synergy between the first and second transfection has been observed. A particular synergy has been observed when MAR elements are present at one or both of the transfection steps. Multiple transfections of the cells with pMAR alone or in combination with various expression plasmids, using the method described above have been carried out. For example, Table 3 shows that transfecting the cells twice with the pMAR-SV40EGFP plasmid gave the highest expression of GFP and the highest degree of enhancement of all conditions (4.3 fold). In contrast, transfecting twice the vector without MAR gave little or no enhancement, 2.8-fold, instead of the expected two-fold increase. This proves that the presence of MAR elements at each transfection step is of particular interest to achieve the maximal protein synthesis.

As a particular example of the transfection method, said purified DNA sequence comprising at least one DNA sequence of interest can be introduced in form of multiple unlinked plasmids, comprising a gene of interest operably linked to a promoter, a selectable marker gene, and/or protein production increasing elements such as MAR sequences.

The ratio of the first and subsequent DNA sequences may be adapted as required for the use of specific cell types, and is routine experimentation to one ordinary skilled in the art.

The defined time for additional transformations of the primary transformed cells is tightly dependent on the cell cycle and on its duration. Usually the defined time corresponds to intervals related to the cell division cycle.

Therefore this precise timing may be adapted as required for the use of specific cell types, and is routine experimentation to one ordinary skilled in the art.

Preferably the defined time is the moment the host cell just has entered into the same phase of a second or a further cell division cycle, preferably the second cycle.

This time is usually situated between 6h and 48 h, preferably between 20h and 24h after the previous transfecting event.

Also encompassed by the present invention is a method for transfecting a eukaryotic host cell, said method comprising co-transfecting into said eukaryotic host cell at least one first purified and isolated DNA sequence comprising at least one DNA sequence of

interest, and a second purified DNA comprising at least one MAR nucleotide selected from the group comprising:

- a purified and isolated DNA sequence having protein production increasing activity,
- a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the combined method or the method comprising at least one filter,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
- a synthetic MAR sequence comprising natural MAR element and/or fragments assembled between linker sequences,

a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

Said first purified and isolated DNA sequence can also comprise at least one MAR nucleotide as described above.

Also envisioned is a process for the production of a protein wherein a eukaryotic host cell is transfected according to the transfection methods as defined in the present invention and is cultured in a culture medium under conditions suitable for expression of the protein. Said protein is finally recovered according to any recovering process known to the skilled in the art.

Given as an example, the following process for protein production might be used.

The eukaryotic host cell transfected with the transfection method of the present invention is used in a process for the production of a protein by culturing said cell under conditions suitable for expression of said protein and recovering said protein. Suitable culture conditions are those conventionally used for in vitro cultivation of eukaryotic cells as described e.g. in WO 96/39488. The protein can be isolated from the cell culture by conventional separation techniques such as e.g. fractionation on immunoaffinity or ion-exchange columns; precipitation; reverse phase HPLC; chromatography; chromatofocusing; SDS-PAGE; gel filtration. One skilled in the art will appreciate that purification methods suitable for the polypeptide of interest may require modification to account for changes in the character of the polypeptide upon expression in recombinant cell culture.

The proteins that are produced according to this invention can be tested for functionality by a variety of methods. For example, the presence of antigenic epitopes and ability of the proteins to bind ligands can be determined by Western blot assays, fluorescence cell sorting assays, immunoprecipitation, immunochemical assays and/or competitive binding assays, as well as any other assay which measures specific binding activity.

The proteins of this invention can be used in a number of practical applications including, but not limited to:

1. Immunization with recombinant host protein antigen as a viral/pathogen antagonist.
2. Production of membrane proteins for diagnostic or screening assays.
3. Production of membrane proteins for biochemical studies.
4. Production of membrane protein for structural studies.
5. Antigen production for generation of antibodies for immuno-histochemical mapping, including mapping of orphan receptors and ion channels.

Also provided by the present invention is a eukaryotic host cell transfected according to any of the preceding transfection methods. Preferably, the eukaryotic host cell is a mammalian host cell line.

As already described, example of useful mammalian host cell lines include human cells such as human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol 36, 59 (1977)), human cervical carcinoma cells (HELA, ATCC CCL 2), human lung cells (W138, ATCC CCL 75),
 5 human liver cells (Hep G2, HB 8065); rodent cells such as baby hamster kidney cells (BHK, ATCC CCL 10), Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77, 4216 (1980)), mouse sertoli cells (TM4, Mather, *Biol. Reprod* 23, 243-251 (1980)), mouse mammary tumor (MMT 060562, ATCC CCL51);
 10 and cells from other mammals such as monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); myeloma (e.g. NS0) /hybridoma cells.

Most preferred for uses herein are CHO cells.

15 The present invention also provides for a cell transfection mixture or Kit comprising at least one purified and isolated DNA sequence according to the invention.

20 The invention further comprises a transgenic organism wherein at least some of its cells have stably incorporated at least one DNA sequence of

- a purified and isolated DNA sequence having protein production increasing activity,
- a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the
 25 combined method or the method comprising at least one filter,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
- a synthetic MAR sequence comprising natural MAR element and/or fragments assembled between linker sequences,

30 a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

Preferably, some of the cells of the transgenic organisms have been transfected according the methods described herein.

35 Also envisioned in the present invention is a transgenic organism wherein its genome has stably incorporated at least one DNA sequence of

- a purified and isolated DNA sequence having protein production increasing activity,
- a purified and isolated MAR DNA sequence identifiable according to the method for identifying a MAR sequence using the described bioinformatic tool, the
 40 combined method or the method comprising at least one filter,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
- a synthetic MAR sequence comprising natural MAR element and/or fragments
 45 assembled between linker sequences,

a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

50 Transgenic eukaryotic organisms which can be useful for the present invention are for example selected from the group comprising mammals (mouse, human, monkey etc) and in particular laboratory animals such as rodents in general, insects (drosophila,

etc), fishes (zebra fish, etc.), amphibians (frogs, newt, etc..) and other simpler organisms such as c. elegans, yeast, etc....

- 5 Yet another object of the present invention is to provide a computer readable medium comprising computer-executable instructions for performing the method for identifying a MAR sequence as described in the present invention.

- 10 The foregoing description will be more fully understood with reference to the following Examples. Such Examples, are, however, exemplary of methods of practising the present invention and are not intended to limit the scope of the invention.

EXAMPLES

Example 1: SMAR Scan® and MAR sequences

A first rough evaluation of SMAR Scan® was done by analyzing experimentally defined human MARs and non-MAR sequences. As MAR sequences, the previous results from the analysis of human MARs from SMART Db were used to plot a density histogram for each criterion as shown in Fig. 1. Similarly, non-MAR sequences were also analyzed and plotted. As non-MAR sequences, all Ref-Seq-contigs from the chromosome 22 were used, considering that this latter was big enough to contain a negligible part of MAR sequences regarding the part of non-MAR sequences.

The density distributions shown in Fig. 1 are all skewed with a long tail. For the highest bend, the highest major groove depth and the highest minor groove width, the distributions are right skewed. For the lowest melting temperature, the distributions are left-skewed which is natural given the inverse correspondence of this criterion regarding the three others. For the MAR sequences, biphasic distributions with a second weak peak, are actually apparent. And between MAR and non-MAR sequences distributions, a clear shift is also visible in each plot.

Among all human MAR sequences used, in average only about 70% of them have a value greater than the 75th quantile of human MARs distribution, this for the four different criteria. Similarly concerning the second weak peak of each human MARs distribution, only 15% of the human MAR sequences are responsible of these outlying values. Among these 15% of human MAR sequences, most are very well documented MARs, used to insulate transgene from position effects, such as the interferon locus MAR, the beta-globin locus MAR (Ramezani A, Hawley TS, Hawley RG, "Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator", *Blood*, 101:4717-4724, 2003), or the apolipoprotein MAR (Namciu, S, Blochinger KB, Fournier REK, "Human matrix attachment regions in-sulate transgene expression from chromosomal position effects in *Drosophila melanogaster*", *Mol. Cell. Biol.*, 18:2382-2391, 1998).

Always with the same data, human MAR sequences were also used to determine the association between the four theoretical structural properties computed and the AT-content. Fig. 2 represents the scatterplot and the corresponding correlation coefficient r for every pair of criteria.

Example 2: Distribution plots of MAR sequences by organism

MAR sequences from SMART DB of other organisms were also retrieved and analyzed similarly as explained previously. The MAR sequences density distributions for the mouse, the chicken, the sorghum bicolor and the human are plotted jointly in Fig. 3.

Example 3: MAR prediction of the whole chromosome 22

All RefSeq contigs from the chromosome 22 were analyzed by SMAR Scan® using the default settings this time. The result is that SMAR Scan® predicted a total of 803 MARs, their average length being 446 bp, which means an average of one MAR predicted per 42 777 bp. The total length of the predicted MARs corresponds to 1% of the chromosome 22 length. The AT-content of the predicted regions ranged from

65,1% to 93.3%; the average AT-content of all these regions being 73.5%. Thus, predicted MARs were AT-rich, whereas chromosome 22 is not AT-rich (52.1% AT).

SMARTest was also used to analyze the whole chromosome 22 and obtained 1387 MAR candidates, their average length being 494 bp representing an average of one MAR predicted per 24 765 bp. The total length of the predicted MARs corresponds to 2% of the chromosome 22. Between all MARs predicted by the two softwares, 154 predicted MARs are found by both programs, which represents respectively 19% and 11% of SMAR Scan® and SMARTest predicted MARs. Given predicted MARs mean length for SMAR Scan® and SMARTest, the probability to have by chance an overlapping between SMAR Scan® and SMARTest predictions is 0.0027% per prediction.

To evaluate the specificity of SMAR Scan® predictions, SMAR Scan® analyses were performed on randomly shuffled sequences of the chromosome 22 (Fig. 4). Shuffled sequences were generated using 4 different methods: by a segmentation of the chromosome 22 into nonoverlapping windows of 10 bp and by separately shuffling the nucleotides in each window; by "scrambling" which means a permutation of all nucleotides of the chromosome; by "rubbling" which means a segmentation of the chromosome in fragments of 10 bp and a random assembling of these fragments and finally by order 1 Markov chains, the different states being the all the different DNA dinucleotides and the transition probabilities between these states being based on the chromosome 22 scan. For each shuffling method, five shuffled chromosome 22 were generated and analyzed by SMAR Scan® using the default settings. Concerning the number hits, an average of 3 519 170 hits (sd: 18 353) was found for the permuted chromosome 22 within nonoverlapping windows of 10 bp, 171 936,4 hits (sd: 2 859,04) for the scrambled sequences and 24 708,2 hits (sd: 1 191,59) for the rubbed chromosome 22 and 2 282 hits in average (sd: 334,7) for the chromosomes generated according to order 1 Markov chains models of the chromosome 22, which respectively represents 185% (sd: 0.5% of the mean), 9% (sd: 1.5%), 1% (sd: 5%) and 0.1% (sd: 15%) of the number of hits found with the native chromosome 22. For the number of MARs predicted, which thus means contiguous hits of length greater than 300, 1 997 MARs were predicted with the shuffled chromosome 22 within windows of 10 bp (sd: 31.2), only 2.4 MARs candidates were found in scrambled sequences (sd: 0.96) and none for the rubbed and for the sequences generated according to Markov chains model, which respectively represents 249% and less than 0.3% of the number of predicted MARs found with the native chromosome 22. These data provide indications that SMAR Scan® detects specific DNA elements which organization is lost when the DNA sequences are shuffled.

Example 4: Analysis of known matrix attachment regions in the Interferon locus with SMAR Scan®

The relevance of MAR prediction by SMAR Scan® was investigated by analyzing the recently published MAR regions of the human interferon gene cluster on the short arm of chromosome 9 (9p22). Goetze et al. (already cited) reported an exhaustive analysis of the WP18A10A7 locus to analyze the suspected correlation between BURs (termed in this case stress-induced duplex destabilization or SIDD) and *in vitro* binding to the nuclear matrix (Fig. 9, lower part). Three of the SIDD peaks were in agreement with the *in vitro* binding assay, while others did not match matrix attachment sites. Inspection of the interferon locus with SMAR Scan® (Fig. 9, top part) indicated that three majors peaks accompanied by clusters of SATB1, NMP4 and MEF2 regulators binding sites

correlated well with the active MARs. Therefore, we conclude that the occurrence of predicted CUEs and binding sites for these transcription factors is not restricted to the cLysMAR but may be a general property of all MARs. These results also imply that the SMAR Scan® program efficiently detects MAR elements from genomic sequences.

5 **Example 5: Accuracy of SMAR Scan® prediction and comparison with other predictive tools**

10 The accuracy of SMAR Scan® was evaluated using six genomic sequences for which experimentally determined MARs have been mapped. In order to perform a comparison with other predictive tools, the sequences analyzed are the same with the sequences previously used to compare MAR-Finder and SMARTest. These genomic sequences are three plant and three human sequences (Table 1) totalizing 310 151 bp and 37
15 experimentally defined MARs. The results for SMARTest and MAR-Finder in Table 1 come from a previous comparison (Frisch M, Frech K, Klingenhoff A, Cartharius K, Liebich I and Werner T, In silico pre-diction of scaffold/matrix attachment regions in large genomic sequences, *Genome Research*, 12:349-354, 2001.).
MAR-Finder has been used with the default parameters excepted for the threshold that has been set to 0.4 and for the analysis of the protamine locus, the AT-richness
20 rule has been excluded (to detect the non AT-rich MARs as was done for the protamine locus).

Sequence, description and reference	Length (kb)	Experimentally defined MARs positions (kb)	SMARTest prediction positions (kb)	MAR-Finder prediction positions (kb)	SMART Scan prediction positions (kb)
Oryza Sativa putative ADP-glucose pyro- phosphorylase subunit SH2 and putative NADPH dependant reductase A1 genes (U70541), [4]	30.034	0.0-1.2 5.4-7.4 17.3-18.5 20.0-23.1	- 6.5-7.0 15.2-15.7 16.2-16.6 17.6-18.3 19.6-20.1 20.7-21.3 23.6-23.9 25.0-25.4 27.5-27.9	- - 15.7-15.9 - 17.5-18.4 19.8-20.4 21.3-21.5 23.9-24.2 24.7-25.1 -	- - 15.6-16 - 17.6-18.2 21.6-22 - 23.4-23.8 - -
Sorghum bicolor ADP- glucose pyrophospho- rylase subunit SH2, NADPH-dependant reductase A1-b genes (AF010283), [4]	42.446	0.0-1.5 7.1-9.7 22.4-24.7 32.5-33.7 41.6-42.3	- - 21.3-21.9 22.9-24.0 - 27.3-27.6 - -	- - - 23.2-24.2 - 26.9-27.5 - -	- 7.4-7.7 21.5-21.8 22.9-23.2 23.6-24.0 27.3-27.6 33.4-33.9 -
Sorghum bicolor BAC clone 110K5 (AF124045), [37]	78.195	~0.9 ~5.8 ~6.3 ~9.3 ~15.0 ~18.5 ~21.9 ~23.3 ~25.6 ~29.1 ~34.6 ~44.1 ~48.5 ~57.9 ~62.9 ~67.1 ~69.3 ~73.7	- - - - 15.1-15.8 - 21.7-22.0 - - - - - - 44.1-44.5 47.9-49.5 - - 63.1-63.7 - - 74.3-74.7	- - - - - - - - - - - - - - 47.9-49.4 - - - - -	- - - - - - 21.4-21.9 - - - 29.2-29.5 - 39.0-40.0 48.1-48.6 48.8-49.3 - - - 74.3-74.6
Human alpha-1-antitry- sin and corticosteroid binding globulin intergenic region (AF156545), [35]	30.461	2.6-6.3 22.0-30.4	5.5-6.0 - 25.7-26.2 27.5-27.8 - -	3.0-3.2 5.1-6.0 24.9-25.3 25.5-25.8 26.2-26.4 27.5-28.2	5.4-5.8 - 25.8-26.4 - - -
Human protamine locus (U15422), [24]	53.060	8.8-9.7 32.6-33.6 37.2-39.4 51.8-53.0	- - - -	8.0-8.9* 33.9-34.8* 33.9-34.8* -	- - - -
Human beta-globin locus (U01317), [21]	75.955	1.5-3.0 15.6-19.0 44.7-52.7 60.0-70.0	- 18.0-18.4 - 34.4-34.9 - 56.6-57.1 59.8-60.3 65.6-66.0	- 15.5-16.0 18.0-18.4 - 50.6-50.8 56.5-57.2 58.1-58.5 63.0-63.6	2.3-2.6 15.3-15.6 - - - - 62.8-63.1 -

			67.6-67.9 68.8-69.1	68.7-69.3 -	66.3-66.7 -
Sum(kb)	310.151	at least 56.1	14.5	13.8	9.5
Total numbers :		37	28	25	22
Average kb /predicted MAR			11.076	12.406	14.097
True positives [number of experimentally defined MAR found]			19[14]	20[12]	17[14]
False positives			9	5	5
False negatives			23	25	23
Specificity			19/28= 68%	20/25= 80%	17/22= 77%
Sensitivity			14/37= 38%	12/37= 32%	14/37= 38%

Table 1: Evaluation of SMAR Scan® accuracy

- 5 Six different genomic sequences, three plant and three human sequences, for which experimentally defined MARs are known, were analyzed with MAR-Finder, SMARTest and SMAR Scan®. True positive matches are printed in bold, minus (-) indicates false negative matches. Some of the longer experimentally defined MARs contained more than one in silico prediction, each of them was counted as true positive match.
- 10 Therefore, the number of true in silico predictions is higher than the number of experimentally defined MARs found. Specificity is defined as the ratio of true positive predictions, whereas sensitivity is defined as the ratio of experimentally defined MARs found. * AT-rich rule excluded using MAR-Finder.
- 15 SMARTest predicted 28 regions as MARs, 19 (true positives) of these correlate with experimentally defined MARs (specificity: 68%) whereas 9 (32%) are located in non-MARs (false positives). As some of the longest experimentally determined MARs contains more than one in silico prediction, the 19 true positives correspond actually to 14 different experimentally defined MARs (sensitivity: 38%). MARFinder
- 20 predicted 25 regions as MARs, 20 (specificity: 80%) of these correlate with experimentally defined MARs corresponding to 12 different experimentally defined MARs (sensitivity: 32%). SMAR Scan® predicted 22 regions, 17 being true positives (specificity: 77%) matching 14 different experimentally defined MARs (sensitivity: 38%).
- 25 As another example, the same analysis has been applied to human chromosomes 1 and 2 and lead to the determination of 23 MARs sequences (SEQ ID N° 1 to 23). These sequences are listed in Annex 1 in ST25 format.

30 **Example 6: Analyses of the whole genome using the combined method (SMAR Scan®-pfsearch)**

In order to test the potential correlation between the structural features computed by SMAR Scan® and the S/MAR functional activity, the whole human genome has been analyzed with the combined method with very stringent parameters, in order to get

35 sequences with the highest values for the theoretical structural features computed, which are called "super" S/MARs below. This was done with the hope to obtain predicted MAR elements with a very potential to increase transgene expression and recombinant protein production. The putative S/MARs hence harvested were first

40 analyzed from the bioinformatics perspective in an attempt to characterize and classify them.

6.1 S/MARs predicted from the analysis of the whole human genome

As whole human genome sequence, all human RefSeq (National Center for Biotechnology Information, The NCBI handbook [Internet]. Bethesda (MD): National Library of Medicine (US), Oct. Chapter 17, The Reference Sequence (RefSeq) Project, 2002 (Available from <http://www.ncbi.nih.gov/entrez/query.fcgi?db=Books>) contigs (release 5) were used and analyzed with the combined method, using SMAR Scan® as filter in the first level processing, employing default settings except for the highest bend cutoff value, whereas a stringent threshold of 4.0 degrees (instead of 3.202 degrees) has been used for the DNA bending criterion.

In the second level processing, predicted transcription factors binding have been sought in the sequences selected from the previous step without doing any filtering on these sequences.

The analysis by the combined method of the whole human genome came up with a total of 1757 putative "super" S/MARs representing a total of 1 065 305 bp (0.35% of the whole human genome). Table 2 shows for each chromosome: its size, its number of genes, its number of S/MARs predicted, its S/MARs density per gene and its kb per S/MAR. This table shows that there are very various gene densities per S/MAR predicted for the different chromosomes (standard deviation represents more than 50% of the mean of the density of genes per S/MAR predicted and the fold difference between the higher and the lower density of genes per S/MAR is 6,5). Table 2 also shows that the kb per S/MAR varies less than the density of genes per S/MAR (standard deviation represents 25% of the mean of kb per S/MAR and the fold difference between the higher and the lower kb per S/MAR is 3.2).

Chromosome	Number of genes per chromosome	Size of the chromosome (millions bp)	Number of S/MARs predicted	Density of genes per S/MAR	Kb per S/MAR
1	2544	230	85	29.9	2705
2	1772	241	143	12.3	1685
3	1406	198	101	13.9	1960
4	1036	190	118	8.7	1610
5	1233	180	116	10.6	1551
6	1247	170	94	13.2	1808
7	1383	160	179	7.7	1754
8	942	145	77	12.2	1883
9	1100	119	48	22.9	2479
10	1003	133	71	14.1	1873
11	1692	132	67	25.2	1970
12	1278	131	78	16.3	1679
13	506	97	70	7.2	1385
14	1168	88	36	32.4	2444
15	895	83	35	25.5	2371
16	1107	81	41	27	1975
17	1421	80	37	38.4	2162
18	396	75	51	7.7	1470
19	1621	56	36	45.02	1555
20	724	60	28	25.8	2142
21	355	34	18	19.7	1888
22	707	34	28	25.2	1214
X	1168	154	170	6.8	905
Y	251	25	30	8.3	833
Sum	26 955	3 050	1 757	457	433 12
Mean	1 123	127	73	19	1 804
Sd	510	72.8	45	10	462

Table 2: Number of S/MARs predicted per chromosome. The number of genes per chromosome

corresponds to the NCBI human genome statistics (Build 34 Version 3) (National Center for Biotechnology Information, The NCBI handbook [Internet]. Bethesda (MD): National Library of Medicine (US), Oct. Chapter 17, The Reference Sequence (RefSeq) Project, 2002 (Available from <http://www.ncbi.nih.gov/entrez/query.fcgi?db=Books>) based on GenBank annotations.

Chromosome sizes are the sum of the corresponding human RefSeq (National Center for Biotechnology Information, The NCBI handbook [Internet]. Bethesda (MD): National Library of Medicine (US), Oct. Chapter 17, The Reference Sequence (RefSeq) Project, 2002 (Available from <http://www.ncbi.nih.gov/entrez/query.fcgi?db=Books>) (release 5) contig lengths

6.2 Bioinformatics analysis of "super" MARS for transcription factor binding sites

The 1757 predicted "super" S/MARs sequences obtained previously by SMAR Scan® were then analyzed for potential transcription factors binding sites. This has been achieved using RMatch™ Professional (Kel AE, Gossling E, Reuter I, Cheremushkin E, KelMargoulis OV, Wingender E, MATCH: A tool for searching transcription factor binding sites in DNA sequences, *Nucleic Acids Res.* 31(13):35769, 2003), a weight matrixbased tool based on TRANSFAC (Wingender E, Chen X, Fricke E, Geffers R, Hehl R, Liebich I, Krull M, Matys V, Michael H, Ohnhauser R, Pruss M, Schacherer F, Thiele S, Urbach S, The TRANSFAC system on gene expression regulation, *Nucleic Acids Research*, 29(1):2813, 2001). Match™ 2.0 Professional has been used with most of the default settings Match™ analysis was based on TRANSFAC Professional, release 8.2 (20040630). The sums of all transcription factors binding prediction on the 1757 sequences analyzed according to Match™ are in Table 3. Based on this table, only the transcription factors totalizing at least 20 hits over the 1757 sequences analyzed were considered for further analyses.

Hereafter are some of the human transcription factors that are the most often predicted to bind on the 1757 putative S/MAR sequences and their Match description: Cdc5 (cell division control protein 5) a transcriptional regulator/repressor, Nkx3A a homeodomain protein regulated by androgen, POU1F1 (pituitaryspecific positive transcription factor 1) which is specific to the pituitary and stimulates cells proliferation. Thus, in addition to SATB1, NMP4 and MEF2, other transcription factors can participate in the activity of MARs.

AP1	1	AR	2	Bach2	1	Brn2	1
C/EBP	20	C/EBPgamma	5	CDP CR3	1	COMP1	2
CREBP1	34	Cdc5	858	Cdx2	35	Evi1	472
FOX	78	FOXD3	79	FOXJ2	244	FOXP3	29
Freac7	272	GATA1	2	GATA3	142	GATA4	125
HFH1	12	HFH3	1	HLF	275	HNF1	337
HNF3alpha	23	HNF3beta	71	HP1	2	Lhx3	22
MEF2	114	MRF2	57	Myc	18	NKX3A	849
Nkx25	2	Oct1	191	PBX	5	POU1F1	483
POU3F2	11	POU6F1	29	Pax3	3	Pax6	20
Pit1	505	SRF	8	TEF	2852	TFIIA	14
TTF1	1	V\$MTATA_B	4	VBP	53	Vmw65	1
XFD1	65	XFD2	418	XFD3	2		

Table 3 is a summary of all transcription factors binding prediction (totalizing 20 hits or more) on the 1757 sequences analyzed.

5 6.3 Bioinformatics analysis of predicted "super" MARs for dinucleotide frequencies

10 Various computer analysis were performed in order to easily identify "super" S/MAR sequences using an explicit criterion that could be identified without computing. Among those, a di-nucleotide analysis was performed on the 1757 superMARs, computing each of the 16 possible dinucleotide percentage for each sequence considering both strands in the 5' > 3' direction.

15 A summary (min., max., median, mean, 25th percentile and 75th percentile) as well as the histograms of each dinucleotide percentage over the 1757 S/MAR sequences are respectively presented in Table 4. A similar analysis was performed on randomly selected sequences from the human genome, representing randomly selected non-S/MAR sequences (which might however contain some MARs). Table 5 represents respectively a summary of the dinucleotide content analysis for these sequences.

Table 4: Dinucleotide percentages over the 1757 S/MAR sequences

20

	AA %	AC %	AG %	AT %
Minimum	0.000	0.0000	0.0000	18.50
25th percentile	4.234	0.9372	0.1408	32.11
Median	7.843	2.2408	0.4777	34.68
Mean	7.184	3.2117	1.0865	34.32
75th percentile	10.110	4.7718	1.5096	36.94
Maximum	17.290	12.9479	8.1230	50.00
	CA %	CC %	CG %	CT %
Minimum	0.0000	0.00000	0.0000	0.0000
25th percentile	0.9695	0.00000	0.0000	0.1408
Median	1.9776	0.00000	0.0000	0.4777
Mean	2.6977	0.14123	0.2709	1.0865
75th percentile	3.7543	0.09422	0.1256	1.5096
Maximum	10.4061	4.24837	7.4410	8.1230
	GA %	GC %	GG %	GT %
Minimum	0.00000	0.0000	0.00000	0.0000
25th percentile	0.08696	0.0000	0.00000	0.9372
Median	0.32616	0.0000	0.00000	2.2408
Mean	0.63347	0.2104	0.14123	3.2117
75th percentile	0.83333	0.1914	0.09422	4.7718
Maximum	5.77889	9.8795	4.24837	12.9479
	TA %	TC %	TG %	TT %
Minimum	28.63	0.00000	0.0000	0.000
25th percentile	33.48	0.08696	0.9695	4.234
Median	35.22	0.32616	1.9776	7.843
Mean	35.29	0.63347	2.6977	7.184
75th percentile	37.14	0.83333	3.7543	10.110
Maximum	50.00	5.77889	10.4061	17.290

25

Considering the results of the predicted S/MAR elements and of the nonS/MAR sequences in the summary tables, noticeable differences can be noticed in the AT et TA dinucleotide contents between these two groups of sequences. AT and TA represent respectively at least 18,5 % and 28.6 % of the dinucleotide content of the predicted S/MAR sequences, whereas the minimum percentages for the same dinucleotides in

nonS/MAR sequences are respectively 0.3 % and 0%. Similarly, the maximum CC and GG content in S/MAR sequences is 4.2 %, whereas in nonS/MAR sequences the percentages for these two dinucleotides can amount up to 20.8 %.

The correlation between AT and TA dinucleotide percentages and the DNA highest bend as computed by SMAR Scan® is depicted in Fig. 17 for the predicted S/MAR sequences and in Fig. 18 for the nonS/MAR sequences. The different scatterplots of these figures show that the TA percentage correlates well with the predicted DNA bend as predicted by SMAR Scan®.

Table 5: Dinucleotide percentages over the 1757 nonS/MAR sequences summary

	AA %	AC %	AG %	AT %
Minimum	0.000	1.735	1.512	0.3257
25th percentile	7.096	4.586	6.466	5.1033
Median	9.106	5.016	7.279	6.8695
Mean	8.976	5.054	7.184	7.0108
75th percentile	10.939	5.494	7.969	8.7913
Maximum	17.922	13.816	12.232	23.1788
	CA %	CC %	CG %	CT %
Minimum	3.571	0.8278	0.0000	1.512
25th percentile	6.765	4.1077	0.4727	6.466
Median	7.410	5.5556	0.8439	7.279
Mean	7.411	5.9088	1.2707	7.184
75th percentile	8.010	7.2460	1.5760	7.969
Maximum	15.714	20.8415	12.6074	12.232
	GA %	GC %	GG %	GT %
Minimum	1.319	0.4967	0.8278	1.735
25th percentile	5.495	3.2615	4.1077	4.586
Median	6.032	4.4092	5.5556	5.016
Mean	6.065	4.7468	5.9088	5.054
75th percentile	6.602	5.8824	7.2460	5.494
Maximum	10.423	16.0000	20.8415	13.816
	TA %	TC %	TG %	TT %
Minimum	0.000	1.319	3.571	0.000
25th percentile	3.876	5.495	6.765	7.096
Median	5.625	6.032	7.410	9.106
Mean	5.774	6.065	7.411	8.976
75th percentile	7.464	6.602	8.010	10.939
Maximum	24.338	10.423	15.714	17.922

Four of the novel super MARs were randomly picked and analyzed for AT and TA dinucleotide content, and compared with the previously known chicken lysMAR, considering windows of 100 base pairs (Table 6).

Surprisingly, Applicants have shown that all of the super MARs have AT dinucleotide frequencies greater than 12%, and TA dinucleotides greater than 10% of the total dinucleotides analysed in a window of 100base pairs of DNA. The most efficient MARs display values around 34% of the two dinucleotide pairs.

Table 6. Summary of %AT and TA dinucleotide frequencies of experimentally verified MARs

CLysMAR (average of CUEs)	AT% : 12.03	TA% : 10.29	SEQ ID No
P1_68	AT% : 33.78	TA% : 33.93	SEQ ID No
P1_6	AT% : 34.67	TA% : 34.38	SEQ ID No

P1 42	AT% : 35.65	TA% : 35.52	SEQ ID No
Mean value for all human "super"MARs	AT% : 34.32	TA% : 35.29	
Mean value for all human non-MARs	AT% : 7.01	TA% : 5.77	

6.4 Analysis of orthologous intergenic regions of human and mouse genomes

- 5 In order to get an insight on S/MAR evolution, orthologous intergenic regions of human and mouse genomes have been analysed with SMAR Scan®. The data set used is composed of 87 pairs of complete orthologous intergenic regions from the human and mouse genomes (Shabalina SA, Ogurtsov AY, Kondrashov VA, Kondrashov AS, Selective constraint in intergenic regions of human and mouse genomes, *Trends*
- 10 *Genet*, 17(7):3736, 2001) (average length ~12 000 bp) located on 12 human and on 12 mouse chromosomes, the synteny of these sequences was confirmed by pairwise sequence alignment and consideration of the annotations of the flanking genes (experimental or predicted).
- 15 Analysis of the 87 human and mouse orthologous intergenic sequences have been analysed with SMAR Scan® using its default settings. Analysis of the human sequences yielded a total of 12 S/MARs predicted (representing a total length of 4 750 bp), located on 5 different intergenic sequences.
- 20 Among the three human intergenic sequences predicted to contain a "super" S/MAR using SMAR Scan® stringent settings, one of the corresponding mouse orthologous intergenic sequence is also predicted to contain a S/MAR (human EMBL ID: Z96050, position 28 010 to 76 951 othologous to mouse EMBL ID: AC015932, positions 59 884 to 89 963). When a local alignment of these two orthologous intergenic sequences is
- 25 performed, the best local alignment of these two big regions correspond to the regions predicted by SMAR Scan® to be S/MAR element. A manual search for the mouse orthologs of the two other human intergenic sequences predicted to contain a "super" S/MAR was performed using the Ensembl Genome Browser (<http://ensembl.org>). The mouse orthologous intergenic sequences of these two human sequences were
- 30 retrieved using Ensembl orthologue predictions (based on gene names), searching the orthologous mouse genes for the pairs of human genes flanking these intergenic regions.
- 35 Because SMAR Scan® has been tuned for human sequences and consequently yields little "super"MARs with mouse genomic sequences, its default cutoff values were slightly relaxed for the minimum size of contiguous hits to be considered as S/MAR (using 200 bp instead of 300 bp). Analysis by SMAR Scan® of these mouse sequences predicted several S/MARs having high values for the different computed structural features. This finding suggests that the human MAR elements are conserved across
- 40 species.

Example 7 : Dissection of the chicken lysozyme gene 5'- MAR

- 45 The 3000 base pair 5'-MAR was dissected into smaller fragments that were monitored for effect on transgene expression in Chinese hamster ovary (CHO) cells. To do so, seven fragments of ~400 bp were generated by polymerase chain reaction (PCR). These PCR-amplified fragments were contiguous and cover the entire MAR sequence when placed end-to-end. Four copies of each of these fragments were ligated in a head-to-tail orientation, to obtain a length corresponding to approximately half of that of

the natural MAR. The tetramers were inserted upstream of the SV40 promoter in pGEGFPControl, a modified version of the pGL3Control vector (Promega). The plasmid pGEGFPControl was created by exchanging the luciferase gene of pGL3Control for the EGFP gene from pEGFP-N1 (Clontech). The 5'-MAR-fragment-containing plasmids thus created were co-transfected with the resistance plasmid pSVneo in CHO-DG44 cells using LipofectAmine 2000 (Invitrogen) as transfection reagent, as performed previously (Zahn-Zabal, M., et al., "Development of stable cell lines for production or regulated expression using matrix attachment regions" *J Biotechnol*, 2001. 87(1): p. 29-42.). After selection of the antibiotic (G-418) resistant cells, polyclonal cell populations were analyzed by FACS for EGFP fluorescence.

Transgene expression was expressed at the percentile of high expressor cells, defined as the cells which fluorescence levels are at least 4 orders of magnitude higher than the average fluorescence of cells transfected with the pGEGFPControl vector without MAR. Fig. 5 shows that multimerized fragments B, K and F enhance transgene expression, despite their shorter size as compared to the original MAR sequence. In contrast, other fragments are poorly active or fully inactive.

Example 8 : Specificity of B, K and F regions in the MAR context

The 5'-MAR was serially deleted from the 5'-end (Fig.6, upper part) or the 3'-end (Fig.6, lower part), respectively. The effect of the truncated elements was monitored in an assay similar to that described in the previous section. Figure 6 shows that the loss of ability to stimulate transgene expression in CHO cells was not evenly distributed.

In this deletion study, the loss of MAR activity coincided with discrete regions of transition which overlap with the 5'-MAR B-, K- and F-fragment, respectively. In 5' deletions, activity was mostly lost when fragment K and F were removed. 3' deletions that removed the F and b elements had the most pronounced effects. In contrast, flanking regions A, D, E and G that have little or no ability to stimulate transgene expression on their own (Fig. 5), correspondingly did not contribute to the MAR activity in the 5'- and 3'-end deletion studies (Fig. 6).

Example 9: Structure of the F element

The 465 bp F fragment was further dissected into smaller sub-fragments of 234, 243, 213 bp and 122, 125 and 121 bp, respectively. Fragments of the former group were octamerized (8 copies) in a head-to-tail orientation, while those of the latter group were similarly hexa-decamerized (16 copies), to maintain a constant length of MAR sequence. These elements were cloned in pGEGFPControl vector and their effects were assayed in CHO cells as described previously. Interestingly, fragment FIII retained most of the activity of the full-length F fragment whereas fragment FII, which contains the right-hand side part of fragment FIII, lost all the ability to stimulate transgene expression (Fig. 7). This points to an active region comprised between nt 132 and nt 221 in the FIB fragment. Consistently, multiple copies of fragments FI and FIB, which encompass this region, displayed similar activity. FIIA on its own has no activity. However, when added to FIB, resulting in FIII, it enhances the activity of the former. Therefore FIIA appears to contain an auxiliary sequence that has little activity on its own, but that strengthens the activity of the minimal domain located in FIB.

Analysis of the distribution of individual motifs within the lysozyme gene 5'-MAR is shown in Fig. 8 A, along with some additional motifs that we added to the analysis. Most of these motifs were found to be dispersed throughout the MAR element, and not

specifically associated with the active portions. For instance, the binding sites of transcription factors and other motifs that have been associated with MARs were not preferentially localized in the active regions. It has also been proposed that active MAR sequences may consist of combination of distinct motifs. Several computer programs (MAR Finder, SMARTest, SIDD duplex stability) have been reported to identify MARs as regions of DNA that associate with the DNA matrix. They are usually based on algorithms that utilizes a predefined series of sequence-specific patterns that have previously been suggested as containing MAR activity, as exemplified by MAR Finder, now known as MAR Wiz. The output of these programs did not correlate well with the transcriptionally active portions of the *cLysMAR*. For instance, peaks of activity obtained with MAR Finder did not clearly match active MAR sub-portion, as for instance the B fragment is quite active in vivo but scores negative with MAR Finder (Fig. 8B, compare the top and middle panels). Bent DNA structures, as predicted by this program, did not correlate well either with activity (Fig. 8B, compare the top and bottom panels). Similar results were obtained with the other available programs (data not shown).

The motifs identified by available MAR prediction computer methods are therefore unlikely to be the main determinants of the ability of the *cLysMAR* to increase gene expression. Therefore, a number of other computer tools were tested. Surprisingly, predicted nucleosome binding sequences and nucleosome disfavoured sequences were found to be arranged in repetitively interspersed clusters over the MAR, with the nucleosome favouring sites overlapping the active B, K and F regions. Nucleosome positioning sequences were proposed to consist of DNA stretches that can easily wrap around the nucleosomal histones, and they had not been previously associated with MAR sequences.

Nucleosome-favouring sequences may be modelled by a collection of DNA features that include moderately repeated sequences and other physico-chemical parameters that may allow the correct phasing and orientation of the DNA over the curved histone surface. Identification of many of these DNA properties may be computerized, and up to 38 different such properties have been used to predict potential nucleosome positions. Therefore, we set up to determine if specific components of nucleosome prediction programs might correlate with MAR activity, with the objective to construct a tool allowing the identification of novel and possibly more potent MARs from genomic sequences.

To determine whether any aspects of DNA primary sequence might distinguish the active B, K and F regions from the surrounding MAR sequence, we analyzed the 5'-MAR with MAR Scan®. Of the 38 nucleosomal array prediction tools, three were found to correlate with the location of the active MAR sub-domains (Fig. 9A). Location of the MAR B, K and F regions coincides with maxima for DNA bending, major groove depth and minor groove width. A weaker correlation was also noted with minima of the DNA melting temperature, as determined by the GC content. Refined mapping over the MAR F fragment indicated that the melting temperature valley and DNA bending summit indeed correspond the FIB sub-fragment that contains the MAR minimal domain (Fig. 9B). Thus active MAR portions may correspond to regions predicted as curved DNA regions by this program, and we will refer to these regions as CUE-B, CUE-K and CUE-F in the text below. Nevertheless, whether these regions correspond to actual bent DNA and base-pair unwinding regions is unknown, as they do not correspond to bent DNA as predicted by MAR Wiz (Fig.9B).

Example 10 : Imprints of other regulatory elements in the F fragment

Nucleosome positioning features may be considered as one of the many specific chromatin codes contained in genomic DNA. Although this particular code may contribute to the activity of the F region, it is unlikely to determine MAR activity alone, as the 3' part of the F region enhanced activity of the minimal MAR domain contained in the FIB portion. Using the MatInspector program (Genomatix), we searched for transcription factor binding sites with scores higher than 0.92 and found DNA binding sequences for the NMP4 and MEF2 proteins in the 3' part of the F fragment (Fig. 8B). To determine whether any of these transcription factor-binding sites might localize close to the B and K active regions, the entire 5'-MAR sequence was analyzed for binding by NMP4 and MEF2 and proteins reported to bind to single-stranded or double-stranded form of BURs. Among those, SATB1 (special AT-rich binding protein 1) belongs to a class of DNA-binding transcription factor that can either activate or repress the expression of nearby genes. This study indicated that specific proteins such as SATB1, NMP4 (nuclear matrix protein 4) and MEF2 (myogenic enhancer factor 2), have a specific distribution and form a framework around the minimal MAR domains of *cLysMAR* (Fig. 10). The occurrence of several of these NMP4 and SATB1 binding sites has been confirmed experimentally by the EMSA analysis of purified recombinant proteins (data not shown).

Example 11 : Construction of artificial MARs by combining defined genetic elements

To further assess the relative roles of the various MAR components, the *cLysMAR* was deleted of all three CUE regions (Fig. 11, middle part), which resulted in the loss of part of its activity when compared to the complete MAR sequence similarly assembled from all of its components as a control (Fig. 11, top part). Consistently, one copy of each CUE alone, or one copy of each of the three CUEs assembled head-to-tail, had little activity in the absence of the flanking sequences. These results strengthen the conclusion that optimal transcriptional activity requires the combination of CUEs with of flanking sequences. Interestingly, the complete MAR sequence generated from each of its components, but containing also BglII-BamHI linker sequences (AGATCC) used to assemble each DNA fragment, displayed high transcriptional activity (6 fold activation) as compared to the 4.8 fold noted for the original MAR element in this series of assays (see Fig. 5).

We next investigated whether the potentially curved DNA regions may also be active in an environment different from that found in their natural MAR context. Therefore, we set up to swap the CUE-F, CUE-B and CUE-K elements, keeping the flanking sequences unchanged. The sequences flanking the CUE-F element were amplified by PCR and assembled to bracket the various CUEs, keeping their original orientation and distance, or without a CUE. These engineered ~1.8 kb MARs were then assayed for their ability to enhance transgene expression as above. All three CUE were active in this context, and therefore their action is not restricted to one given set of flanking sequences. Interestingly, the CUE-K element was even more active than CUE-F when inserted between the CUE-F flanking sequences, and the former composite construct exhibited an activity as high as that observed for the complete natural MAR (4.8 fold activation). What distinguishes the CUE-K element from CUE-F and CUE-B is the presence of overlapping binding sites for the MEF-2 and SatB1 proteins, in addition to its CUE feature. Therefore, fusing CUE-B with CUE-F-flanking domain results in a higher density of all three binding sites, which is likely explanation to the increased activity. These results indicate that assemblies of CUEs with sequences containing binding sites

for proteins such as NMP4, MEF-2, SatB1, and/or polyPpolyQ proteins constitute potent artificial MAR sequences.

Example 12 : Expression vectors

Three expression vectors according to the present invention are represented on Figure 12.

Plasmid pPAG01 is a 5640 bp pUC19 derivative. It contains a 2960 bp chicken DNA fragment cloned in *Bam*H1 and *Xba*I restriction sites. The insert comes from the border of the 5'-end of the chicken lysozyme locus and has a high A/T-content.

Plasmid pGEGFP (also named pSV40EGFP) control is a derivative of the pGL3-control vector (Promega) in which the luciferase gene sequence has been replaced by the EGFP gene sequence from the pEGFP-N1 vector (Clontech). The size of pGEGFP plasmid is 4334bp.

Plasmid pUbCEGFP control is a derivative of the pGL3 with an Ubiquitin promoter.

Plasmid pPAG01GFP (also named pMAR-SV40EGFP) is a derivative of pGEGFP with the 5'-Lys MAR element cloned in the MCS located just upstream of the SV40 promoter. The size of the pPAG01EGF plasmid is 7285bp.

Example 13 : Effect of the additional transfection of primary transfectant cells on transgene expression

One day before transfection, cells were plated in a 24-well plate, in growth medium at a density of 1.35×10^5 cells/well for CHO-DG44 cells. 16 hours post-inoculum, cells were transfected when they reached 30-40% confluence, using Lipofect-AMINE 2000 (hereinafter LF2000), according to the manufacturer's instructions (Invitrogen). Twenty-seven microliters of serum free medium (Opti-MEM; Invitrogen) containing 1.4 μ l of LF2000 were mixed with 27 μ l of Opti-MEM containing 830 ng of linear plasmid DNA. The antibiotic selection plasmid (pSVneo) amounted to one tenth of the reporter plasmid bearing the GFP transgene. The mix was incubated at room temperature for 20 min, to allow the DNA-LF2000 complexes to form. The mixture was diluted with 300 μ l of Opti-MEM and poured into previously emptied cell-containing wells. Following 3 hours incubation of the cells with the DNA mix at 37°C in a CO₂ incubator, one ml of DMEM-based medium was added to each well. The cells were further incubated for 24 hours in a CO₂ incubator at 37°C. The cells were then transfected a second time according to the method described above, except that the resistance plasmid carried another resistance gene (pSVpuro). Twenty-four hours after the second transfection, cells were passaged and expanded into a T-75 flask containing selection medium supplemented with 500 μ g/ml G-418 and 5 μ g/ml puromycin. After a two week selection period, stably transfected cells were cultured in 6-well plates. Alternatively, the cell population was transfected again using the same method, but pTKhygro (Clontech) and pSVdhfr as resistance plasmids. The expression of GFP was analysed with Fluorescence-activated cell sorter (FACS) and with a Fluoroscanner.

Fig.13 shows that the phenotype of the twice-transfected cells (hereafter called secondary transfectants) not only was strongly coloured, such that special bulb and filter were not required to visualize the green color from the GFP protein, but also contained a majority of producing cells (bottom right-hand side FACS histogram) as compared to the parental population (central histogram). This level of fluorescence corresponds to specific cellular productivities of at least 10 pg per cell per day. Indeed,

cells transfected only one time (primary transfectants) that did not express the marker protein were almost totally absent from the cell population after re-transfection. Bars below 10^1 units of GFP fluorescence amounted 30% in the central histogram and less than 5% in the right histogram. This suggested that additional cells had been

Strikingly, the amount of fluorescence exhibited by re-transfected cells suggested that the subpopulation of cells having incorporated DNA twice expressed much more GFP than the expected two-fold increase. Indeed, the results shown in Table 2 indicate that the secondary transfectants exhibited, on average, more than the two-fold increase of GFP expected if two sets of sequences, one at each successive transfection, would have been integrated independently and with similar efficiencies. Interestingly, this was not dependent on the promoter sequence driving the reporter gene as both viral and cellular promoter-containing vectors gave a similar GFP enhancement (compare lane 1 and 2). However, the effect was particularly marked for the MAR-containing vector as compared to plasmids without MAR- (lane 3), where the two consecutive transfections resulted in a 5.3 and 4.6 fold increase in expression, in two distinct experiments.

Type of plasmids	Primary transfection	Secondary transfection	EGFP fluorescence Fold increase
pUbCEGFP	4'992	14'334	2.8
pSV40EGFP	4'324	12'237	2.8
pMAR-SV40EGFP	6'996	36'748	5.3

Type of plasmids	Primary transfection	Secondary transfection	EGFP fluorescence Fold increase
pUbCEGFP	6'452	15'794	2.5
pSV40EGFP	4'433	11'735	2.6
pMAR-SV40EGFP	8'116	37'475	4.6

Table 7. Effect of re-transfecting primary transfectants at 24 hours interval on GFP expression. Two independent experiments are shown. The resistance plasmid pSVneo was co-transfected with various GFP expression vectors. One day post-transfection, cells were re-transfected with the same plasmids with the difference that the resistance plasmid was changed for pSVpuro. Cells carrying both resistance genes were selected on 500 µg/ml G-418 and 5µg/ml puromycin and the expression of the reporter gene marker was quantified by Fluoroscanner. The fold increases correspond to the ratio of fluorescence obtained from two consecutive transfections as compared to the sum of fluorescence obtained from the corresponding independent transfections. The fold increases that were judged significantly higher are shown in bold, and correspond to fluorescence values that are consistently over 2-fold higher than the addition of those obtained from the independent transfections.

The increase in the level of GFP expression in multiply transfected cells was not expected from current knowledge, and this effect had not been observed previously.

Taken together, the data presented here support the idea that the plasmid sequences that primarily integrated into the host genome would facilitate integration of other plasmids by homologous recombination with the second incoming set of plasmid molecules. Plasmid recombination events occur within a 1-h interval after the plasmid DNA has reached the nucleus and the frequency of homologous recombination

between co-injected plasmid molecules in cultured mammalian cells has been shown to be extremely high, approaching unity (Folger, K.R., K. Thomas, and M.R. Capecchi, Nonreciprocal exchanges of information between DNA duplexes coinjected into mammalian cell nuclei. *Mol Cell Biol*, 1985. 5(1): p. 59-69], explaining the integration of multiple plasmid copies. However, homologous recombination between newly introduced DNA and its chromosomal homolog normally occurs very rarely, at a frequency of 1 in 10^3 cells receiving DNA to the most [Thomas, K.R., K.R. Folger, and M.R. Capecchi, High frequency targeting of genes to specific sites in the mammalian genome. *Cell*, 1986. 44(3): p. 419-28.]. Thus, the results might indicate that the MAR element surprisingly acts to promote such recombination events. MARs would not only modify the organization of genes in vivo, and possibly also allow DNA replication in conjunction with viral DNA sequences, but they may also act as DNA recombination signals.

Example 14 : MARs mediate the unexpectedly high levels of expression in multiply transfected cells

If MAR-driven recombination events were to occur in the multiple transfections process, we expect that the synergy between the primary and secondary plasmid DNA would be affected by the presence of MAR elements at one or both of the transfection steps. We examined this possibility by multiply transfections of the cells with pMAR alone or in combination with various expression plasmids, using the method described previously. Table 3 shows that transfecting the cells twice with the pMAR-SV40EGFP plasmid gave the highest expression of GFP and the highest degree of enhancement of all conditions (4.3 fold). In contrast, transfecting twice the vector without MAR gave little or no enhancement, 2.8-fold, instead of the expected two-fold increase. We conclude that the presence of MAR elements at each transfection step is necessary to achieve the maximal protein synthesis.

Table 8

Primary transfection		Secondary transfection		
Type of plasmid	EGFP-fluorescence	Type of plasmid	EGFP-fluorescence	Fold increase
pMAR	0	pMAR	0	0
		pSV40EGFP	15'437	2.3-2.5
		pMAR-SV40EGFP	30'488	2.6-2.7
pMAR-SV40EGFP	11'278	pMAR-SV40EGFP	47'027	4.3-5.3
		pMAR	12'319	1.0-1.1
pSV40EGFP	6'114	pSV40EGFP	17'200	2.8
		pMAR	11'169	1.8-2.3

Interestingly, when cells were first transfected with pMAR alone, and then re-transfected with pSV40EGFP or pMAR-SV40EGFP, the GFP levels were more than doubled as compared to those resulting from the single transfection of the later plasmids (2.5 and 2.7 fold respectively, instead of the expected 1-fold). This indicates that the prior transfection of the MAR can increase the expression of the plasmid used in the second transfection procedure. Because MARs act only locally on chromatin structure and gene expression, this implies that the two types of DNA may have integrated at a similar chromosomal locus. In contrast, transfecting the GFP expression vectors alone, followed by the MAR element in the second step, yielded little or no improvement of the GFP levels. This indicates that the order of plasmid

transfection is important, and that the first transfection event should contain a MAR element to allow significantly higher levels of transgene expression.

If MAR elements favoured the homologous recombination of the plasmids remaining in episomal forms from the first and second transfection procedures, followed by their co-integration at one chromosomal locus, one would expect that the order of plasmid transfection would not affect GFP levels. However, the above findings indicate that it is more favourable to transfect the MAR element in the first rather than in the second transfection event. This suggests the following molecular mechanism: during the first transfection procedure, the MAR elements may concatemerize and integrate, at least in part, in the cellular chromosome. This integrated MAR DNA may in turn favour the further integration of more plasmids, during the second transfection procedure, at the same or at a nearby chromosomal locus.

Example 15 : MARs as long term DNA transfer facilitators

If integrated MARs mediated a persistent recombination-permissive chromosomal structure, one would expect high levels of expression even if the second transfection was performed long after the first one, at a time when most of the transiently introduced episomal DNA has been eliminated. To address this possibility, the cells from Table 3, selected for antibiotic resistance for three weeks, were transfected again once or twice and selected for the incorporation of additional DNA resistance markers. The tertiary, or the tertiary and quaternary transfection cycles, were performed with combinations of pMAR or pMAR-SV40EGFP, and analyzed for GFP expression as before.

Table 9

Tertiary transfection			Quaternary transfection		
Type of plasmid	EGFP-fluorescence	Fold increase	Type of plasmid	EGFP-fluorescence	Fold increase
pMAR	18368	2.2	pMAR	43'186	2.4
			pMAR-SV40EGFP	140'000	7.6
pMAR-SV40EGFP	16544	2.0	pMAR-SV40EGFP	91'000	5.5
			pMAR	33'814	2.0

Table 9. MARs act as facilitator of DNA integration.

The pMAR-SV40EGFP/ pMAR-SV40EGFP secondary transfectants were used in a third cycle of transfection at the end of the selection process. The tertiary transfection was accomplished with pMAR or pMAR-SV40EGFP, and pTKhygro as selection plasmid, to give tertiary transfectants. After 24 hours, cells were transfected again with either plasmid and pSVdhfr, resulting in the quaternary transfectants which were selected in growth medium containing 500 µg/ml G-418 and 5µg/ml puromycin, 300 µg/ml hygromycin B and 5µM methotrexate. The secondary transfectants initially exhibited a GFP fluorescence of 8300. The fold increases correspond to the ratio of fluorescence obtained from two consecutive transfections as compared to the sum of

fluorescence obtained from the corresponding independent transfections. The fold increases that were judged significantly higher are shown in bold, and correspond to fluorescence values that are 2-fold higher than the addition of those obtained from the independent transfections.

These results show that loading more copies of pMAR or pMAR-SV40EGFP resulted in similar 2-fold enhancements of total cell fluorescence. Loading even more of the MAR in the quaternary transfection further enhanced this activity by another 2.4-fold. This is consistent with our hypothesis that newly introduced MAR sequences may integrate at the chromosomal transgene locus by homologous recombination and thereby further increase transgene expression.

When the cells were transfected a third and fourth time with the pMAR-SV40EGFP plasmid, GFP activity further increased, once again to levels not expected from the addition of the fluorescence levels obtained from independent transfections. GFP expression reached levels that resulted in cells visibly glowing green in day light (Fig.14). These results further indicate that the efficiency of the quaternary transfection was much higher than that expected from the efficacy of the third DNA transfer, indicating that proper timing between transfections is crucial to obtain the optimal gene expression increase, one day being preferred over a three weeks period. We believe that MAR elements favour secondary integration events in increasing recombination frequency at their site of chromosomal integration by relaxing closed chromatin structure, as they mediate a local increase of histone acetylation (Yasui, D., et al., SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature*, 2002. 419(6907): p. 641-5.]. Alternatively, or concomitantly, MARs potentially relocate nearby genes to subnuclear locations thought to be enriched in trans-acting factors, including proteins that can participate in recombination events such as topoisomerases. This can result in a locus in which the MAR sequences can bracket the pSV40EGFP repeats, efficiently shielding the transgenes from chromatin-mediated silencing effects.

Example 16 : Use of MARs identified with SMAR Scan® II to increase the expression of a recombinant protein.

Four MAR elements were randomly selected from the sequences obtained from the analysis of the complete human genome sequence with SMAR Scan® or the combined method. These are termed 1_6, 1_42, 1_68, (where the first number represents the chromosome from which the sequence originates, and the second number is specific to the predicted MAR along this chromosome) and X_S29, a "super" MAR identified on chromosome X. These predicted MARs were inserted into the pGEGFPControl vector upstream of the SV40 promoter and enhancer driving the expression of the green fluorescent protein and these plasmids were transfected into cultured CHO cells, as described previously (Zahn-Zabal, M., et al., *Development of stable cell lines for production or regulated expression using matrix attachment regions*. J Biotechnol, 2001. 87(1): p. 29-42). Expression of the transgene was then analyzed in the total population of stably transfected cells using a fluorescent cell sorter (FACS) machine. As can be seen from Fig. 19, all of these newly identified MARs increased the expression of the transgene significantly above the expression driven by the chicken lysosyme MAR, the "super" MAR X_S29 being the most potent of all of the newly identified MARs.

Example 17: Effect on hematocrit of *in vivo* expression of mEpo by electrotransfer of Network system with and without Human MAR (1-68).

5 The therapeutic gene encodes EPO (erythropoietin), an hormone used for the treatment of anemia. The EPO gene is placed under the control of a doxycycline inducible promoter, in a gene switch system described previously called below the Network system (Imhof, M. O., Chatellard, P., and Mermod, N. (2000). A regulatory network for efficient control of transgene expression. J. Gene. Med. 2, 107-116.). The
10 EPO and regulatory genes are then injected in the muscle of mice using an *in vivo* electroporation procedure termed the electrotransfer, so that the genes are transferred to the nuclei of the muscle fibers. When the doxycycline antibiotic is added to the drinking water of the mice, this compound is expected to induce the expression of EPO, which will lead to the elevation of the hematocrit level, due to the increase in red blood
15 cell counts mediated by the high levels of circulating EPO. Thus, if the MAR improved expression of EPO, higher levels of hematocrit would be expected.

In vivo experiments were carried out on 5 week-old C57BL6 female mice (Iffa Credo-Charles River, France). 30µg of plasmid DNA in normal saline solution was delivered by
20 trans-cutaneous injections in the tibialis anterior muscle. All injections were carried out under Ketaminol (75 mg/kg) and Narcoxyl (10 mg/kg) anesthesia. Following the intramuscular injection of DNA, an electrical field was applied to the muscle. A voltage of 200 V/cm was applied in 8 ms pulses at 1Hz (Bettan M, Darteil R, Caillaud JM, Soubrier F, Delaere P, Branelec D, Mahfoudi A, Duverger N, Scherman D. 2000. "High-level protein secretion into blood circulation after electric pulse-mediated gene transfer into skeletal muscle". *Mol Ther.* 2: 204-10).

16 mice were injected by the Network system expressing EPO without the 1_68 MAR and 16 other mice were injected with the Network system incorporating the MAR in 5' of
30 the promoter/enhancer sequences driving the expression of the activator and EPO genes. In each group, half of the mice were submitted to doxycycline in drinking water from the beginning of the experiment (day 0 – the day of electrotransfer) and in the other half, doxycycline was put in drinking water starting at day 21.

35 Blood samples were collected using heparinated capillaries by retro-orbital puncture at different times after the injection of plasmids. Capillaries were centrifugated 10 minutes at 5000 rpm at room temperature and the volumetric fraction of blood cells is assessed in comparison to the total blood volume and expressed as a percentile, determining the hematocrit level.

40 As can be deduced from Fig. 16 The group of mice injected by MAR-network, induced from the beginning of the experiment, display a better induction of the hematocrit in comparison of mice injected by original network without MAR. After 2 months, haematocrits in "MAR-containing group" is still at values higher (65%) than normal
45 hematocrit levels (45-55%).

More importantly, late induction (day 21) is possible only in presence of MAR but not from mice where the Network was injected without the MAR. Thus the MAR likely
50 protects the transgenes from silencing and allows induction of its expression even after prolong period in non-inducing conditions.

Overall, the MAR element is able to increase the expression of the therapeutic gene as detected from its increased physiological effect on the hematocrit.

CLAIMS

1. A purified and isolated DNA sequence having protein production increasing activity characterized in that said DNA sequence comprises
 - a) at least one bent DNA element,
 - b) and at least one binding site for a DNA binding protein.
2. The purified and isolated DNA sequence of claim 1 characterized in that the bent DNA element contains at least 10% of dinucleotide TA, and/or at least 12% of dinucleotide AT on a stretch of 100 contiguous base pairs.
3. The purified and isolated DNA sequence of claim 2 characterized in that the bent DNA element contains at least 33% of dinucleotide TA, and/or at least 33% of dinucleotide AT on a stretch of 100 contiguous base pairs.
4. The purified and isolated DNA sequence of claims 1 to 2, characterized in that it comprises a MAR nucleotide sequence selected from the group comprising the sequences SEQ ID Nos 1 to 27, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.
5. The purified and isolated DNA sequence of claims 1 to 2, characterized in that it comprises a cLysMAR element and/or fragment, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.
6. The purified and isolated DNA sequence of claim 5, characterized in that said part thereof is a nucleotide sequence selected from the B, K and F regions.
7. The purified and isolated sequence of claims 1 to 6, characterized in that said DNA binding protein is a transcription factor.
8. The purified and isolated sequence of claim 7, characterized in that the transcription factor is selected from the group comprising the polyQpolyP domain proteins.
9. The purified and isolated sequence of claim 7, characterized in that the transcription factor is selected from the group comprising SATB1, NMP4, MEF2, S8, DLX1, FREAC7, BRN2, GATA 1/3, TATA, Bright, MSX, AP1, C/EBP, CREBP1, FOX, Freac7, HFH1, HNF3alpha, Nkx25, POU3F2, Pit1, TTF1, XFD1, AR, C/EBPgamma, Cdc5, FOXD3, HFH3, HNF3 beta, MRF2, Oct1, POU6F1, SRF, V\$MTATA_B, XFD2, Bach2, CDP CR3, Cdx2, FOXJ2, HFL, HP1, Myc, PBX, Pax3, TEF, VBP, XFD3, Brn2, COMP1, Evl, FOXP3, GATA4, HFN1, Lhx3, NKX3A, POU1F1, Pax6, TFIIA and Vmw65 or a combination of two or more of these transcription factors.
10. A purified and isolated cLysMAR element and/or fragment having protein production increasing activity, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.
11. The cLysMAR element and/or fragment of claim 10 consisting of at least one nucleotide sequence selected from the B, K and F regions.

12. A synthetic MAR sequence comprising natural MAR elements and/or fragments assembled between linker sequences.

13. The synthetic MAR sequence of claim 12, characterized in that the MAR comprises a cLysMAR element and/or fragment, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

14. The synthetic MAR sequence of claims 12 to 13, characterized in that the linker sequences are BglII-BamHI linker.

15. A method for identifying a MAR sequence using a Bioinformatic tool comprising the computing of values of one or more DNA sequence features corresponding to DNA bending, major groove depth and minor groove width potentials and melting temperature.

16. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 15, characterized in that said Bioinformatic tool contains algorithms, adapted to the use of profiles or weight-matrices, to compute values for one or more of said DNA sequence features corresponding to DNA bending, major groove depth and minor groove width potentials, and melting temperature.

17. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 16, characterized in that said profiles or weight-matrices are based on dinucleotide recognition.

18. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 17, characterized in that said Bioinformatic tool computes values for all of said DNA sequence features,

19. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 18, characterized in that said Bioinformatic tool is SMAR Scan®.

20. The method for identifying a MAR sequence using a Bioinformatic tool according to claims 15-19, characterized in that the identification of one or more DNA sequence features further comprises a feature corresponding to one or more binding sites for DNA binding proteins.

21. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 20, characterized in that said DNA binding protein is a transcription factor.

22. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 21, characterized in that the transcription factor is selected from the group comprising polyQpolyP domain proteins or transcription factors.

23. The method for identifying a MAR sequence using a Bioinformatic tool according to claims 20 to 21, characterized in that the DNA binding protein is selected from the group comprising SATB1, NMP4, MEF2, S8, DLX1, FREAC7, BRN2, GATA 1/3, TATA, Bright, MSX, AP1, C/EBP, CREBP1, FOX, Freac7, HFH1, HNF3alpha, Nkx25, POU3F2, Pit1, TTF1, XFD1, AR, C/EBPgamma, Cdc5, FOXD3, HFH3, HNF3 beta, MRF2, Oct1, POU6F1, SRF, V\$MTATA_B, XFD2, Bach2, CDP CR3, Cdx2, FOXJ2, HFL, HP1, Myc, PBX, Pax3, TEF, VBP, XFD3, Brn2, COMP1, Evl, FOXP3, GATA4,

HFN1, Lhx3, NKX3A, POU1F1, Pax6, TFIIA and Vmw65 or a combination of two or more of these transcription factors.

24. The method for identifying a MAR sequence using a Bioinformatic tool according to claims 15-23, characterized in that values for the identification of DNA bending are comprised between 3 to 5 °.

25. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 24, characterized in that values for the identification of DNA bending are comprised between 3.8 to 4.4 °.

26. The method for identifying a MAR sequence using a Bioinformatic tool according to claims 15-25 characterized in that values for the identification of the major groove depth are comprised between 8.9 to 9.3 Å and values for the identification of minor groove width are comprised between 5.2 to 5.8 Å.

27. The method for identifying a MAR sequence using a Bioinformatic tool according to claims 26, characterized in that values for the identification of major groove depth are comprised between 9.0 to 9.3 Å and values for the identification of minor groove width are comprised between 5.4 to 5.7 Å.

28. The method for identifying a MAR sequence using a Bioinformatic tool according to claims 15-27, characterized in that the melting temperature is comprised between 55 to 75 ° C.

29. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 28, characterized in that the melting temperature is comprised between 55 to 62 ° C.

30. The method for identifying a MAR sequence using a Bioinformatic tool of claims 15 to 29, characterized in that it further comprises at least one filter predicting DNA binding sites for DNA transcription factors.

31. The method for identifying a MAR sequence using a Bioinformatic tool according to claim 30, characterized in that the filter is applied before or after the Bioinformatic tool.

32. The method according to claims 30 to 31, characterized in that the filter detects clusters of DNA binding sites using profiles or weightmatrices.

33. The method according to claim 32, characterized in that the filter detects densities of clusters of DNA binding sites.

34. A method for identifying a MAR sequence characterized in that it comprises at least one filter detecting clusters of DNA binding sites using profiles or weightmatrices.

35. A purified and isolated MAR DNA sequence identifiable according to claims 15 to 33 or claim 34.

36. The purified and isolated MAR DNA sequence of claim 35, containing at least 10% of dinucleotide TA on a stretch of 100 contiguous base pairs.

37. The purified and isolated MAR DNA sequence of claim 36, containing at least 33% of dinucleotide TA on a stretch of 100 contiguous base pairs.

5 38. The purified and isolated MAR DNA sequence of claims 35 to 37, further containing at least 12% of dinucleotide AT on a stretch of 100 contiguous base pairs.

39. The purified and isolated MAR DNA sequence of claim 38, further containing at least 33% of dinucleotide AT on a stretch of 100 contiguous base pairs.

10 40. The purified and isolated MAR DNA sequence of any of claims 35 to 39, comprising a sequence selected from the sequences SEQ ID Nos 1 to 27, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

15 41. The purified and isolated DNA sequence of claim 40, comprising a sequence selected from the sequences SEQ ID Nos 24 to 27, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

20 42. The use of a purified and isolated DNA sequence comprising a first isolated matrix attachment region (MAR) nucleotide sequence which is a MAR nucleotide sequence selected from the group comprising:
 - a purified and isolated DNA sequence of claims 1 to 9,
 - a purified and isolated MAR DNA of claims 35 to 41,
 25 - the sequences SEQ ID Nos 1 to 27,
 - a purified and isolated cLysMAR element and/or fragment,
 - a synthetic MAR sequence of claims 12 to 14,
 a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants for increasing
 30 protein production activity in a eukaryotic host cell.

43. The use of the purified and isolated DNA sequence of claim 42, characterized in that said purified and isolated DNA sequence further comprises a promoter operably linked to a gene of interest.

35 44. The use of the purified and isolated DNA sequence of claims 42 or 43, characterized in that said purified and isolated DNA sequence further comprises at least a second isolated matrix attachment region (MAR) nucleotide sequence which is a MAR nucleotide sequence selected from the group comprising:

40 - a purified and isolated DNA sequence of claims 1 to 9,
 - a purified and isolated MAR DNA of claims 35 to 41,
 - the sequences SEQ ID Nos 1 to 27,
 - a purified and isolated cLysMAR element and/or fragment,
 - a synthetic MAR sequence of claims 12 to 14,
 45 a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants for increasing protein production activity in a eukaryotic host cell.

50 45. The use of the purified and isolated DNA sequence of claim 44, characterized in that said first and at least second MAR sequences are located at both the 5' and the 3' ends of the sequence containing the promoter and the gene of interest.

46. The use of the purified and isolated DNA sequence of claim 44, characterized in that said first and or at least second MAR sequences are located on a sequence distinct from the one containing the promoter and the gene of interest.

5 47. The use of the purified and isolated DNA sequence of any of claims 42 to 46, characterized in that said purified and isolated DNA sequence is in the form of a linear DNA sequence as vector.

10 48. A method for transfecting a eukaryotic host cell, said method comprising
a) introducing into said eukaryotic host cell at least one purified DNA sequence comprising at least one DNA sequence of interest and/or at least one purified and isolated DNA sequence consisting of a MAR nucleotide sequence or other chromatin modifying elements,
15 b) subjecting within a defined time said transfected eukaryotic host cell to at least one additional transfection step with at least one purified DNA sequence comprising at least one DNA sequence of interest and/or with at least one purified and isolated DNA sequence consisting of a MAR nucleotide sequence or other chromatin modifying elements
20 c) selecting said transfected eukaryotic host cell.

49. The method of claim 48, characterized in that said DNA sequence of interest is a gene of interest coding for a protein operably linked to a promoter.

25 50. The method of claims 48 and 49, characterized in that the selected transfected eukaryotic host cells are high protein producer cells with a production rate of at least 10 pg per cell per day.

51. The method of claims 48-50, characterized in that the MAR nucleotide sequence is selected from the group comprising:
30 - a purified and isolated DNA sequence of claims 1 to 9,
- a purified and isolated MAR DNA of claims 35 to 41,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
35 - a synthetic MAR sequence of claims 12 to 14,
a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

52. The method of claims 48-50, characterized in that the MAR nucleotide is a purified and isolated sequence according to claims 1 to 9, a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.
40

53. The method of claims 48 to 52, characterized in that the defined time corresponds to intervals related to the cell division cycle.
45

54. The method of claim 53, characterized in that the defined time is the moment the host cell just has entered into a second cell division cycle.

50 55. A method for transfecting a eukaryotic host cell, said method comprising co-transfecting into said eukaryotic host cell at least one first purified and isolated DNA sequence comprising at least one DNA sequence of interest, and a second and isolated purified DNA comprising at least one MAR nucleotide selected from the group comprising:

- a purified and isolated DNA sequence of claims 1 to 9,
- a purified and isolated MAR DNA of claims 35 to 41,
- the sequences SEQ ID Nos 1 to 27,
- a purified and isolated cLysMAR element and/or fragment,
- a synthetic MAR sequence of claims 12 to 14,

5

a sequence complementary thereof, a part thereof sharing at least 70% nucleotides in length, a molecular chimera thereof, a combination thereof and variants.

56. A process for the production of a protein wherein

10

a) a eukaryotic host cell transfected according to claims 48 to 54 or claim 55, is cultured in a culture medium under conditions suitable for expression of said protein and

b) said protein is recovered.

15

57. A eukaryotic host cell transfected according to any one of claims 48 to 54 or claim 55.

58. A cell transfection mixture or kit comprising at least one purified and isolated DNA sequence according to claims 1 to 9, 10 to 11, 12 to 14 or 35 to 41.

20

59. A transgenic organism characterized in that at least some of its cells have stably incorporated at least one DNA sequence of claims 1 to 9, 10 to 11, 12 to 14 or 35 to 41.

60. A transgenic organism characterized in that its genome has stably incorporated at least one DNA sequence of claims 1 to 9, 10 to 11, 12 to 14 or 35 to 41.

25

61. The transgenic organism of claims 59 and 60 characterized in that some of its cells have been transfected according to the method of claims 48 to 54 or claim 55.

30

62. A computer readable medium characterized in that it comprises computer-executable instructions for performing the method for identifying a MAR sequence of claims 15 to 33 and/or claim 34.

SEL PCT 012.ST25
SEQUENCE LISTING

<110> Selexis S.A.

<120> HIGH EFFICIENCY GENE TRANSFER AND EXPRESSION IN MAMMALIAN
CELLS BY A MULTIPLE TRANSFECTION PROCEDURE OF MAR SEQUENCES

<130> SEL PCT 012

<150> US 60/513,574

<151> 2003-10-24

<150> EP 04 002 722.9

<151> 2004-02-06

<160> 241

<170> PatentIn version 3.1

<210> 1

<211> 320

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(320)

<223> MAR of human chromosome 1, nt from 36686 to 37008

<220>

<221> misc_binding

<222> (1)..(320)

<223> MAR of human chromosome 1, genomic contig; 36686 to 37008

<400> 1

ttatattatg ttgttatata tattatatta tgttattaga ttatattatg ttgttatatt 60

atataataat attatattat atattatata ttatattata taatatataa taatattata 120

taattatata ttacattata taatatataa taatattata taattatata ttacattata 180

taatatataa taatattata taataatata taattatata atatataata atattatata 240

atattatata atattatata atatataaat atataataat atatattata ttatataata 300

gtatataata ttatataata 320

<210> 2

<211> 709

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(709)

<223> MAR of human chromosome 1, nt from 142276 to 142984

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(709)

<223> MAR of human chromosome 1, genomic contig; 142276 to 142984

<400> 2

tacaatatat ttctattat atatatttg tattatatat aatatacaat atatttcta 60

ttatatataa tatatttgt attatatata ttacaatata tttgtatta tataatatat 120

aatacaatat ataatatatt gtattatata ttatataata caatatatta tatattgtat 180

tatatattat atataaact atataatata ttgtattata tattatatat aatactatat 240

aatatatttt attatatatt atatataata caatatataa tatattgtat tataatacaa 300

tgtattataa tgtattatat tgtattatat attatatata atacaatata taataatata 360

ttataatata taataataat ataataat aataatatat attgtattat atattatata 420

atacaatata taatatattg tattatatat atttattac atataatata taatacata 480

tataatatat tttgtattat atataatata tttattatg tattatagat aatatatttt 540

attatatatt atatataata caatatataa tatatttgt attgtatata atatataata 600

caatatataa tatattgtat tatatataat attaatatat tttgtattat atatttatat 660

tttatattat aattatgttt tgcattatat attcatatt atatatacc 709

<210> 3

<211> 409

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(409)

<223> MAR of human chromosome 1, nt from 1368659 to 1369067

<220>

<221> misc_binding

<222> (1)..(409)

<223> MAR of human chromosome 1, genomic contig; 1368659 to 1369067

<400> 3

tacacataaa tacatatgca tatatatattat gtatatatac ataaatacat atgcatatac 60

attatgtata tatacataaa tacatatgca tatacattat gtatatatac ataaatacat 120

atgcatatac attatgtata tatacataaa tacatatgca tatacattat gtatatatac 180

ataaatacat atgcatatac attatgtata tatacataaa tacatatgca tatacattat 240

gtatatatac ataaatacat atgcatatat tatatacata aattatatta tatacataat 300

acatacatat atattatgtg tatatatata taaatacata tacatatatt atgtgtatat 360

atacatgata catatacata tattatgtat atatatacat aaatacata 409

<210> 4

<211> 394

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(394)

<223> MAR of human chromosome 1, genomic contig; 2839089 to 2839482

SEL PCT 012.ST25

<400> 4

tatgtatata tacacacata tgtatatata cacacatatg tatatacgta tatatgtata 60
 tatacacaca tatgtatata cgtatatatg tatatatata cacatatgta tatacgata 120
 tatgtatata tacacacata tgtatatatg tatatatgta tatatacaca catatgtata 180
 tatgtatata tacacacata tgtatatatg tatatatgta tatatacaca catgtgtata 240
 tatatatata catatgtata tatgtatata tacacacata tgtatatatg tgtatgtata 300
 tatacacaca tatgtatata tacacatata tatgtatata tacacacata cttatatata 360
 cacatatata tgtatatata cacatatgta taca 394

<210> 5

<211> 832

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(832)

<223> MAR of human chromosome 1, genomic contig; 1452269 to 1453100

<400> 5

tatattacta tatatacaat atacatatta ctatatatac catgtattac tatatatatc 60
 tactatatat attactatat atacaaaata tatattacta tatatacaat atacatatta 120
 ctatatatac catatattac tatatatatc tactatatat attactatat atacaaaata 180
 tatattacta tatatactat atattactgt atatacaata tatattacta tatatatact 240
 atatattact atatatacac tatatattac tatatatata caatatatat attactatat 300

SEL PCT 012.ST25

atacacaatg tatataacta tatatacaat atatattact atatatacta tatatattac 360
 tatacatact atatattact ctatatatac aatatataa ttacaatata tactacatat 420
 tactacatat actttatata ttactatata tactatata t actgtatat acaatatata 480
 ttactaaata tacacaatat atattactat atatacaaa tatatatatt actatatata 540
 cacattatat atgactatat atacacacta tatatattac tatatatata caatatataa 600
 ctatatatac acagtataca tattactata tatacacaaat atatatatta ctatatatac 660
 actatatatt actatatata cacaatatat attactctat gtatacacta tatatattac 720
 tatatatata gaatatatat aactatatat acactatatt actatatata ctatatatta 780
 ctatatgtac tatatatatt actatatata ctatatatta ctatatatac ac 832

<210> 6

<211> 350

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(350)

<223> MAR of human chromosome 1, genomic contig; 831495 to 831844

<400> 6

aatatataat atataaatat taatatgtat tatataaat atattaatat attatattat 60
 attactatat aaataatatt aatatattat attaaaatat taataaatat atcatattaa 120
 atattatatt aattaaatat taataaatat attatattaa tatatttata tattaaacct 180
 ataacatatg catatactta ttatatata acatgcatgt acttatttat atatacaata 240
 tatatttata tattatataa tatattatat gtatttatat attatatatc atatattata 300

SEL PCT 012.ST25

tgtatttata tattatatat catataatat atatatttat attatatata 350

<210> 7

<211> 386

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(386)

<223> MAR of human chromosome 1, genomic contig; 1447225 to 1447610

<400> 7

acatttaatt taattatata ctgctatata taattaaatc tatatatcta tataacttat 60

aattttattt aatttaatta tatatactat atagttatat atacatatat gtaattatat 120

atagtataat tatagtatat atgtatatat aatgtaagta aatatatagt atatatttat 180

atatactata tatttataca tatgtcttta tatatactaa tatatataca catatgtaat 240

atgtacatat ggcatatatt ttatagtgtat tatatacata tatgtaatat atatagtaat 300

atgtaaatat atagtacata tttaattata tggtaatatata tacacatata tgtaatatgt 360

gtatttatagt acatatttta tagtat 386

<210> 8

<211> 585

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(585)

<223> MAR of human chromosome 1, genomic contig; 4955365 to 4955949

<400> 8

atacacacat atacacatat gtacgtatat atactatata tacacacata tacacatatg 60
 tacgtatata tactatatat acacacatat acacatatgt acgtatatat actatatata 120
 cacacatata cacatatgta cgtatatata ctatatatac acacatatatac acatatgtac 180
 gtatatatac tatatatata cacatatata catatgtacg tatatattat atatacacac 240
 atatacacat atgtacgtat atatactata tatacacaca tatacacata tgtacgtata 300
 tatactatat atacacacat atacacatat gtacgtatat atactatata tacacacata 360
 tacacatatg tacgtatata tactatatat acacacatat acacatatgt acgtatatat 420
 actatatata cacacatata cacatatgta cgtatatata ctatatatac acacatatatac 480
 acatatgtac gtatatatac tatatatata cacatatata catatgtacg tatatatata 540
 atatataccc atacacatac gtatatatacgt acatatatat acgta 585

<210> 9

<211> 772

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(772)

<223> MAR of human chromosome 1, genomic contig; 5971862 to 5972633

SEL PCT 012.ST25

<400> 9

agtaaacata tatatagtaa atatatatag tgtatatata gtaaatatat atagtgcata 60
 tatatagtc atatatatag tgtatatata gtaaatatat agtgtatata tatagtaa 120
 atatatatgt tatatatagt aaatatatat agtaaata tatatactat atatagtaa 180
 tatatatata ctatatatag taaatatata tatagtatat atatagtaa tatatatata 240
 gtatatatat agtaaata tatatagtat atatatagta aatatatata tagtatatat 300
 agtaaata tatagtatat atatagtaa tatatatata gtatatatat agtaaata 360
 tatatagtat atatatagta aatatatata tagtatatat atagtaaata tatatagtat 420
 atatatagta aatatatata gtatatatat agtaaata tatagtatat atatagtaa 480
 tatatatata ctgtatatat atagtaaata tatatact gtatatatat agtaaata 540
 tatatactgt atatatatag taaatatata tacactgtat atatatagta aatatatata 600
 cactgtatat acatagtaa tatatatata ctgtatat acatagtaa tatatact 660
 gtatatatat agtaaata tatatactgt atatacatag taaatatata tacagtgtat 720
 atacatagta aatatatata cagtgtatat acatagtaa tatatatata gt 772

<210> 10

<211> 304

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(304)

<223> MAR of human chromosome 1, genomic contig; 6221897 to 6222200

<400> 10

SEL PCT 012.ST25

atatataata tatataatta tatttatatat aatatataat atatataatt atattatata 60
 ttatatataa tatatttatat attatatata taatatatat tatatattaa atatataatta 120
 tatatataat atatattata tattaaatat atattatata tataatatat attatatata 180
 atatataata tatatattat atatataatta tatatttatat atatataatta tatatatata 240
 atatataata tatatattat atataatata tatttatatat atataatata tataatatat 300
 atta 304

<210> 11

<211> 311

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(311)

<223> MAR of human chromosome 1, genomic contig; 9418531 to 9418841

<400> 11

tatatataat atttatatat aatattcatg tatttatata taaatattta tatatttata 60
 tataaatatt tatatattta tatataaata ttatatatt tatatataat atttatacat 120
 tatatataat atttatatat tatatataat atttatatat aatatttata tatttatatat 180
 aatatttata tatttatatg tataatatat attttatata tgtatgtata atatataatt 240
 tatatatgta tgtataatat attttatata tgtatgtata atatattatt atatataata 300
 tataatttat a 311

<210> 12

<211> 302

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(302)

<223> MAR of human chromosome 1, genomic contig; 15088789 to 15089090

<400> 12

atataatata tatattatat atataaatat atataaatat ataacatata tattatatat 60

aaatatatat aaatatataa catatatatt atatatataa atatatataa atatatataa 120

tatatattat atatatataat atatatataat atataacata tatattatat atataaatat 180

atattatata ttatatata taatatatat aaatatataa tatatatatta tatatatataat 240

atatatataat atataatata tatatttata tataatatat ataaatatat aatatataat 300

at 302

<210> 13

<211> 461

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(461)

<223> MAR of human chromosome 1, genomic contig; 6791827 to 6792287

SEL PCT 012.ST25

<400> 13

tatataatat atattatata tacacatata taatatatat tatatatata catatataat 60

atatattata tatacacata tataatatat attatatata cacatatata atatattatta 120

tatatacaca tatataatat atattatata tacacatata taatatatat tatatatata 180

catatataat atatattata tatacacata tataatatat attatatata cacatatata 240

atatattatta tatatacaca tatgtaatat atattatata cacacatata atatattatta 300

tatacacata tataatatat attatatata catatataat atatattata tatacacata 360

tataatatat attatatata cacatatata atatattatta tatatacaca tataatatat 420

aatatatata catatataat atatattatta tatatgcaca t 461

<210> 14

<211> 572

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(572)

<223> MAR of human chromosome 1, genomic contig; 163530 to 164101

<400> 14

atattataat tatatatatt atatataatt atataaaata tatattataa ttatatatat 60

tttatataat atatattatta taattaatat attatatata atatatatat tatatatata 120

atatattatta tatatattat atataatata tataatatat ataatatata atataatata 180

tatattatat ataatatata atatatatata tatattataa tataatatat ataatatata 240

atataatata tataatatat aatataatat ataatatata atatatatata tatataatat 300

SEL PCT 012.ST25

aatatataat atatataata tataatataa tatataatat atataatata ttataatata 360

atatatataa tatataatat aatatatata atatataata taatatataa tatataatat 420

atatttaata tatttattaa ttatttgta tatatttatt aatatataat atataatata 480

tttaatatat tataactata tatttatatta taattatata tatttatatat atacaattat 540

aattatatat tatatatact tataatatat at 572

<210> 15

<211> 357

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(357)

<223> MAR of human chromosome 1, genomic contig; 1842332 to 1842688

<400> 15

tatatctata tatatctata tatatataat atagataata tctatatata taatatagat 60

aatattatct atatataata tagataatat tatctatata taatatagat aatattatct 120

atatataaaa ttatattata tctatatata ttatatatat aaaattatat tatatctata 180

tataatatag ataatatcta tatataaata gataatatct atatatataa tatagatatt 240

atctatatatta tagatataga taatattatc tatattatag atattatcta tatataatat 300

agataatatt atctatatatta tatatataat atatctatat tatctataat attatct 357

<210> 16

<211> 399

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(399)

<223> MAR of human chromosome 1, genomic contig; 2309560 to 2309958

<400> 16

attatatata atatatatta tatattatat atatcaagca gcagatatata tatataatat 60

atataatata tataatatat attgtatatt atataatata taatatatat aatatatatt 120

gtatattata taatatataa tatatataat atatattgta tattatataa tatataatat 180

atataatata tattgtatat tatataatat ataatatatg taatatatta tgtaatatat 240

tatataatat atattatata ttatatataa tatataattat atataatata tattacataa 300

tatattacat atattacgta atatatgtta tatattacat ataatatata acatatatta 360

cgtaatatat gtaatatatt acatataata tatacatta 399

<210> 17

<211> 394

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(394)

<223> MAR of human chromosome 1, genomic contig; 2231759 to 2232152

SEL PCT 012.ST25

<400> 17

atatatactt ataaattata tacttatata tacttataaa ttatatactt atatatactt 60

ataaattata tacttatata tacttataaa ttatatactt atatatactt ataaattata 120

tacttatata tacttataaa ttatatactt atatatactt ataaattata tacttatata 180

tacttataaa ttatatactt atatataatt ataaattata tacttatata taattataaa 240

ttatatactt atatataatt ataaattata tacttatata taattataaa ttatatactt 300

atatataatt ataaattata tacatatata taattataaa ttatatacat atataattat 360

aaattatata catatataat tataaattat atac 394

<210> 18

<211> 387

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(387)

<223> MAR of human chromosome 1, genomic contig; 7406524 to 7406910

<400> 18

tatattatat ataatatata ttatatataa tataaataat atatattata tataatatat 60

aaataatata taatatataa ataatatata atatataata tataaataat atataatata 120

taacatatata ataatatata taatatataa ataatatata taatatataa ataatatata 180

taatatataa aaatatataa tatataatac atatataaat aatatattat attatatatg 240

atacataata tattatatat aatatattat atgatacata atatattata tagaatatat 300

tatatgatac ataatatatt atatagaata tattatatga tacataatat attatatgat 360

SEL PCT 012.ST25

acataatata ttatatataa tatatta

387

<210> 19

<211> 370

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(370)

<223> MAR of human chromosome 1, genomic contig; 9399572 to 9399941

<400> 19

catatataca tatatacaca tatatacaca tatatataca catacatatg tacacatata 60

tatacacata tgtatacaca tatatacaca tatatacaca catatataca catatataca 120

cacatatata cacatatata cacatatata cacatatata catatataca catatataca 180

tatatacaca tatatataat atacacacat atatatacac atatatacac acatatatac 240

acatatatac acatatatat acacatatat acacatatat acatatatac acatatatat 300

acatatatac acatatatac atatatacac atatatacat atatacacac atatatacac 360

atacatatac

370

<210> 20

<211> 377

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(377)

<223> MAR of human chromosome 1, genomic contig; 12417411 to 12417787

<400> 20

attatatata ataatatata ttatatattt atatataaat tataataaat acatatatt 60

atatatttat atataaatta tatataataa ataatatata ttatatatat ttataaatta 120

taataaatac atataattac atatatttat atatgaatta tatataataa ataatatata 180

ttatatatat ttatatgtag attatatata aatatatata atttatatat ataataatat 240

atataattta tatatatataa tatatatata ataatatat ataatttata tatataatta 300

tatatatataa aatatatataa taatatatat aatttatata tataattata tatataataa 360

atatatatataa ttatat 377

<210> 21

<211> 1524

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1524)

<223> MAR of human chromosome 1, genomic contig; 1643307 to 1644830

<400> 21

tataaatata tataaatata taaatatata taaatatata aatatatata aatatatata 60

aatatataaa aatatataaa tatatatataa tatatatataa tatataaaaa cataaaaaata 120

SEL PCT 012.ST25

tatataaata tatataaata tataaaaaata tataaatata taaatatata aaaatatata	180
aatatataaa tatatacata aatatatata aatatatata aatatataaa aatatatata	240
aatatataaa tatatataaa tatatataaa tatatataaa tatataaaaa tatatataaa	300
tatataaata tataaaaaata tatataaata tataaatata taaatatata taaatatata	360
aatatataaa taaatatag tatttatgaa tatatatgaa tatataaata tataaaaaat	420
atatataaat atataaatat atataaatat ataaatatat acatatatac atatataaat	480
aaataaatat aagtatttat gaatatatat gaatatataa atatataaaa aatatatata	540
aatatataaa tatatataaa tataaatata taaaaatata taaaaatata tataaatata	600
taaatatata taaatatata aatatatata aatatatata aatatataaa tatatataaa	660
tatatataaa tatataaata tataaatata tataaatata tataaatata taaatatata	720
aatataaata tataaatata tataaatata tataaatata taaatatata taaatatata	780
taaatatata taaatatata taaatatata aatatatata aatatatata taaatatata	840
taaatatata aatatataaa tatataaaaa tatataacaa tatataaata tatataaaaa	900
tatataacaa tatataaata taaatatata taaaaatata taacaatata taaatatata	960
tatatataaa tatataaata taaatatata aaatatatat aaatatataa atatataaa	1020
atatataaat gtataaatat atataaaaaat atataacaat atataaatat ataaatatat	1080
aacaatatat aaatatataa aaatatataa caatatataa atataaatat atataaaaaat	1140
atataacaat atataaatat aaatatatat ataaatatat aaatatataa ataaaaaata	1200
tatataaata tataaatata tatataaata tatataaata tataaatgta taaatatata	1260
taaatatata aatatataaa aatatataaa tatatataaa tatatataaa tatataaata	1320
taaatatata aatatatata aatatataaa tataaatata taaacatata taaatatata	1380
taaataaaca tatataaaga tatataaaga tataaagata tataaatata taaatatata	1440
aagatatata aatatataaa gatataataa tatataaaga tatataaata tataaagata	1500
tataaatata atatataaat atat	1524

SEL PCT 012.ST25

<211> 664

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(664)

<223> MAR of human chromosome 1, genomic contig; 1398763 to 1399426

<400> 22

```

acacatatat atataaaata tatatatata cacacatata tataaaatat atatatatac   60
acacatatat ataaaatata tatatacaca catatatata aaatatatat atacacacat   120
atatataaaa tatatatata cacacatata tataaaatat atatatatac acatatatat   180
aaaatatata tatacacaca tatatataaa atatatatat acacacatat atataaaata   240
tatatatata cacatatata taaaatatat atatacacac atatatataa aatatatata   300
tacacacata tatataaaat atatatatac acacatatat aaaatatata tatacacaca   360
tatataaaat atatatatac acatatatat aaaatatata tatacacata tatataaaat   420
atatatacac acatatatat aaaatatata tatacacaca tatatataaa atatatatat   480
acacatatat ataaaatata tatatacaca tatatataaa atatatatat atacacatat   540
atataaaata tatatacaca catatatata aagtatatat atacacacat atatatataaa   600
tatatatata cacatatata taaaatatat atatacacat atatatataaa tatatatata   660
caca                                     664

```

<210> 23

<211> 1428

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1428)

<223> MAR of human chromosome 2, genomic contig; 17840365 to 17841792

<400> 23

aatttattat atattatata ttatatatat tatatatatt atatattata tatattatat 60
atattatata ttatatatat tatatatatt atatttatat ataatatata tctaatatat 120
atattagata taatatatat ctaatatata tatattttat atatataata tatctctaata 180
atatatattt tatatgtata taatatatct ctaatatata tatattttat atgtatataa 240
tatatctcta atatatatat ttttatata taatatatct ctaatatata tattttatat 300
atataatata tatctaatat atataatata tatattagat atatataaaa tatatatgat 360
atatttatta tatatataat atataatata taatatatat attatattat atacatatat 420
attatataca atatatatta tatatatttt atatacatta tatattatat atattttata 480
tacaatatat attatatatt ttatatacaa tatatatatt atatatttta tatttttata 540
tacaatatat attatatata ttttatatat aatatatatt atatatattt tatataaat 600
atattatata tatttttatat ataatatata ttatataaat tatatataat atatattata 660
ataaattata atatttttta tatatataat atgtatttta tatataatat attataaat 720
atattttata tataatatat tataatatat attttatata taatatatta taatatatat 780
tttatattat aatatattat aatatatatt ttatatataa tatattataa tatatatttt 840
atatataata tattataata tatattttat atataatata ttataatata tatattataa 900
tatatatttt atatataata tattatcata tatatattaa atatatattt tatatataat 960
atattataat atatatatta taatatatat ttatatata atatattata atatatatat 1020

SEL PCT 012.ST25

tataatatat atttatata taatatatta taatatatat ttatatata atatattata 1080

atatatattt tatatataat atattataat atatatttta tatataatat aatatatatt 1140

ttatatataa tatattataa tatatatfff atatataata tattataata tatattttat 1200

atataatata ttataatata tttttatat ataatatatt ataatatata ttttatatat 1260

aatatattat aatatatatt ttatatataa tatattataa tatatatfff atatataata 1320

tattataata tatattttat atataatata ttaattaaat ttattaattt attaattatt 1380

aatatttatt atattattaa ttaataatat ataaattatt aatatata 1428

<210> 24

<211> 4624

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(4624)

<223> MAR 1_6 of chromosome 1

<400> 24

ggatcttaaa tctatfffat ttatttattt ttcatgtggc caataccctc cacccttc 60

ttctgtctct ttcaacttat tgtggttacc ttgaggctac ctgagacagt aggcttgggt 120

ggggaagtat gcattctaag tgtaaagttt gatgagcttt gacaaatgtc aacctatgta 180

ccagaacatt ttatcacccc ataaaatctc cttgtgtca ctgcccagtc agtgtctatt 240

ctagtatcca actcctggct ccaagaaacc attgaactgt ttctgtcac tataaattag 300

atttgtcttt tctagagttt catgtaaatg gaatcataca ctaagtactc ttgtgcctg 360

gcttctgctc agcataatgt tttgagaat cattcatgct gctgcatgtt ttcagtagtt 420

SEL PCT 012.ST25

catttttta aataggtgaa ttgtaactca ttctgtgaat ataccatatt ctgtcttcca 480
 ttatctgtt agtggatctt taggtcggtt ctagtgttg gctattgcaa ataaagctgc 540
 tgtaaatatt aatgcacaag ttcccatgt tcatatgttt catttcactt aggaaaatac 600
 ctaagagagg aattgcacat attaaaaaaa tttaaaaaac tactaagctg ttctccaaaa 660
 tgggtgtaca attttattc ccaagagcaa tatgagtgtt taattgctcc acattctcac 720
 caacactgg tgctgttag ttattttc attgtttca ttgtatgtc tgtgaggcag 780
 cattgatgtg catgtctctg agtgtcatct tagcggtgat gctgagcatc agttcacgtc 840
 ctataggcc gttgtatat ctgcttgtg aaatgtctgt tcaaacttt tgcctatttt 900
 aaattgagtt gtgttctct tcttaggatt aagtaatgag ttaaaaatat ttctgataca 960
 aatcttcat tataatttc taatgcttc tcatctatag ttattttct catattcttt 1020
 aactgtatct ttgaagagc aaatttact ttgattatg cccaatttat caagttttta 1080
 tatggctctt ttgattatgc ccataatcac attagacttt gcctaacca agtttcaga 1140
 gatttttct ttatgcttt tatctagaaa ttgttagtt ttaggtttta aaaaagtta 1200
 attatttat ttgagacagg gtattgctct ttacatatac tggagtgcag tgatgcaatc 1260
 atggctcact gcagcctcaa cctctgggc tcaagcgggt ctcccatctc agagtctga 1320
 gtagctggcc aggtgcatgc cagctcaat gtgttttca ttgcatttc cctgataatt 1380
 attgacgttg agcattttt tcatatatca gttagctatt tgtacgtctt ctttgagaa 1440
 acatctattc gggcttttg cccattttaa agtcagatgg ttgtttgtc agctattgag 1500
 ttgttgagt tcttgtata ttctggatat taccatcttg tcagatgcac agttgcaaa 1560
 tttttttt ctatttgta ggtgtctct ttctgttg ttctctcgg tatgcagaag 1620
 tttttagtg tgatgtaatt tcattgtct gttttgctt ttgtgcctg tactttctta 1680
 ttcttatcca aaaaacttt atctagatca atgtcacgaa gagtttctc tctgtttct 1740
 tcgagtagtt tttaataatt ttgggtatac atttaagtct ttaatctatt tggaattgat 1800
 ttgtcatat ggtgagagat cagagtctaa ttctactt ttggatgtgg aaagctagtt 1860
 tttcagcac catttattga agagactgtc tcttctcaa tgtgtgtct ttgtgcctc 1920

SEL PCT 012.ST25

gtcaaaaatc agttggctgt gcgtaggattt atttctgtgt tctctatttt gttccattgg 1980
tctagtttta gccttaaatt taggtctgca atttttttt tttgtatat ggtgtgaagt 2040
aagagtcaaa gttcattatt ttcatatgg atatgtaatt actccagtac catcatttag 2100
ttgaatgga ctgtccttc tccatggaat tacatgggca tctttgtct gaaaccaatt 2160
atgtatgttt acgtatgtgt atgtttatgc atatgttata ggtttaatat atattaatat 2220
atataatata taatatataa atattaatat gtattatata atatatatta atatattata 2280
ttatattact atataaataa tattaatata ttatattaaa atattaataa atatatcata 2340
ttaaatatta tattaattaa atattaataa atatattata ttaatatatt tatatattaa 2400
acctataaca tatgcatata cttatttata tataacatgc atgtacttat ttatatatac 2460
aatatatatt tatatattat ataatatatt atatgtattt atatattata tatcatatat 2520
tatatgtatt tatatattat atatcatata atatatatat ttatattata tatattatat 2580
gatatataat attatataat gtattaatat atattaaacc tatatttata attctggact 2640
cactattttg ttccattgggt gtctgtgtgt atctaaccct atgccaataa tgtactatct 2700
taattaccat agctttatag taagctttga aatcagatag tgtatttttt atcattgttt 2760
ttaaaataa tagtttatct tttatttga atttgaatc agctagtcag ttctgcaaa 2820
aagcttactg ggattttgct tggaattatg ttacatctgt agcatgtact atccaatatt 2880
ctagccttta tccacatgtg gctattaagg tttaaattaa ttaaattaa atttaattaa 2940
ttaaattaa aacttaataa ttggttcctc attcacacta ccatatgtca agtgttcaat 3000
agccacatat ggtcaatgtc ttgaaaagt caatacagta catttccatt attgcagtaa 3060
gttctgtcaa acagcactat cgtagaccga ttaggagaga actgacttaa cagtattgga 3120
tgctccagtc aatgaacatc tttttttt tcatatttt cagtagtctc tgcagtatat 3180
tatagatttc agtttacata tttgcatat atttattaa atgtataacg gtagaagtac 3240
tattattgga tgatgtgtc tatagatgta ttttaggtca agtttgttga tagtgtgtt 3300
taaactctgt atacctcttg attttttat ttactgttc ttgaattac tgagacagga 3360
atgttatatc ctttaactata ttgtgaatt tattcacttc ttccctcagt tctgttaact 3420

SEL PCT 012.ST25

ttgcttagg tgcttttaa aaatgaaact ttcaatctct gcctttaat ttagcattt 3480
 agaccattta cattcaatgt aattatcaat atcagtttat ttaagtctga agttgtgcaa 3540
 ttttcctct acctatatta taaatcttc tatatacaaa acacatgcta tgtttctgc 3600
 atatgttta aatgacaccc ggaaagcatt gacactatt ttgcttagg ttatcttca 3660
 aagatgttaa aaatgagaaa gaaatattct gcatttatcc atacacttat tatttgcaaa 3720
 ggtttttta aataccttg ttagatttc agttaccaac ttgtattcc tttagctga 3780
 agaacttaca atttctgta ggacaggctct ctgacaacaa attatctcag ctttctttg 3840
 tctaaaaaag ttattgcctt ttttttaa atatatcttc actggatatt gaatttagg 3900
 tgataatctt tttttttg tttagcattt aaatatgtct tctaagtcc tcttgcttc 3960
 atagttctg atgagaagtc tactgttatt agtatctctt tgtgtgtgtc tctctttt 4020
 ccctctctgc tattatggct attttttt tttttttt tttaggtcac tgggtgcagc 4080
 aatttaatta tgggtgtcct tggatgttt tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 4140
 tgtgtgtgtg tagctgatgt tcttgagct ttagaatctg tgagtttga gtttcatca 4200
 attattttt ctttcattc ctttattta ctcatgttcg tgtttattt tatatttta 4260
 agaattttgt gcgtatttgt aataactgtt taaatgtcat ttgtgaattc cattgcttct 4320
 aggtaggatt ctattgacag atatttttc cctgacgaga ggtcatactt tccttattct 4380
 tcatgtatct agtgggtttt ggttgaatac tggatatatt gaattttatg ggagtgtctga 4440
 attctacaat attccttaaa aatgtgttgg atttgtttt agcagatagc tatcttactt 4500
 gaagatcaat ttcatattt ttgatgttca tttttcatt tattaaagaa taggtccatg 4560
 gtagagtta ctgatatcaa ctttctggt gtctctaata aatgcaacat attcaataag 4620
 atcc 4624

<210> 25

<211> 3616

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(3616)

<223> MAR 1_68 of chromosome 1

<400> 25

gactctagat tataccaacc tcataaaata agagcatata taaaagcaaa tgctcttacc 60
ttgcagatcc ctgaactgag gaggcaagat cagtttggca gttgaagcag ctggaatctg 120
caattcagag aatctaagaa aagacaaccc tgaagagaga gaccagaaa cctagcagga 180
gtttctccaa acattcaagg ctgagggata aatgttaccat gcacaggggtg agcctccaga 240
ggctgtcca ttagcaactg ctacagtttc attatctcag ggatcacaga ttgtgctacc 300
tattgcctac catctgaaaa cagttgcttc ctatattca tccagtttaa tatttattta 360
aaccaagaag gttaatctgg caccagctat tccgttgtga gtggatgtga aagtaccaat 420
tccattctgt ttactatta actatccttt gccttaatat gtatcagtag gtggcttgtt 480
gctaggaaat attaaatgaa tggcatgttt cataggttgt gtttaaagtt gtttttgag 540
ttaaattctt cttaataat actttctgat gtcaaaaaca cttagaagtc atggtgttga 600
acatctatat agggttggat ctaaaatagc ttcttaacct ttctaacca ctgttttgt 660
ttgtttgtt ttaactaagc atccagtttg ggaaattctg aattagggga atcataaaag 720
gtttcatttt agctgggcca cataaggaaa gtaagatata aaattgtaa aatcgtaag 780
aactctatc ccatctgaag tgtgggttag gtgcctcttc tctgtgtcc ctaaacatcc 840
tattttatct gtatatatat atattcttcc aaatatccat gcatgggaaa aaaaatctga 900
tcataaaaat attttaggct gggagtgttg gctcacgcct gtaatccag cactttggga 960
ggctgaggtg ggcggatcat gaggtaaga gatcgagacc atcctgacca atatggtgaa 1020
accccatctc tactaaagat acaaaactat tagctggacg tgggtggcacg tgctgtagt 1080

SEL PCT 012.ST25

cccagctact cgggaggctg aggcaggaga acggcttgaa cccaggaggt ggaggttgca 1140
 gtgagctgag atcgcgccac tgcactccag cctgggcgac agagcgagac tctgtctcaa 1200
 aaaaaaata tatatatata tatatatata catatatata taaaatatat atatatacac 1260
 acatatatat ataaaatata tatatatata cacatatata taaaatatat atatatacac 1320
 acatatatat aaaatatata tatacacaca tatatataaa atatatatat acacacatat 1380
 atataaaata tatatatata cacatatata taaaatatat atatacacac atatataata 1440
 aatatatata tacacacata tatataaaat atatatatat acacatatat ataaaatata 1500
 tatatacaca catatatata aaatatatat atacacacat atatataaaa tatatatata 1560
 cacacatata tataaaatat atatatacac acatatataa aatatatata tacacacata 1620
 tataaaatat atatatacac atatataata aatatatata tacacatata tataaaatat 1680
 atatacacac atatataata aatatatata tacacacata tatataaaat atatatatat 1740
 acatatatat aaaatatata tatacacata tatataaaat atatatatat acacatatat 1800
 ataaaatata tatacacaca tatatataaa gtatatatat acacacatat atataaaata 1860
 tatatatata catatatata aaatatatat atacacatat atataaaata tatatatata 1920
 catatatata aaaatatata tatatatatt taaaatatatt ccaattgtct cactttgtgg 1980
 atgagaaaaa gaagtagtta gaggtcaagt aactggcct acatctttc tcaagattgt 2040
 aaactcctag tgagcaataa ccacatcttc attttcttg tataaaacaa gaaagtttag 2100
 catgaaaaag gtactcaatt acaaagtgtg tggattgaat tgaagaccct tggaagggga 2160
 tttgtacct gaggatctct ttctttggc catattgttc aatggacaaa atttagcctt 2220
 cgaaggcagg ccgatttgag gtaatacta cctttaccac ttgatagcta tgtgaccttg 2280
 gccatgtggt ttcaacagtc tgaacctcat ttctctgtg tatgtgtggt cctccttaca 2340
 agtttgtaa aaatgtgaag tccttagcca tgatagccca atataacagg ctaaataata 2400
 ataggtttat gtcttttcc ttatatctct cagataagca ctgtccaagt ttgaggtgtt 2460
 ttgaggtctc gcctgatttg gattgtttga gtttatgcta ttctttgaat tctttgagct 2520
 gttctgaagc agtgtatcat gaacaaaaac atccccagtt cagtccaaac ccctggttac 2580

SEL PCT 012.ST25

atatcattct tatgccatgt tataaccagt ttgagagtgt tccctctgtt attgcattta 2640
 agtttcagcc tcacacagaa attcagcagc caatttctaa gccctaagca taaaatctgg 2700
 ggtggggggg ggggatggcc tgaagagcag cattatgaat agcaccatta taattaatga 2760
 tctctcagga agatttacia tcacaggtag cagataaaac aaatagtact gcttctgcac 2820
 ttcccctct tttattcgct atgaaatttt atgggaaatc agtcagtgga aaaatgtaag 2880
 ctcttaatct tcccagaaa tcctacctca ttgatgaat actttgaggg aatgaattag 2940
 agcatttttt tcttttatag tctacttcgc atttacgaag tgaggacggg agcttaggct 3000
 gcctggccaa ctgatgagaa ggtcagaggc atttttagag acctctgttg tcttcattc 3060
 atgttcattt tccacaaggc aagtaatttc caacaaatca gtgtcttcat tagtaataag 3120
 attattaaca acaataatag tcatagtaac tattcagtga gagtcatta tatatcaggc 3180
 attctacaag gtactttata tacatctgag taaacctcac acaattctac agggagggtat 3240
 ttctatcccc atttaacaaa taaggaaacg aagtccaagt aaattaactt gcccaaggct 3300
 acacagatag tacctggcag aacaggaatt taaacctaaa ttgtccaac tccaaaagca 3360
 gccttctatt tgtataaat gctgcctctc attatcacat attttattat taacaacaac 3420
 aaacatacca attagcttaa gatacaatac aaccagataa tcatgatgac aacagtaatt 3480
 gttatactat tataataaaa tagatgtttt gtatgttact ataattctga atttgaatag 3540
 aaatttgcat ttctgaaagc atgttcctgt catctaatat gattctgtat ctattaaaat 3600
 agtactacat ctagag 3616

<210> 26

<211> 4660

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(4660)

<223> MAR 1_42 of chromosome 1

<400> 26

gatcccttga ggtcagtagt ttaagaccag cctgaccaac atggtgaaac ccatctctac 60
taaaaataca aaaattagcc aggcgtggtg gcgggggcct gtaaccccag ctactcagga 120
ggctgaggca caagaatctc ttgaacccgg gaggcggagg ttgcagtga ctgagattgt 180
gtcactgcac tccagcctgg gcaacagtgc cagactctgc cttaaaaaaa aaaaaaaaaa 240
aaaaaaggcc gggcgcggtg gctgacgcct gtaatcccag cactttggga ggccgaggcg 300
ggtggatcat gaggtcagga gatcgagacc acagtgaaac cccgtctcaa ctaaaatac 360
aaaaaattag ccgggcgcgg tggtagggcg ttgtagtccc agctactcag gaggctgagg 420
caggagaatg gcgtgaacct gggaggcgga gcttgcaagt agccgagatg gcaccactgc 480
actccagcct gggcgaaaga gtgagactcc gtctcaaaaa aaaaaaaaaa ttagctgggt 540
atggtggtgc gtgcctgtaa tcccagctac tcgggaggct gaggcaggag aatcccttga 600
acctgggagg cggaggttgc agtgatctgc catcctgtca ctgcatcact acactccagc 660
ctgggtgaca gagcgagact ctgtctcaaa aaaaaaaaaa aaaaaaaaag ctgggtgttg 720
tggtatgcac cagcttagt cccagctact tgggaggctg agttgggggg attgcctgag 780
ccaggagggt cgaggcttca gggagccatg attatgccac tgcactccag cctggggccac 840
agagtgaac ctctgtcaa aaacaaaaaa acaaaaaaac acagtgtgtt agatctgtct 900
agacttggtg atataattaa gaggccatta tgggcagaac tgtgccccct tccaaaattc 960
atatataaat atatataaat atatataaat atataaatat atataaatat ataaatatat 1020
ataaatatat aaatatatat aaatatataa atatataaa atatataaa atatataaaa 1080
atatataaat atatataaat atatataaat atataaaaac ataaaaatat atataaatat 1140
atataaatat ataaaaatat ataaatatat aaatatataa aaatatataa atatataaat 1200
atatacataa atatataaa atatataaa atatataaaa atatataaa atatataaat 1260

SEL PCT 012.ST25

atatatataat atatatataat atatatataat atataaaaat atatatataat atataaatat 1320
 ataaaaaatat atataaatat ataaatatat aaatatatat aaatatataa atatatataat 1380
 aaatataagt atttatgaat atatatgaat atataaatat ataaaaaata tatataaata 1440
 tataaatata tataaatata taaatatata catatataca tatataaata aataaatata 1500
 agtatattatg aatatatatg aatatataaaa tatataaaaa atatatataa atatatataat 1560
 atatatataat ataaatatat aaaaatatat aaaaatatat ataaatatat aaatatatat 1620
 aaatatataa atatatataa atatatataa atatatataat atatatataat atatatataat 1680
 atataaatat ataaatatat ataaatatat ataaatatat aaatatataa atataaatat 1740
 ataaatatat ataaatatat ataaatatat aaatatatat aaatatatat aaatatatat 1800
 aaatatatat aaatatataa atatatataa atatatatat aaatatatat aaatatataa 1860
 atatatataat atataaaaaat atataacaat atataaatat atataaaaaat atataacaat 1920
 atataaatat aaatatatat aaaaatatat aacaatatat aaatatataat atatatataat 1980
 atataaatat aaatatataaa aatatatata aatatataaaa tatatatataa tatataaatg 2040
 tataaatata tataaaaaata tataacaata tataaatata taaatatata acaatatata 2100
 aatatataaaa aatatataac aatatataaa tataaatata tataaaaaata tataacaata 2160
 tataaatata aatatatata taaatatata aatatataaata taaaaaatat atataaatat 2220
 ataaatatat atataaatat atataaatat ataaatgat aaatatatat aaatatataa 2280
 atatatataaa atatatataat atatatataat atatatataat atataaatat aaatatataa 2340
 atatatataa atatatataat ataaatatat aaacatatat aaatatatat aaataaacat 2400
 atataaagat atataaagat ataaagatat ataaatatat aaatatataa agatatataa 2460
 atatataaag atatatataat atataaagat atataaatat ataaagatat ataaatatataa 2520
 tatataaata tataaagata tataaatata atataaaaaat atataaatat atattaaaaa 2580
 tatatacata taaatatatg tatattttt tgagatgggg tctcgctcag ccaccacgc 2640
 tggagtgcag tggcagcagc tgggctcact gcaaccactg tctctgggt ccaagcaatt 2700
 ctgtctcagc ctccaagta gctgggatta caggcacctg ccatcatgcc cggctaattt 2760

SEL PCT 012.ST25

ttgtatttta gtagagatgg agttcacca tgttgccag gttggtctcg aattcctgac 2820
ctcaggtgat ctgccggcct cggcctccca gtgctgggat tacaggcatg agtcaccacg 2880
cccggcccta tatatatttt tgagacaagc tctgtgtctc ccaggctgga gtgcagcagc 2940
atgatcatga ctactgtag cctagacctc cagggctcaa gtgattctcc cacctcagcc 3000
tccaagtag ctgggattac aggcattgcac caccacaccc agctaatttt tgtttgttt 3060
tgtttggag acagaatctc tctctgtcac ccaggctgga gtgcagtggg gtgatctcag 3120
ctcagtgcaa cctccacctc ctgggttcaa gtgattctca tgcctcagcc tctgagtag 3180
ctgggactac aggcgtgagc caccacgccc tgataaattt tgtatttttt ttttcagatg 3240
gagtctcact ctgtcatact caggctggag tgcagtggcg tgattttggg ttattgcaac 3300
ctctgcttcc tgggttcaag cgattctcct gcctcagcct ccagagtagc tgggattaca 3360
ggcgctgcc accatgcca cctagctaac tttttttt ttttttga gatagagtct 3420
cactctgtca ccaggctgg agtgcaatgg ggcgatattg gctcactaca acctccacct 3480
cccaggttca agcgattctc ctgcctcagc ctctgagta gctgggatta cagggtgggtg 3540
ccaccacgcc agactaatat ttgtattatt agtagagacg gggtttcacc acattggta 3600
ggctggtttc gaactcctgg cctcagggtga tctgcctgcc tcggcctccc aaagtgtctg 3660
gattacaggc atgagccact gcggctggcc caatttttg attttttg tagagacggg 3720
ggtttacta tgcttccag gctgggtctca aactcctgga ctcaagcgat ctgcctgtct 3780
cagcctcca aagtgcaggg attacagtca tgagccacca ctgcacggcc caaaattta 3840
tttatttat tattattatt atttttaga tggagtctcc ttctgttgc agattggaat 3900
gcagtgccac gatctcagct cactgcaacc tccccctct gggatcaagt gattctttt 3960
ttttaagac tctgtctcaa aaaaaaagaa aaaaaaaaaa aatatatata tatatatata 4020
tatacacgaa tttgggcca ggcacagtgg cttatgcctg taatcccagc actttgggag 4080
ggccgagggtg ggtggatcac aaggctcagga gtttgagacc agcctggcca atatggtgaa 4140
accctgtctc tactaaaaaa taaaaaatt agctgggcgt ggtggcacga gcctgtaatc 4200
ctggctactc tggaggctaa ggcaggagaa tcgcttgaac cggggaggca gaggttgag 4260

SEL PCT 012.ST25

tgagccagga tcgcatcact gcactccagc ctgggtaaca gagcaagact ctgtctcaaa 4320
 aaacaaacaa aacaaaacaa aacaaaataa ataacgggtgc aaaattgaat atgccttttt 4380
 gactctctaa atgcctcaga tccatttacc ctggggattt gtcctttcta gccccaccac 4440
 catctcccct ctggaagact gctgacctat aaggataaag accagactct tgagcaggca 4500
 cttagggtct tctgcccac cccatcccc aactccccct cagtaatttt ggctactagt 4560
 atttctccac atctgaggct atcgtgggtc tcccttcagt ggtcatgaag gacaagggtg 4620
 gagaagtttg cctcgtgag tctgatgagg gattgggtgg 4660

<210> 27

<211> 3354

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(3354)

<223> MAR X_S29 of chromosome X

<400> 27

gatcccttta taaaaccaca atataatgga gtgctataat ttcaaacagt gtttggctctg 60
 ctggcagagt ggtcattcta acagcagtca cagtagagta gaaataagac tgcagtatat 120
 ctaaggcaaa aagctgaggt ttcaggagct tgaaggtaaa gaggaagaaa gaaatgggaa 180
 tgggaattgg aaagacaaat atcgttaaga gaaaattgct tttaggagag gggaaagaat 240
 ctatgtgtac ttaagactat ggaatcaatc ccatttaagc tgggaaacta gtttcatata 300
 taactaataa attttattta cagaatatct atttacctga tctaggcttc aagccaaagg 360
 gactgtgtga aaaaccatca gttctgtcat attcctaaaa aaaaattaaa aagttaaaaa 420

SEL PCT 012.ST25

taaataaata ataaaacttc ttttcttca aaataatcaa ggtgcttatt cacatccatt 480
 ccaatttggg gaaatactta ttttctatg attagcgaag agaaaagtaa ctgcatctc 540
 aattcaagtt gatacatgtc actttaaga ggtcaactaa tatttgctag ttgagctaac 600
 catataggct ttaaatactt tcatagtaga aagaaaatga aaatcattag tgaactgtat 660
 aaaatagatc atacttttg aaagaatcag actgaagttt ccgaaaaaaaa gaagtaagct 720
 tcaatgaaaa ggtaagtga tttagcattt actcagcatc tactatggac ttaacaccta 780
 acagtagata atctgaaggc aaacatattt gtatagggac tgcagaatga tagatgataa 840
 atatcatctc ttctatttga atgaatattt ttcaaatct ttcacacaca gtggtttgct 900
 atggaaagat tttagtaga ttaaacaat ctgaagatgg agttagaaag cttaggctat 960
 gtttgagca caacatataa tttctctgtg attgtttctt catctttcaa atgaggttac 1020
 tgtgaagatt aatgagata actaaatgat gataaaataa tgaatctta gcagcacctt 1080
 attaatctg tgcaacaact ctgtgaagt agtagggctc agcttcagtc acttctctgc 1140
 catttattaa ctaagatagt ttggaagtt acccatctct tcagctgtaa aatgatgagg 1200
 atcacaccta tttatgggg ctgcttttag gtacaaatat acaggcaagc actttgttaa 1260
 tactaaagca ttacaccaat tagttttact cttttccatt cacacatgaa attaatgtaa 1320
 tcagaattct gtagattacc taaatcttct gttaacacgt gatatgcagt tcagggtaaa 1380
 tgtcagttga gttaccaaag cacatacata ctccaccacc tatccaaatc tacaagcctc 1440
 ccagtttgc ttactattt tggttaaatt aatatgaatt cctagatgaa aatttcactg 1500
 atccaaatga aataaaaaat atattacaaa actcacacct gtaatctcaa cattttgga 1560
 ggccaaggca ggtagatcac ttgaggccag gagttcaaga ccagcctgat caacatggtg 1620
 aaaccctgtc tctactaaaa atacaaaaat tagccagggtg tgggtggcatg tgctgtagt 1680
 cctacctact cgggaggctg aggcacaaga atcgcttgaa tggggagggt ggagggtgca 1740
 gtgacctgag atcgtgccac tgcactccag cctaggcaac agagtgagat catgtgtcat 1800
 atatatatat atatatatat atatatatat atatatatat acacacacac acatatatat 1860
 atacacatat atatacgtat atatatatat gtatatatat acatatatat acatatatat 1920

SEL PCT 012.ST25

atatacgtat atatacgt atataatat caatgtaaat tatttgggaa atttgggtatg 1980
aatagtcttc cctgtgaaca cagatcataa aatcatatat caagcagaca aataagtagt 2040
agtcacttat atgcttatac ttgtaactta aagtaaaaga attacaaaag catatgacaa 2100
agactaattt taagatatcc taatttaa atgttttctaa aagtgtgtat accattttac 2160
ctatcatatg aataatttag aaacatgttt ataaaattaa tgcctaaatc cattcaaaag 2220
tttgtaatg cagatcacc acaacaacaa agaatcctag cctattaaaa aagcaacacc 2280
acctacatat aatgaaatat tagcagcatc tatgtaacca aagttacaca gtgaatttgg 2340
gccatccaac actttgagca aagtgttgaa ttcacaaat gaatgtgtaa tcatttactt 2400
actaatgcca atacacttta aggtaatctt aagtagaaga gatagagttt agaattttt 2460
aaatttatct ctgtgtgaa agcaatagac ttgaataaat aaattagaag aatcagtc 2520
tcaagccacc agagtatttg atcgagattt cacaaactct aactttctga taccattct 2580
cccaaaaacg tgtaacctcc tgcgatagg aacaaccac tgcagggatg ttctcgtgg 2640
aaaaaggaaa ttcttttgc attggtttca gacctactg gttacaagaa aaaccaaagg 2700
ccattgcaca atgctgaagt actttttca aatttaaaat ttgaaagtg ttcttaaaat 2760
ctatcattta tttaaaata cggatgaatg agaaagcata gatttgataa agtgaattct 2820
tttctgcaat ctacagacac ttcaaaaat cactacagac actacagaca ctacagaaaa 2880
tcataaataa acaagtgcta gtatcaatat ttctacaaa aaatggcatt cttagaattt 2940
tttataggct agaagggttg tacaaactaa tctgccacgg attttaa atgagtgaat 3000
aaattatatt gcaaaaaaaaa tcaggttaca gagaactggc aaggaagact ctatgtaaa 3060
acacagaaaa catacaaaac gtatttttaa gacaaataaa aacagaactt gtacctaga 3120
tgatactgga gattgtgtg acatattagc attatcactg tcttgctaaa acataaaaat 3180
aaaaagatgg aagatgaaat tacaatacaa atgatgattt aaacatataa aaggaaaata 3240
aaaattgttc tgaccaacta ctaaaggaag acctactaaa gatatgcat ccagcacatt 3300
gccactctac atgtgtctg taaaccagca gcataggat cctctagcta gagt 3354

SEL PCT 012.ST25

<211> 677

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(677)

<223> MAR of chromosome 1 genomic contig; 12803267..12803943

<400> 28

```

ttatatagta tatataatag tatatattat atagtataca taatagtata tattatatag   60
tatacataat agtatatatt ataataataca taattgtata tatcatatag tatacattat   120
agtatatata atagtatata cattatagta tatatcatat agtatacatt atagtatata   180
tatcatatag tatacattat agtatatata atagtatata cattatagta tatatcatat   240
agtatacgtata atagtatata tcatatagta tacgtaatag tatatatcat atagtatacgt   300
taatagtata tatcatatag tatacgtaat agtatatata atagtatata cgtaatagta   360
tatatcatat agtatacgtata atagtatata tcatatagta tacgtaatag tatatatcat   420
atagtatacgt taatagtata tatcatatag tatacgtaat agtatatata atagtatata   480
tattatatag tatatatcat atagtatata ttatatagta tatatcatat agtatatatt   540
atagtatata tatcatatag tatatattat atagtatata tcatatagta tatataatag   600
tatatatcat atagtatata taatagtata tatcatatag tatatatata atactatatt   660
atatatagta tacataa

```

677

<210> 29

<211> 332

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(332)

<223> MAR of chromosome 1 genomic contig; 13079684..13080015

<400> 29

tttaattatat tatatatatt atataattat atattaatat atattaatta tattatatat 60

attatataat tatatatataa tatatatataa ttatattata tatattatat aattatatat 120

taatatatat taattatatt atatatatta tataattata tattaatatata tattaattat 180

attatatata ttatatatta taattatata ttatataatt ataatatata tgtaatatata 240

atatatatataa ttaatatata attaaaacta tttaattata tgtatattat atataatatg 300

tattatttaa ataataaata tattatttat at 332

<210> 30

<211> 479

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(479)

<223> MAR of chromosome 1 genomic contig; 15682296..15682774

<400> 30

SEL PCT 012.ST25

acaagtacat atatatatag tatatatata caagtacata tatatagat atatatatat 60
 acaagtacat atatatagta tatatatata tacaagtaca tatatatagt atatatatat 120
 acaagtacat atatatagta tatatatata caagtacata tatatatagt atatatatat 180
 acaagtacat atatatagta tatatatata caagtacata tatatatagt atatatatat 240
 acaagtacat atatatagta tatatatata caagtacata tatatatagt atatatatat 300
 acaagtacat atatatatag tatatatata tacaagtaca tatatatata gtatatatat 360
 atacaagtac atatatatag tatatatata tatatacaag tacatatata tagtgtatat 420
 atatatatac aagtacatat atatacttgt attagtatat atatatatat atacaagta 479

<210> 31

<211> 531

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(531)

<223> MAr of chromosome 1 genomic contig; 15694611..15695141

<400> 31

tataatatat ataatacata atagatatat tatattatat aatagatata taattataaa 60
 cataataata tataatgaat ataataataa ataaatataa taaaatatat aatatatcta 120
 ttatgtatta tatattatat atgtttatat ataataataat tatatatggt tatatataat 180
 ataattatat atgtttatat ataataataat tatatattat atattataga tataatatat 240
 aatatactat atattataga tataatatat aatatactat atattataga tataatatat 300
 aatatactat atattataga tataatatat aatatactat atattataga tataatatat 360

SEL PCT 012.ST25

aatatactat atattataga tataatatat aatatatatt atatattata gatataatat 420

ataatatatt atatattata tctatatata atatattgta tattatatat aatatattgt 480

atattatata taatatattg tatattatat ataatatatt gtatattata t 531

<210> 32

<211> 378

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(378)

<223> MAR of chromosome 1 genomic contig; 886276..886653

<400> 32

ttatattata tatcttacct aaattatata tatatattac ataaattata tacaatataa 60

attatatata atataattta tatataaaat ataaattata taaataattt atatataaaa 120

tataaattat ataaataatt tatatataaa atataaatta tgtataaaat ttatatataa 180

aatataaatt gtgtataaaa ttatatataa aatataaatt gtgtataaaa ttatatata 240

aaatataaat tatatataat ttatatatta taatataaat tatatataat atatataata 300

aaatataaat tatatataat atatataata agatataaat tatatataat atatataata 360

agatataaaa tatataat 378

<210> 33

<211> 595

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(595)

<223> MAR of chromosome 1 genomic contig; 3326732..3327326

<400> 33

```

aaaatatata aatatatata aaaatatata aaaatatata aatatatata aaaatatata   60
aatatatata aatatatata aaaatatata aatatatata aatatatata aaaatatata   120
atatatataa aatatatata aatatatata aatatatata aaaatatataa tatatatataa   180
aatataaata tatataaata tatataaaaa tataaatata tataaatata tataaatata   240
taaatatata taaatatata taaatatata aatatatata aatatatata aatatatata   300
aatatatataa tatataaaaa tatatatataa tatataaata tatataaata tataaatata   360
taaaaatata tataaatata taaatatata taaatatata taaatatata tataaatata   420
tataaatata tatatatata aatatatata aatatatata taaatatata taaatatata   480
tatatatata taaatatata taaatatata taaatatata tataaatata tataaatata   540
tataaatata tatataaata tatataaata tatatatataa tatatatataa tatat      595

```

<210> 34

<211> 738

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(738)

SEL PCT 012.ST25

<223> MAR of chromosome 1 genomic contig; 4485716..4486453

<400> 34

ataatagata atatatatta tatgatagat atataatata ttatataata tataatatat 60
 tatatatcta tcatataata tatataatat ataatatatt atatatctat catataatat 120
 aatatatata atataataa tatatcatat tatattgtat ataatatata tcatattata 180
 ttgtatataa tatatatcat attatattgt atataatata tatcatatta tattgtatat 240
 aatatatatac atattatatt gtatataata tatatcatat tatattgtat ataatatata 300
 tcatattata ttgtatataa tatatatcat attatattgt atataatata tatcatatta 360
 tattgtatat aatatatatac atattatatt gtatataata tatatcatat tatattgtat 420
 ataatatata tcatattata ttgtatataa tatatatcat attatattgt atataatata 480
 tatcatatta tattgtatat aatatatatac atatatattc tattatattg tatataatat 540
 atattatata ttatctatta tattgtatat aatatatatt atatatattc tattatattg 600
 tatataatat atattatata ttatctatta tattgtatat aatatataat aaatatagta 660
 tatataatag ataatatata gtatatatga tatatttatat atactatata ttatatatca 720
 tatatactat atactata 738

<210> 35

<211> 386

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(386)

<223> MAR of chromosome 1 genomic contig; 5423067..5423452

SEL PCT 012.ST25

<400> 35

taaatatata aaaatatata taaaaatata aaaatatata tataaatata taaaaatatt 60

tatataaata tataaatata taaatatata ttatatataa tatataaata tataaatata 120

taaatatata ttatatataa tatataaata tatatttata taaatatata aatatatata 180

aaatatataa atatatattt atataaatat ataaatatat ataaaaatata taaatatata 240

tattttatat aaatatataa atatatataa aatatataaa tatatatatt ttatatataa 300

atataaatat atataaaata tataaatata tatattttat atatttatat atataaatat 360

atatatttca tatatcacat atatga 386

<210> 36

<211> 584

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(584)

<223> MAR of chromosome 1 genomic contig; 5805559..5806142

<400> 36

taaatatattt taaaatatat atattttata atatatattt tatattataa tgtgtacata 60

atatatatata taatatataa tatataaatat tgtatatattt attatatata ttataatata 120

tattattata tatttatatta tatataaatat aatatatatatt ataatatatt atattatata 180

tattataatg tattataata tatatttatat tatatatattt aatatatatatt atattatata 240

ttataaatata tatttatatta tatatttataa tatatatattt attatatattt atatatata 300

SEL PCT 012.ST25

atacatatta taatacatat tatataatat attataatat gtattataat acatattata 360
 taatatatta taatatatta tatataataa tatattataa tacatattat atataatata 420
 tattatgtat attatatata atatatatta caatgtatat tatgtatatt atatatatta 480
 tatatcatat aatatatatt atataataa tgatatataa tatatattat ataatatatt 540
 atatgatata tataatatgt attacatgta atatatatca taat 584

<210> 37

<211> 345

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(345)

<223> MAR of chromosome 1 genomic contig; 10802644..10802988

<400> 37

tgtatatata tactatatat atactatata tatagtgtat atatatacta tatatatact 60
 atatatatag tgtatatata tactatatat atagtatata gtatatatag taatatatat 120
 atatagtata tatatacact atatatagta tatatagtat atatatatgg tgtatatagt 180
 atatatatag tgtatatata gtatatatat attgtatata tagtatatat attgtgtata 240
 tatagtatat atatagtata tatagtatat atagtatata tatagtatat atatactata 300
 tatatagtat atatatatgg tatatatata ctatatatat agtat 345

<210> 38

<211> 474

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(474)

<223> MAR of chromosome 1 genomic contig; 13496468..13496941

<400> 38

atattatata taaataaatt atatctataa ttatatatta tatataatat aattatatat 60

ctataattat atattatata taatatatat tatatataat atataaattat atataattta 120

tataatataa tatataatat ataattatat ataattatat aatataatat ataatatata 180

attatatata atttatataa tataatatatat aatatataat tatatatatt tatataatat 240

aattatatat aatatataat tatatataat ttatataata taattatata taatatataa 300

ttatatataa ttatatataat ataattatat ataattatat atttatata atttatataa 360

tataaattata tataatatatat aattatatat aatatataat tatatataat tatatataat 420

atataaattat atataaattta tataatataa ttatatatta tatatatatt atat 474

<210> 39

<211> 483

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(483)

<223> MAR of chromosome 1 genomic contig; 2509163..2509645

SEL PCT 012.ST25

<400> 39

caaaatacat aatatataat agtattatat aatagtatgt atagttataa tatatagtat 60
aattacaata tatgatatgg ttatatatt atatatagta taatataata taacataata 120
ctattataat atataaacta tataatatat actattataa tatatgaact attataatat 180
ataaactata tataatatat aatatgtact attataatat ataaactatt ataataatat 240
atataaacta ttataatata taaactatta taatatatat aatactatgt atacatatat 300
tacattatgt acatactaca ttatgtatta tgtatgtata tatacacaaa atacataata 360
tataatagta ttatataata gtatatatag ttataatata tagtataatt acaatatata 420
atatggttta tatattatat atagtataat acaatataac ataatactat tatatataaa 480
cta 483

<210> 40

<211> 641

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(641)

<223> MAR of chromosome 1 genomic contig; 2776349..2776989

<400> 40

tgttatatat atataacata gatattatat atacatgtta tatatataac atagatatta 60
tatatacatg ttatatatat aacatagata ttatatatat aacatagata ttatatatac 120
atgttatata taacatagat attatatata catgttatat ataacagata ttatatatac 180

SEL PCT 012.ST25

atgttatata taacatagat attatatatg tatgttatat ataacataga tattatatat 240
 gtttatataa tatataacat atgtttaaca tatataatat ataacatggt tatataatat 300
 ataacataat tatatgttat atatgatata aaacatatat attatatacg ttatatgtaa 360
 tatataacat atattgtata cgttatatgt aatatataac atatattgta tacgttatat 420
 gtaatatata acatatattg tatacgttat atgtaatatata taacatatat tgcatacgtt 480
 atatgtaata tataacatat attgtatacg ttatatgtaa tatgtaacat atattgtata 540
 cgttatatgt aatatgtaat atataatata tataacatgt atatataaca tatatgtata 600
 taacatatat ataacatatata taacatatat gttatattat a 641

<210> 41

<211> 745

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(745)

<223> MAR of chromosome 1 genomic contig; 2858703..2859447

<400> 41

atatttatat atgtaataat atataatata tttatatgta ttgtatatg taataatata 60
 tatataataa aatatgtaat aatatataat atatttatat ataaatatat tatattatat 120
 atatattatt atatttataa tataatatat atttatatta tatattataa atatattatta 180
 tataatatat attataaata tatattatat aatatatatt ataaatatat attatattat 240
 atattataaa tatattattat ataatatata ttataaatat atatttatata atatattatta 300
 taaatatata ttatatttat aatatatatt ttgtatatatt atatattata tattataaat 360

SEL PCT 012.ST25

attattatat ttataatata ttatatatt tatatataat atatgatata tattataaat 420
 atatcttata aatatatata ttatatata tatattataa atatataaat ataaatatat 480
 aatataatat aatataatat aataaatata atatataata tatataatat ataataaata 540
 taataaatat aaatatatca tataaatata aatataaata taaatatatc atataaatat 600
 atatatttat atgatatatt atagtatata taaatatatt tatataattat aaaatattta 660
 tataatatat aattataata tatttatata tataaattaa ctaatatata taaactaata 720
 taatatataa tgtaataata tagta 745

<210> 42

<211> 307

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(307)

<223> MAR of chromosome 1 genomic contig; 945522..945828

<400> 42

catatataat atatattacc tatgttatat aggtcatata taacataaat atattacata 60
 tatgtaatat atattaaata taaatatata acatatatgt gtaactatat atgtaaatat 120
 gtacatatat atatatgtaa atatataata tatatttaca ttatattata taatatatat 180
 ttacattata tatttatata tacattatat atatttacet tataaatatt tatataatat 240
 atatttacet tatattacat tatataaaat acaatatatt acattataat acattataac 300
 agataaaa 307

<210> 43

SEL PCT 012.ST25

<211> 357

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(357)

<223> MAR of chromosome 1 genomic contig; 3402743..3403099

<400> 43

aatatttat taaatataat atattaatat ttaatatt taatataata ttaaataaat 60

atattataaa taaattataa tatataaata tatattatgt atttatgtat aatatataaa 120

aattatatat aatatatata tttttataaa tatataaata tataataaat aaatatatta 180

aataaataat aatatattaa atattaatat attaaatatt atatattaaa tataatatgt 240

aatatgaaat atattaaata ttatatatta aatataatat ataattgtaa atatattaaa 300

tattatatat taaatataat atataatatg aaatatatta aatatttat attaaat 357

<210> 44

<211> 323

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(323)

<223> MAR of chromosome 1 genomic contig; 3485830..3486152

SEL PCT 012.ST25

<400> 44

atatttatag actatatatt tatatatffa gtgtatttgt atactatata tttatatagt 60

tagtatattt gtatactata ttttatata tttagtatat ttgtatacta tatattfata 120

tatttagaat atttgatac tatatatffa tatatttagt atatttgat actatatatt 180

tagtatattt gtatactata ttttatata tttagtatat ttgtatacta tatattfata 240

tatttagat atttatatac tatatactta tatatttagt atatttatat actatafact 300

tatatatffa gtatatttat ata 323

<210> 45

<211> 498

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(498)

<223> MAR of chromosome 1 genomic contig; 3548336..3548833

<400> 45

aattattact atattgttaa tataattatt atataatata atataattat atcactatta 60

ttatattata gtattaatat aatagtgtat aacattaata taatatagta ttaatataat 120

agcgtataac attaatataa tatagtatta atataatagc gtataacatt aatataatat 180

agtattaata taatagtgtat tattaatata atatagtatt aatatataat attaatataa 240

tatatcaata taatagtata taatataata taatatatca atataatagt atataatata 300

atataatata tcaatataat agtatataat ataataat atatcaatat aatagtatat 360

SEL PCT 012.ST25

aatattaata taatataata tcaatataat agtatataat attaatatat taatataata 420
 gtatataata ttaatgtaat ataattattaa cataatgtat ataataatat aatagtatat 480
 aatactaata taatataa 498

<210> 46

<211> 400

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(400)

<223> MAr of chromosome 1 genomic contig; 4595109..4595508

<400> 46

aaatatatta tattatatat tatatattat tcaatataact ataatatata ttatatatgt 60
 ttaatacaat atataatatt tacatatatt cccatttatt tatataacat atattatatg 120
 atattatata ttactccata taatataata tattatacat aatatattac tcagtataat 180
 acataatata tataatatat tactcggtat aatatataat attatatggt atgcaatata 240
 atatataata ttatatataa tacattattc aatataatat ataattattat atataataca 300
 ttattcaata taatatataa tacactattc aatataatat acaattattat atataataca 360
 ttattcaata taatatatat tatataatat atatatttat 400

<210> 47

<211> 403

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(403)

<223> MAr of chromosome 1 genomic contig; 7205509..7205911

<400> 47

agtatatata tgtgtatata tatgagtata tatatgtgta tatatatgag tatatatatg 60
tgtatatata tgagtatata tatgtgtata tatatgagta tatatatgtg tatatatatg 120
agtatatata tgtgtatata tatgagtata tatatgtgta tatatatgag tatatatatg 180
tgtatatata tgagtatata tatgtgtata tatgagtata tatatgtgta tatatgagta 240
tatatatatg tgtatatatg tgagtatata tatgtgtata tatatgagta tatatgtgta 300
tatatatgag tatacatatg tgtatatata tgagcatata tgtgtatata tatgagtata 360
tatatgtgta tatatatgag tatatatgtg tatatatatg agt 403

<210> 48

<211> 309

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(309)

<223> MAR of chromosome 1 genomic contig; 7507280..7507588

<400> 48

SEL PCT 012.ST25

tataaaatat atattattta tatattatat ataaaatata tattatatta tatattatag 60
 atataataaa taaataatat ataatatatt atataattat ttatacataa ttatatataa 120
 ttatatgtaa ttgtacaatt atatataatt atatacaatt atacacataa ttatatacaa 180
 ttatacaatt atatacataa ttatatatat aatatacata attatatatt aattatacaa 240
 ttatatocat aattatatat aattatacaa ttatatocat aattattatg tatattatat 300
 tatataata 309

<210> 49

<211> 516

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(516)

<223> MAR of chromosome 1 genomic contig; 3581085..3581600

<400> 49

atatatatat atatatatat atttatatat atatatatatta atatatatatta tatataaaaa 60
 tatataaaaat ttatatatat aatttatata tataaaaata tataaaattt atatatatata 120
 ttatatata taaaatatata taaaatttat atatataatt tatatatata aaaatatata 180
 aaatttatat atataattta tatatatataa aatatataaa atttatatat ataatttata 240
 tatataaaaa tatataaaaat ttatatatat aatttatata tataaaaata tataaaattt 300
 atatatatataa ttatatata taaaatatat aaatttatata tataattata tatataatat 360
 aaaatttatat atataattat atatataata taaaattata tatataatta tatatatataat 420
 ataaaattat atatatatig tatatatata aatatatacaa aatttatata tataaaaatat 480

SEL PCT 012.ST25

aaaatataca taaaaataaaa tatatataat ttatat 516

<210> 50

<211> 534

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(534)

<223> MAR of chromosome 1 genomic contig; 3084851..3085384

<400> 50

atataatata tatgactata tattttatat tataattctat ttcaataaaaa tatttatatt 60
 ttattatata ttataatata taattatata tgtaataata tataatataat aatatatatt 120
 ttatattata ttttatattt atttttatat ttatattat attttattat atatattata 180
 atatataatt atatatgcaa taatatatta tatattataa tatataatta tatatgcaat 240
 aatatattat atattataat atataattat atatgcaata atatattata gattataata 300
 tataattata tatgcaataa tatattatat attatatatt agataatata ttaatatata 360
 ttataacata taatatataa catataatat ataatatatt atctaataata taatataaca 420
 tataatataat aatatattat ataatatatt attacatata taatatattg taatatataa 480
 tattacatat atcttcaaaa agagttatgt gtatataata catatatata ccat 534

<210> 51

<211> 583

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(583)

<223> MAR of chromosome 1 genomic contig; 160087..160669

<400> 51

tatttatata aaatatataa aatatattat atataaatat attatatata atatatttat 60

atattatata atatatttat atatttatata taatatattt tatataatat acataatata 120

ttttatatat tatatataat atattttata tataatgtac aatataattt atatattata 180

tataatatat tttatatata ctatacaata tttttatat attatatatt ttatatatat 240

ttttcatgta acatatatat tttatatata atatatatac catatataat atattttata 300

tataatatat ataccatata taatatattt tatatataat atgtatatca tatatagtat 360

attttatata taatagggtat accatatata atatatttta tatataatag gtataacata 420

tataatatat tttatatata atagtatac catatataat atattttata tattatagat 480

accatatgta atatacttta tatataatat agataccata tgaatatatac tttatatata 540

atatagatac catatgtaat atactttata tataatatag ata 583

<210> 52

<211> 314

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(314)

SEL PCT 012.ST25

<223> MAR of chromosome 1 genomic contig; 4350424..4350737

<400> 52

```
tatgtgtata taaatatatg tatatatgtg tatataaata tatataaata tatgtatata    60
tgtatata catatattta tatataaata tatgcatata tttatatata aaatatatgc    120
atatatgtat atatataaaa tatatacata tatgtatata tataaaatat atacatatat    180
gtatatatat aaaatatata catatatgta tatatataaa atatatacat atatgtatat    240
atataaata tatacatata tgtatatata taaaatatat acatatatft atatataaa    300
aataccaagt cttat                                     314
```

<210> 53

<211> 828

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(828)

<223> MAR of chromosome 1 genomic contig; 8443267..8444094

<400> 53

```
tattatata ttatatatac tatataatta tataatatat agttatatag tatatataat    60
atatataata tatactatag tatatataat atatataata tatactatag tatatataat    120
atataattat atataatata tataatatag tatatattat atatataffa tatatatata    180
atatatatat aatatatata atatagtata tataatatat aattatatat aatatataat    240
atagtatata taatatataa tatatatata attatatact ataatatata taatatataa    300
```

SEL PCT 012.ST25

ttatatatta tatactatag tatatattat tatatataat agatataata tatataatta 360
 ttatataata tagtatatat aatatataat tatatataat agatataata taatataatt 420
 atatataata tagtatatat aatatataat tatattatat tatatataat atataaattat 480
 aatatataat tatattatat aatatatata atatataatt atattatata attatattat 540
 ataatatata taatatataa ttatattata taatatatat aatatataat tatattatat 600
 aatatatata atatataatt atattatata atatatataa tatataatta tattatataa 660
 tatatataat atataaattat atattatata taatatagta tatataatat gtaattatat 720
 atcatataat atataacatt gtatataata tataattaca tattatataa tgtatataat 780
 atataaattat atacattata taatatagta tataattata tattatgt 828

<210> 54

<211> 573

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(573)

<223> MAR of chromosome 1 genomic contig; 8703190..8703762

<400> 54

tatattatat ataaaatata catataatat acctataata tacatataat atataatata 60
 tattatgtac atataatata catataatat atataatata taatgtacat ataatatata 120
 tataatatat gttatatatt atatataaaa tataggatat atataatata gaatatatat 180
 actatattgt atatataaga tatataatat atagtatata tactatataa tatataatat 240
 atagtatata taatatataa tatagaatat atatacaata tataatatag aatataggat 300

SEL PCT 012.ST25

atatatagaa tatacatata taatatgtat atattatata ttatattata tattatataa 360
 aaatatataa tatataatat aaaaatatat tatatattat ataataaaa atatattata 420
 tattatatat tatataatat aaaaatatatt atatattata tattatatat aaaaatatatt 480
 atatattata tattatatat aaaaatatat tatatattat atattatata taaaaatata 540
 ttatatatta tatataaaaa tatatattat tac 573

<210> 55

<211> 597

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(597)

<223> MAR of chromosome 1 genomic contig; 8819076..8819672

<400> 55

acatatotta tatataaaat atataaatat acacatattt tatatataat atatattata 60
 tatatgaaat atacacatat ttttatatat ataatatata tattatatat aatatatgca 120
 tatattatat ataaaatata tatattatat ataaaatag catatattat atataatata 180
 tataatataa aatatataat atatattata tattatatat aatatatatt atatataata 240
 catatatata atatataata tatataaaat ataatatata tattatataa tatatatata 300
 aatatatata atatatatat aatatatata ttatatataa aatatatatt atagtataaa 360
 tatataatat atataatata tatattatat gtaaaatata tattatatat aaaaatatata 420
 atatataaaa tatatattat atataaaata tataatatat aaaaatatata atatataata 480
 aatatataat atatataaat atatattata tataaaaatat ataatatata taaatatata 540

SEL PCT 012.ST25

ttatatataa aatatataat atatatataat atatattata tataaaatat atattat 597

<210> 56

<211> 646

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(646)

<223> MAr of chromosome 1 genomic contig; 759619..760264

<400> 56

taatatatat aatatatatt atataataat atataatata tattatatta taatatataa 60

tatattatat aataatatat attatataat atataataat atatataata catattattt 120

aataatatat aatatatatt atataataat atataatata tattatataa taatatatat 180

tatattatat aatatataat atatataata tatattatat aataatatat aatatatatt 240

atagaatgat atattagata ttatataatt atatatataa tattatatat tatataataa 300

tatataatat atattatata attatataa taatattata tattatataa ttatatataa 360

tatattatat aattatatat ataattattat atattatata attatataa atatatatta 420

tataattata tatataatac tatatattat ataattatat ataatactat atattatata 480

atttatataa ttatatatat tatatattat ataattatat atattatata ttatataata 540

acatatatat tatatattat ataataacat atatattata tattatataa tacatatata 600

ttatatatta tataatacat tattatataa tatataatat atatta 646

<210> 57

<211> 752

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(752)

<223> MAR of chromosome 1 genomic contig; 1226710..1227461

<400> 57

taaacatata tataaatata tataaatata tatataaata tatataaata tataaatata	60
taaatatata tgaatatata aatatatata aatatatatg aatatataaa tatatatata	120
aatatatata aatatatata taaatatata taaatatata taaatatata taaatatata	180
taaataaata tataaatata tataaatata taaatatata tataaatatg taaataaata	240
tatataaata tataaatata tataaatata tataaatata tatagaaata tatatagaaa	300
tatatataaa tatatataga aatatataga aatatatata gaaatatata taaatatata	360
taaatataga aatatatata aatatatata aatatatata gaaatatata atatataaa	420
atatataaa atatataaat atatataaa atatatatat aaatatatat aaatatatat	480
aaatatatat aaatatatat aaatatatat attaatatat aaatctatat taatatatat	540
taatatataa atctatatta atatatatta atatatatat taatatatat taatatataa	600
atatatatat taatatataa atatataaa atatatatgt aaatatatat ataaatatat	660
ataaatatat atataaatac atataaatat atatataaat atatataaat atatataaa	720
atatataaa atatatatat aaatatatat aa	752

<210> 58

<211> 300

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(300)

<223> MAR of chromosome 1 genomic contig; 1119049..1119348

<400> 58

taatatacat ttatataat atatgtaata tatatfttat atatatgtaa tatatatftt 60

atataatata tgtaatatat attttatata tatgtaatat atattfttata taatatatgt 120

aatatatatt ttatataata tatgtaatat atattfttata taatatatgt aatatatatt 180

ttatataata tatgtaatat atattfttata taatatatgt aatatatatt ttatataata 240

tatgtaatat atattfttata taatatatgt aatatatatt ttatatatat gtaatacata 300

<210> 59

<211> 617

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(617)

<223> MAR of chromosome 1 genomic contig; 3603613..3604229

<400> 59

aaaatataat atatataata tataatatat ataatatatt atatataaaa tatataatat 60

SEL PCT 012.ST25

ataatatata taataaaata tacataatat ataattgata ataaaatata cataatatat 120
 aatatataat aaatatataa tatataatat ataataaaat atataatata taatatataa 180
 taaaatatat aatatattat atataataaa atatataata tattatatat aataaaatat 240
 ataatatatt atatataata aaatatataa tatattatat ataataaaat atataatata 300
 ttatatataa taaaatatat aatatattat atataataaa atatataata tataatatat 360
 aataatatat ataatatata atatataata taaaatatat ataatatata atatataata 420
 taaaatatat aatatataat atatataata aaatatatat gatataatat atatataata 480
 aaatatatga tatataatat atataataaa atatataata tataatatata tatataatat 540
 atatactaaa aaatatataa tatataataa aaaatatata atatataata tatataatat 600
 ataataaaat atatata 617

<210> 60

<211> 674

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(674)

<223> MAR of chromosome 1 genomic contig; 2592460..2593133

<400> 60

taagcttata tatatatata agcttatata tatatatata agcttatata tatatagaaa 60

gcttatatat atatagaaag cttatatata taagaagctt atatataaaa gcttatgtat 120

aaatatatat aaatatattt atttatgctt atagatacat atataaatat atttatttat 180

atttatatat aaacatatat ttatatatat ttatatataa ttattttatt atataaataa 240

SEL PCT 012.ST25

atatataata aataataaat atatataata tatttattgt attatttata taaatttatt 300
 aatataatat ataataaaat aataattata taaatatata aatatctata aatatatata 360
 aatatatata atatctataa atatatatata atataaatat atataaatatc tataaatata 420
 gataaatata aatatatata atatctataa atagataa atataaatat atataactat 480
 atataaatat atataactat atataaatat atatataaat atataaact atatatataa 540
 ctatatatat aaatatatat aactatatat ataaatatat atataaatat atataactat 600
 atatataaat atataaact atatataaat atatatatata atatatatata ctatatatat 660
 aaatatatat ataa 674

<210> 61

<211> 1694

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1694)

<223> MAR of chromosome 1 genomic contig; 2891680..2893373

<400> 61

atattgaata catatattat atattgcatat atacattgcat atgtatatat atatattata 60
 tatgcatata tacattgcat tgtatatata tatataaagt atgattatat ataatatata 120
 catgtatatg tatatacatg tatatattat attatatatt attatacat attattatgt 180
 ctatatataa tataatatat acatattaat aatataatc ataatatat ataatatatt 240
 atataatata taatatata taatatatta tataatacat aatataat atatatattat 300
 atgatacata atataatata atattattata tgatacataa tataatatata tacatatata 360

SEL PCT 012.ST25

taatatatta ttattattaa tataatatat acatatata atatatata 420
 ttatatataa tatacatata atataatag taatattata tataatataa tacataatat 480
 aatacatatt aataatatat tattaataag ataatatata tgtatctata atatatatat 540
 atagtatat gtagtatat attatagata tacatgttta tacatgtata tattatagat 600
 atatacatgt atatacatgt atatatata gatatatata tgtatatatg tatatatatt 660
 agatatatat gtatatatgt atatatatta tagatataat atatacaaga atataagaat 720
 atatatataa taatatataa tacacataat acgtatatat tatatatata tgtatatatt 780
 atagtatat atatacatgt atattatata tacatgtata ttatatatac atgcatatta 840
 tatatatatt tatatatat atccatgtat attatgtata ttgtgtata ttatatatac 900
 atgtatatata tatatacatg catattatat atattttat atataatata catatatatt 960
 atgtatatgt gtgtatatata tatatacaca tatattatat atacatggat attatatata 1020
 cacatatatt atatatatat atatatata tatacacata tattatatat acatgtatat 1080
 tatatatata cgtatatatt atatacacac gtatatata tatacacgta tattatatat 1140
 acacacgtat attatatata cacgtatat atatatatac acgtatatata tatatacacg 1200
 tatattatat atatacacgt atattatata tacacgtata ttatatatac acacgtatat 1260
 tatatatata cgtatatatt atatacacac gtatatata tatacacgta tattatatat 1320
 acacacgtat attatatata cacgtatat atatatatac acgtatatata tatatacatg 1380
 tatattatat atacatgtat attatatata cacatgtata ttatatatac atgtatatata 1440
 tatatacaca tgtatatatt atagcatgt atattatata tacacatgta tattatatat 1500
 acacatgtat attatatata catatatatt atatatatat gtatatattg tatacatata 1560
 tattatatat acatgtatat tatagataca tatatatata atatacatgt atattatgta 1620
 tacatatata ttaaataat atgtatatgt tatatacata tatattatat acatgtatat 1680
 tacatgtata cata 1694

<210> 62

<211> 587

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(587)

<223> MAR of chromosome 1 genomic contig; 3432560..3433146

<400> 62

gaattatata tatatagctg aattatatac atatataata tatacaatat atattatata 60

tttatatag atatatacaa tatatattac atattatata tacaatatat aatatataat 120

atataatatt atatattata tattgtatat aatatatatt atataacatt atataatata 180

taatattata tattatatat tgtatataat atatattata taacattata taatatatac 240

tattatatat tataatatat aatatataat aatatataat agtatatatt atatatattg 300

tatatattat atataaatat ataatatata atatatatta tataatatat attatataat 360

atatattatt atatattata ttttatata taatatatat tatatatatt atattttata 420

tataaatata taatatataa taatatataa tttaatatat ataatatata caatatataa 480

tatataatat attaatatat ataatatata caatatataa tatataatat ataatatata 540

atataaatta ttatatataa tatatattat atatagctga attatat 587

<210> 63

<211> 313

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(313)

<223> MAR of chromosome 1 genomic contig; 3805392..3805704

<400> 63

tatataatat gtatattatg taatatttta tatagcatat atgtatatta tatataatct 60

tttatatata gtatataata tgtatattat atattatata attatataat tatgtattat 120

ataaaatata ttatataata tataattata tattttttga aatatagatt atatataata 180

tatatggcag tgagctgaga tataatatat attatctata ctatataata tatattatat 240

atactctata ttatatatgt atatattata tataatatat acatatataa tgtgtatata 300

ttatatataa taa 313

<210> 64

<211> 349

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(349)

<223> MAR of chromosome 1 genomic contig; 4521378..4521726

<400> 64

ttatatacac tatataatat gtattttat atacttatat acactatata tgtattttata 60

tataattata tacactatat aatatgtatt tatatataat tatatacact atataatatg 120

tatttatata taattatata cactatataa tatgtattta tatataattg tatacactat 180

SEL PCT 012.ST25

ataatgtata ttatatata attgtataca ctatataatg tatatttatg tataattgta 240

tacactatat aatgtatatt tatgtataat tgtatacact atataatgta tatttatgta 300

taattgtata taccatataa tgtatattta tgtataattg tatatacca 349

<210> 65

<211> 500

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(500)

<223> MAR of chromosome 1 genomic config; 3240166..3240665

<400> 65

ttaatatata atatatatta tatatttata tattaatata taatatatat ttatatataa 60

tatatattat atatttatat tacatatatt tatatgttaa tatatatttt atatatttat 120

atatttata tatttatata ttatatattt atatattata ttatatatt atatatttat 180

attatatatt tatatattat atttatatat tatatattta tatttatatat ttatatattg 240

tatatttata ttatatattt atatattgta ttatatatt atatatttat atactatata 300

tatttatata tatttatatat ttatatatta tatatattta tatatattat atatttatat 360

attatatata ttatatata ttatatattt atatattata tatatttata tatattatat 420

atatttatat atatttatata ttatatata atatatatta tatattttat ctatatattt 480

atatattaat atatattata 500

<210> 66

<211> 866

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(866)

<223> MAR of chromosome 1 genomic contig; 409429..410294

<400> 66

atatatataa tatattatat atattatata ttatatatat aatacatata ttatatatat 60
 aatatataat acatatatta tatatattat atattatata taatatataa tacatatatt 120
 atatataata tataatatat aatatattat ataataaata tatataatta tataatatata 180
 tataatatat aatattatat aattatataa tatatataat tatattatat attataaata 240
 ttatataata tatatattac aaatatatat tatatatatt ataaatatta tataacatat 300
 atattatata atatatataa tatataatat atataaaaat ataatatata agatatatat 360
 aatatatgat atatatgata tataatatat gatatatatg atatatataa tatatgatat 420
 atatgatata tatgatatat ataatatatg atatatatga tatatatgat atatgatata 480
 tatgatatat gatatatatg atatatatga tatatgatat atatgatata tatgatatat 540
 gatatatatg atatatatga tatatgatat gatatatata atatatgata tgatatatat 600
 aatatatgat atatatgata tatgatatgt aatatatatg atatattata tataatatat 660
 aatatataca taatatataa tatataatat ataatatata taatatgtga tatatataat 720
 atatgatata tgatatatga tatatattat ataatatata taatatatat tatatataat 780
 atatattata taatatatat aatatatatt atatataata tataagatat aagatataat 840
 atatataata tataatatat ataata 866

<210> 67

SEL PCT 012.ST25

<211> 335

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(335)

<223> MAR of chromosome 1 genomic contig; 614754..615088

<400> 67

acccaatata tgtgtatata tgtatgtata tatacatata catacataca tatatgtaca 60

tacatatata catacataca tatatatgta catacatata tacatacata catatatata 120

tataacatat atacacacat atatacagat atacatatat acatacatat atacatataa 180

catatatata tacatatata catataaacac atacatacat acatatatac atacaacata 240

tatacataca tatatacata tgtatacata catatatgta tacatatatg tatacatata 300

tgtatacata tatgtatata tatattgtta tatat 335

<210> 68

<211> 455

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(455)

<223> MAR of chromosome 1 genomic contig; 1299520..1299974

SEL PCT 012.ST25

<400> 68

ggatatatat attattagtt gttatattat tatatattat atatattatt atatataata 60

tattatatca tatatattat tatatataat atattatatt atatattatta ttatataata 120

tattatatca tatatattat tatatataat atatattata tatattatta tatataatat 180

atattatata tattattatg tataatatat atattatata ttatttatat atatataaat 240

tatataataa tatataatta attatacata tatacatata taagtataca tataatatat 300

ttatatagta tatataaata tatatacaat atatttatat attatatatt atatataaat 360

atatacaata ttttatattc atatatttta tatatgatac atataatata tattatatat 420

gatatataat atatattcata tatgatatat aacat 455

<210> 69

<211> 404

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(404)

<223> MAR of chromosome 1 genomic contig; 1970778..1971181

<400> 69

atatataata tgtataatat ataatatata tcatatattg ttctatgtat attacatata 60

atatgcatta tatatttatat attgcatata atatgcatta tatatttatat attgcatata 120

atatgcatta tatatttatat attgcatata atatgcatta tatatttatat attgcatata 180

atatgcatta tatatttatat attgcatata atatgcatta tatatttatat attgcatata 240

SEL PCT 012.ST25

atatgcatta tatatttat aatatataca catataatat atataattta tatatattta 300

tatataattta catttattat atatttatta tatataaata tatttttata tattacttat 360

atattatata taatatatat aatatatata ttatatataa tata 404

<210> 70

<211> 605

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(605)

<223> MAR of chromosome 1 genomic contig; 3562918..3563522

<400> 70

tatatatata aaatacatat atattatata tattatatat aatacatata ttatatatta 60

tatataatac acgtatataa tatataatat ataatacata taatatatat gatataatat 120

acataataa tatatgatat ataatacata tataatatat atgatatata atacatatat 180

aatatatatt atataataa catatataat atatattata tataatacat atataatata 240

tattatatat aatacatata taatatatat tatatataat acatatataa tatatattat 300

ataatacata tataatatat attatataat acatgtatat aatatatatt atataataa 360

catatatatt atataataca tgtatataat atatattata tataatacat atatattata 420

tattatatat taatatattt atataatagt aatatataat attaatatat tatatatatt 480

aatattatat ataatacata tattatatat aatataaata tatataatac atataataa 540

cacatattat atataataca tatatttat ataatatata tattatatat aatatatatg 600

taata

605

SEL PCT 012.ST25

<210> 71

<211> 317

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(317)

<223> MAR of chromosome 1 genomic contig; 189743..190059

<400> 71

tatttttat atttatatat tatatatatt ttatatgta atatattata tataaaatta 60

tataatttta ctacatataa tatataaaat tatataattt tactacatat aatatataaa 120

attatataat ttactatat ataatatata aaattatata attttatata taatatatat 180

tataatatat atttatgca atatataatta tatattatat tataatatat tgtatatttt 240

tgtatataaa atatataata tataatatat ttatagacaa taatatataa tataatatat 300

aaaattttat atataaa 317

<210> 72

<211> 522

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(522)

SEL PCT 012.ST25

<223> MAR of chromosome 1 genomic contig; 229111..229632

<400> 72

gatatatata tttatatata taaaagatat atattattta tatataaaga tatatattta 60
 tatatatata agatatatat ttttatata tataaaagat atatatttat atatatgata 120
 tatattattt atatatataa aagatatata tttatatata tgatatatat ttttatata 180
 taaaagatat atataaaaga tatatattat ttatatatat aaaagatata tatataaaag 240
 atatatatta tttatatata taaatgatat atattattta tatataaaag atatatatta 300
 tttatatata aaagatatat attatttata tatataaaag atatacatat aaaagatata 360
 ttttatata taaaagatat atatatttat atataaaaga tacatatatt tatatatata 420
 aaagatatat atatttttat atataaaata tatattatat atataaaaga tatatatata 480
 tatatatatc ttttatatat aaaagatata tataaatata ta 522

<210> 73

<211> 1110

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1110)

<223> MAR of chromosome 1 genomic contig; 1138030..1139139

<400> 73

tatgtatgta tacataatat attatatatg tatattatgt atacataata tattatatat 60
 gtatattatg tatacataat atattatata tgtatattat gtatacataa tatattatat 120

SEL PCT 012.ST25

attatatgta tattatgtat acataatata ttatatatta tatgtatatt atgtatacat 180
 aatatattat atattatatg tatattatgt atacataata tattatatat tatatgtata 240
 ttatgtatac ataatatatt atatattata tgtatattat gtatacataa tatattatat 300
 attatatgta tattatgtat acataatata ttatatatta tatgtatatt atgtatacat 360
 aatatattat atattatatg tatattatgt atacataata tattatatat tatatgtata 420
 ttatgtatac ataatatatt atatattata tgtatattat gtatacataa tatattatat 480
 attatatgta tattatgtat acataatatt tatatattat atgtatatta tgtatacata 540
 atatattata tattatgtat atattatgta tacataatat gtacacataa tatttatata 600
 ttatatgtat attatgtata cataatatt atatattata tgtatattat gtatacataa 660
 tatttatata ttatatgtat attatgtata cataatatt atatattata tgtatattat 720
 gtatacataa tatttatata ttatatgtat attatgtata cataatatat tatatattat 780
 atgtatatta tgtatacata atatattata tattatgtat atattatgta tacataatat 840
 attatatatt atatgtatat tatgtatata taatattat atattatatg tatattatgt 900
 atacataata tattatatat tatatgtata ttatgtatac ataatatatt atatattata 960
 tgtatattat gtatacataa tatattatat attatatatg tatattatgt atacataata 1020
 tattatatat tatatatgta tattatgtat tatattatat attatgtata ttatagatta 1080
 tgtatgcata cataatatgt attgtatatt 1110

<210> 74

<211> 521

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(521)

SEL PCT 012.ST25

<223> MAR of chromosome 1 genomic contig; 2863407..2863927

<400> 74

```

aatatatata aatatataaa tatatatata tatatatata tataaatata taaatatata   60
tatgtaaata tatgtaaata tatgtaaata tatgtatatg tatatatatg taaatgtatg   120
taaatatata taaatatatg taaatatata taaatatatg taaatatata aatatatata   180
actatatata aatatatata aatatataaa tataaatata tataaatata tataaatata   240
taaataaata catataaata tataataaaa tacatatata tatatatata tatataaaaa   300
tatatatata tatatatata aatatataaa catatatata tatataaata tatataaata   360
tataaatata taaatatatg aatatatata aatatatata atatatata atatatata   420
atatagataa atatatata atatatata atatatata atatatata atatatata   480
aaatatata atataaaaa atatatata atataaaaa a                                     521

```

<210> 75

<211> 560

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(560)

<223> MAR of chromosome 1 genomic contig; 5712303..5712869

<400> 75

```

atataattat atatatatta tatattatat ataattatat attatatata atgtataatt   60
atatattata tataatatat ataatatat atattttta tataaatata ttatatattt   120

```

SEL PCT 012.ST25

atatattata tataaattta tatatatataa ttttatata ttatatatat ttatatatta 180
 tatattgtat atatttatat attacatatt gtatatattt atatattata tattatatat 240
 ttatatatta tatatttatat atttatatat tatatattat atatatttat atatttatata 300
 taaattttt atatataata tataaatata tattatataa tataaatttg tatatataat 360
 atatatttat atttatatata aaatatttat atttatatata aaatataata taaatatata 420
 catataatat atatattata tatttataat tatatattat atataatata tataatatat 480
 aatatataat acatatatat catatatgaa atatatatca tatattatac atatttatata 540
 taacatatat atttatatc 560

<210> 76

<211> 479

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(479)

<223> MAR of chromosome 1 genomic contig; 8578812..8579290

<400> 76

tatggtatac atatagtata tatggggtac atatatggta tatatatggg ttatatatat 60
 gatatatatt atatatgtat atggtatata tatggtatat atattatata tgcatatggg 120
 atgtatatgg tatatatatg atatatacat atggtgtata tatatgttat atatgatata 180
 tataagggtat atatatggta tatataagggt atatatagta tatatatggg atatataagg 240
 tatatattgt atatatatgg tatatataag gtatatatat tgtatatatg gtatatatat 300
 ggtttatata tatggtgtgt atatatgggt tttatatata cactttatat actatatatt 360

SEL PCT 012.ST25

atatacacac tatatataat atatattata tatagttaaa tatatggtat atgcaattag 420

atatatggta tatgtaatta tatatatggt atatagatgg tgtatatatg gtatatata 479

<210> 77

<211> 477

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(477)

<223> MAR of chromosome 1 genomic contig; 8579294..8579770

<400> 77

tatagtatat atacacacta taggtaatat actacatatt atatacacac tataaataaa 60

atatataata tataatatft tctatatagt atatattata tattgtatat actatatata 120

atatatacta tagacagtag atactttata tactatagac agtatatact atatactgta 180

tacactatag acagtatata ctatatactg tatacagtat atgtagtgta tatgtagtgt 240

atataatata tagtatatat tatctatact atatacagta tatatagtgt atacataata 300

tatattatat attatatata ctatatacag tatacatagt gtatatgtag tgtataatat 360

atataatgtg tatataaaat atatatacta tatataatat atattatata taatatatac 420

actatatata ctatagatac actatatatt cactatatat actatatata ctatata 477

<210> 78

<211> 331

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(331)

<223> MAR of chromosome 1 genomic contig; 8580024..8580354

<400> 78

actatatgtt atatacataa gatatagtat ataccatata ttatatacat tatatatagt 60

gtatactata tataatgtat ataatatata gtatatatac actatatata ctatgtatat 120

atacactata tatactatgt atatatacac tatatatact atgtatatat acactatata 180

tactatgtat atatacacta tatatactat gtatatatac actatatata ctatgtatat 240

atacactata tatactatgt atatatacac tatatatact atgtatatat agtgtatata 300

tactgtatat gttatagtgat atatatagta t 331

<210> 79

<211> 410

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(410)

<223> MAR of chromosome 1 genomic contig; 8580705..8581114

<400> 79

tatagtctat attatataca gtctatataa tatatagtat atactatata tacttttcct 60

SEL PCT 012.ST25

cattctgact atatactata tatatactat atatagtata ttagtgtat atatacacca 120
 tatatactat atatagtata cataccatat atagtatact atacatacca tatatagtat 180
 acataccgta tatagtatac tatacttacc atatatagta tacatactat atataatata 240
 tctggtgtat atatacacta tatatactat atatactata tatagtatat gtacactatt 300
 tatagtattt atagtatata tactgtatat atagtatgta gtatatatac tatatattat 360
 gtagactata tataatatag actatgtgta gagtatatat actatatata 410

<210> 80

<211> 433

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(433)

<223> MAR of chromosome 1 genomic contig; 12979167..12979599

<400> 80

atatataata tatatatgtc ctatatataa aatatatcat atatataaat atatatgata 60
 tattttatat attaaatata taattatata taaatatata ttatatata aatatattat 120
 ttcaatatat ataaatatat ttaaatatat ttaaatagaat tattaaatat ataaatatat 180
 aattatattt aatatataaa tatatatata atatatatt atatttaata tatataaata 240
 tatattaaat atataattat atatttatat atttattata tataaatata tatttgttct 300
 aaataaatat atattctaaa tatataatat ttatatatt ataatatata atataaaata 360
 tataataaat atataatata taaataaata aatatttatt ataaaataca tataaatatt 420
 aaatatatat taa 433

SEL PCT 012.ST25

<210> 81

<211> 385

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(385)

<223> MAR of chromosome 1 genomic contig; 16336644..16337028

<400> 81

tttatataaa tatctatata aataaatata taaatatata aatataaata tatataaata 60

tataaataaa tatataaata tatataaata taaatatata tataactatg aatttatatt 120

tatataaata tatatctata tgaatataaa tatatattta tataaatata aatatatata 180

taaatatata tatftatata gatataaata tatatataaa tatatatatt tatatagata 240

taaatatata tctatatatg aatatatatc tataggaata taaatatata tctatataaa 300

tataaatata tataagtata aatatatata aatatatatc tatataaata taaatatata 360

tataaatata aatatatata taaat 385

<210> 82

<211> 363

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(363)

<223> MAR of chromosome 1 genomic contig; 20624448..20624810

<400> 82

tatatatata gttatatata tatttatata tatagttata tatatatattt tatatagtta 60

tatatatagt tatatatata gttatatata tatagttata tatatagtta tatatatagt 120

tatatatata tagttatata tatagttata tatatagtta tatatatagt tatatatata 180

tagttatata tatagttata tatatatagt tatatatata gttatatata tatagttata 240

tatatagtta tatatatagt tatatatata gttatatata tagttatata tatatagtta 300

tatatatata gttatatata tatagttata tatatagtta tatatatata gttatatata 360

tag 363

<210> 83

<211> 310

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(310)

<223> MAR of chromosome 1 genomic contig; 566025..566334

<400> 83

tatatataat atatattgta tatattatat attgtatata taatatatat tgtatatatt 60

atatattgta tatataatat atattgtata tattatatat tgtatatata atatatatat 120

tgtatatatt atatattgta tatataatat atatattgta tatattatat attgtatata 180

SEL PCT 012.ST25

taatatatat attgtatata ttatatattg tatatataat atatatttg tatatattat 240

atattgtata tataatatat atattgtata tattatatat agtatatatt atatatagta 300

tatataatat 310

<210> 84

<211> 1236

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1236)

<223> MAR of chromosome 1 genomic contig; 1171429..1172664

<400> 84

aaagtattat atgtattata tgtatatgta ttatatatta catatgtatt atatataata 60

tatattatat attattatat attatatatt atatattatt atttatataa tgtattatat 120

attatatagt atatatagta tatataatgt attatatatt atatagtata tatagtatat 180

ataatgtatt atatatagta tatataatgt attatatagt atatatacta tataatgtat 240

tacatattat gtagatgata tgtaatgtat tatatattat atagtatatg taatgtatta 300

tatgtattat atagtatata ttatatatga tgtattattt agtatatata atatatatga 360

tgtattatat aacatatata atatatatga tgtattatat agcatgtata gtatatatga 420

tgtattatat agcatgtata gtatatatga tgtattatat atagcatgta tagtatatat 480

gatgtattat atatagcatg tatagtatat atgatgtatt atatatagca tgtatagtat 540

atatgatgta ttatatatag catgtatagt atatatgatg tattatatat agcatgtata 600

gtatatatga tgtattatat atagcatgta tagtatatat gatgtattat atatagcatg 660

SEL PCT 012.ST25

tatagtatat atgatgtatt atatatagca tgtatagtat atatgatga ttatatatag 720
 catgtatagt atatatgatg tattatatat agcatgtata gtatatatga tgtattatat 780
 attatatatg gtatatatga tgtattatat attatatatg gtatatatga tgtattatat 840
 attatatatg gtatatatga tgtattatat attatatata atatatatga tgtattatat 900
 attatatata atatatatga tgtattatat atgatgtatt atataaata tatatgatgt 960
 attatatata ttattatcta ttatatacga tgtattatat gcaagttatt atgtataata 1020
 tataatgtat tatatattat ataatgtata atataaaat atataaatat ataattatgt 1080
 ataaatatag aaatatatac attatacatt atatacatta taatgtataa tatataaata 1140
 tattatatat aaatgtatac attatatata aatatattat atacattata tataaaatat 1200
 gtatatagtt attatacctt atatatacta aacagt 1236

<210> 85

<211> 309

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(309)

<223> MAR of chromosome 1 genomic contig; 1925173..1925481

<400> 85

atatatttat ataaatatat ttatatataa tatatattat ataattattat aatatatgtt 60
 atattatata tattttatatac aatatataat atattattata tatattttat acaatatata 120
 atatatatta tatatatttt atataatata taatatatat tatatatatt ttatacaata 180
 tataatatat attatatatt atataatata tattatatat attttatata atatatata 240

SEL PCT 012.ST25

tatatatttat acaatatata atgtatatca ttatattata taatgtatat catattatat 300

aatgtatat 309

<210> 86

<211> 312

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(312)

<223> MAR of chromosome 1 genomic contig; 4396756..4397067

<400> 86

cacagtgtat atatagtata tatactgtat atatactgtg tatatacact gtatatacac 60

agtgatatata cagtatatat actatatata cactgtgtat atatagtata tataaattct 120

aggaatatat atactatata tatactatat atataaattc taggaatata tacacactat 180

atatacacta tatatacaca tatatacact atatatatta tacacatata ttatatatat 240

acactatata tacacgagat atataacata tacactatat actatacata acatatatac 300

tatatatact at 312

<210> 87

<211> 398

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(398)

<223> MAR of chromosome 1 genomic contig; 56057..56454

<400> 87

atatatatta catatttat atataatata tattatataa tatatattat attatataat 60

atataatata aatataatat aaatttatatt atataatata taatataaat ataataataa 120

ttatataaat ataatatata ttttattata taatataata tatatttat aaatataata 180

tataaattat ataataatat atatattata taatataata tattttatta tataaatata 240

tattatatta tataatatat attttattat ataatatata ttatatatt atagaatata 300

atatatattt tattatataa tatatattat ataatatata ttatatttat atataacata 360

tattattata taaaatatgt ataatatata ttatataa 398

<210> 88

<211> 391

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(391)

<223> MAR of chromosome 1 genomic contig; 56984..57374

<400> 88

tactataata catatttat ataattattat atactatata ttactatatt attatattat 60

atataattaa actatattat agtatataat atataatata tactatatgt aatattacta 120

SEL PCT 012.ST25

tgatactgat attatattat atataattaa attatattat attaatatat aaattatata 180
 taatacataa tatataaatt atattatatt atttatatat aatgtatgcc atataattta 240
 tatataatgc atttatata atttatatat aatgcattaa atataaatta tatataatgc 300
 atttatata atttatata atgcattata tataatttat atttaatat taaatttata 360
 ttaatatat ttatatatta tatataataa a 391

<210> 89

<211> 309

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(309)

<223> MAR of chromosome 1 genomic contig; 469547..469855

<400> 89

atatatatgt aatatatatg ttatatatgt aatatatatg ttatgttata tatgttatat 60
 atatgttata tataatatat atgttatata tacgttatat gttatatata tgttatatat 120
 aatatatgtt atatatatgt tatatgttat atatgttata tataatatat gttatatata 180
 atatatgtta tatatgttat atataatata tgttatatat atttatata atatatgtta 240
 tatatattat atataatata taatatatgt gatataatat ataaaatata tgtgatatat 300
 attatatat 309

<210> 90

<211> 441

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(441)

<223> MAR of chromosome 1 genomic contig; 546190..546630

<400> 90

```

atacacaaca tatgtgtata tatatagtat atatacaca catatgtgta tatatatagt   60
atatatacac aatatatgtg tatatatata gtatatatac acaatatatg tgtatatata   120
gtataaatat atactatata tagtatatat agtataaata tatactatat atagtatata   180
catagtataa atatatacta tatatagtat atacatagta taaatatata ctatatatag   240
tatatacata gtataaatat atactatata tagtatatac atagtataaa tatatactat   300
atatagtata tacatagtat aaatatatac tatatatagt atatacatag tataaatata   360
tactatatat agtatataca tagtataaat atatactata tatagtatat acatagtata   420
aatatatact atatatagta t                                     441

```

<210> 91

<211> 1367

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1367)

<223> MAR of chromosome 1 genomic contig; 124643..126009

SEL PCT 012.ST25

<400> 91

atatttat gatataat atatataa ttatatataa tattatat gatataac 60
 attatata attatatg atatatta tatattat atagatata taatatat 120
 aatattat atgattat atacatata taatatata aatattat atgatatata 180
 atatataa tattatat attatatata ttatatata tatataat tctaatata 240
 taattata tgatatata gattatatac attatatata atatata ttatatga 300
 tatataat tatatacat atatataa tataatgat ataattat atattatata 360
 ttattat atacaatga tataatata tatatcat atatttat tatatacat 420
 gtataata ttatatata tatataat tatatacat gtataata tatattat 480
 atatttat tatatacat gtataata tatattat atatttat tatatacat 540
 gtataata tatattat atatttat tatatacat gtataata ttatatata 600
 tatattat attatata tatacaatg atatattata tattatat ttattata 660
 tacaatgat atattatata ttatatatt atattatata caatgat attatatatt 720
 atatttat attatata atgatatat tatattat atatttat tatatacat 780
 gtatatata tatattat attatata tatataatg atgaatatt atatattata 840
 tattatatt atatataatg tatgaatata tatattat atatttat tatataat 900
 gtatgaata ttatatata tatattata ttatatata tgatgaat attatatatt 960
 atatattat attatatata atgattga tattatat tatattata tattatat 1020
 aatgatga atattatata ttatatatt atattatata taatgat aatattat 1080
 attatatatt tatattat ataagtat taattata tattatat ttattata 1140
 tataatgat gaattat atattatata ttattat atataatga tgattata 1200
 tatattat attatata tatataatg atataatatt atatattata tattatatt 1260
 gtataata ttatatata tatattata ttatatata tatattat atatttat 1320
 tgatatata attatatatt atatttat attatatata atgata 1367

SEL PCT 012.ST25

<210> 92

<211> 458

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(458)

<223> MAR of chromosome 1 genomic contig; 58908..59365

<400> 92

tatatgatata atagatata tatgggatata atagatata tatgatata atggatata	60
tatgatata agtatatag atatatatgg tatatatatg atatatagta tatatgatata	120
ataggtata tatgatata agtatatag atatatatgg tatatatggg atatatatga	180
tatatgatata atagatata tatgatata gatatatatg atatatatga tatatatggg	240
atatatgata tatatgggata ataggtata tatatgatata atagatata tatggatata	300
atagatata tatgatata atggatata tatgatata atagatata tggatata	360
atagatata tgatatata ggtatatata tgatatata gatatatatc atatatatgg	420
tatatatatg atatatatga tatatatcat atatatgg	458

<210> 93

<211> 330

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(330)

<223> MAR of chromosome 1 genomic contig; 306867..307196

<400> 93

ataatatata aatatatatg atatatatct atatatatca tatataaata tatatgatat 60

atatctatat atatcatata taaatatata tgatatataa atatatatga tatatatcta 120

tatatatcat atataaatat atagtatata taaatatata tgatatatat ctatatatat 180

catatataaa tatatatgat atatatctat atatcatata taaatatata tgatatatat 240

ctatatatat catatataaa tatatatgat atctatctat atatatcata tataaatata 300

tatgatatct atctatatat atcatatata 330

<210> 94

<211> 353

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(353)

<223> MAR of chromosome 1 genomic contig; 636899..637251

<400> 94

tatgtataca tatacacata tacgtatata tatacatata tacacatata cgtatatata 60

tacgtataca tacatatgta tatgtatacg tatacacaca tatgtatatg tatacgtata 120

cacacatata cgtatatatg tatacgtata cacacatata cgtatatgta tacatatata 180

SEL PCT 012.ST25

tgtgtacata tacgtatata cgtatatgta tacatatata cgtttatgta tatatacgta 240

tatacgata tatgtatatg tatacatata tacatatatg tgtatatatcg tatatacgta 300

tatgtgtata tatacaatat acatacatgc acatatatgt gtatatgcac ata 353

<210> 95

<211> 345

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(345)

<223> MAR of chromosome 1 genomic contig; 1435510..1435854

<400> 95

atcatatata ttatatatca tatatatgat atataaaaat tatatatcat atatatgata 60

tataaaaatt atatatatca tatataatat atataatata ttatatatat aaattatata 120

taatatatat aaattatata tatcatatat atgatatata atttatatat catatatatg 180

atatatatata tatattatatt atataataa tatttatatat tatataatat gtaatatata 240

ttatatatta catattatat tatttataaa taatatttta taatatatat aatattatat 300

aatatagaat atttatatatt atatattaca tattatataa tatat 345

<210> 96

<211> 521

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(521)

<223> MAR of chromosome 1 genomic contig; 39695..40215

<400> 96

tatatatata atagatatta tatatctatt atatatctat tatatatata atagatatta 60

tatatctatt atatatataa tagatattat atatctatta tatatataat agatattata 120

tatctattat atataatata tatctattat atattatata tctattatat ataatatata 180

tctattatat atattatata tctattatat atataataga tattatatat ctattatata 240

taatatatat ctattatata ttatatatct attatatata tgtatctatt atatatatta 300

tgtatctatt atatataata tatatctatt atatatatat tatatataat atatattata 360

tatattatat atctattata tataatatat atctattata tatattatat atctattata 420

tatattatat atctattata tataatatat atctattata tatattatat atctattata 480

tataatatat attatatata tattatatat tgtatatcta t 521

<210> 97

<211> 484

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(484)

<223> MAR of chromosome 1 genomic contig; 1286007..1286490

SEL PCT 012.ST25

<400> 97

atatcatata tattatatat catatatatg atatataaaa attatatatc atatatatga 60

tatatatata ttatatatat catatatat atatataata tattatatat ataaattata 120

tataatatta tatataaatt atatatcaca tatatgacat ataaattata tatcacatat 180

atgatatata atttatatat cacatatatg atatataatt tatatatcat atatatgata 240

tataatttat atatcatata tatgatatat aatttatata tcatatatat gatatatata 300

atatattatt tatatatatt atattatata ttatatata tgtaatatat attatatatt 360

atataatag tatatatat tatattattac atattatatt atttataaat aatattttat 420

aatatatata atattatata atatagaata ttatatatta tatattacat attatataat 480

atat 484

<210> 98

<211> 244

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(244)

<223> MAR of chromosome 1 genomic contig; 73556..73879

<400> 98

attatatatt atattatata atatataata atattatata attatatatt acattatata 60

atatataata atattatata ataafatata attatataat atataataat attatataat 120

attatataat attatataat atataaatat ataataatat atattatatt atataatagt 180

atatattata ttatatata tatgttatta tattatataa tataaactat tatataatat 240

SEL PCT 012.ST25
244

aata

<210> 99

<211> 463

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(463)

<223> MAR of chromosome 1 genomic config; 179038..179500

<400> 99

tacaatatat ttctattat atatatttg tattatatat aatatacaat atattttcta 60

ttatatataa tatattttgt attatatata ttacaatata tttgtatta tataatatat 120

aatacaatat aatatattgt attatataat atataatact atataatata ttgtattata 180

tattatatat aatactatat aatatatttt attatatatt atatataata ctatataata 240

tattttatta tatattatat ataatacaat atataatata ttgtattata atacaatgta 300

ttataatgta ttatatataa tatataatac aatatataat attatatata ttatatata 360

tatatatatt gtattatata tttgtatta tatatatatt gtattatata ttatatatt 420

atattataat tatgttttgc attatatatt tcatattata tat 463

<210> 100

<211> 390

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(390)

<223> MAR of chromosome 1 genomic contig; 55617..56006

<400> 100

tgtataatat atatacttta tatataatat atatacttta tatatatact atatactaata 60

atatataata tatactatat ataatatata ctaatatata taatatatac actatatata 120

atatatacta atatatatta tatatacttt atataatatata tactaatata tataatatat 180

atactttata tataatatat actaatatat ataattgata tactttatat ataatatata 240

ctaatatata atatatatac ttatatata atatatacta atatatatta tatatacttt 300

atatatataa tatatacttta tatattatat atgcttatat ataatatata cactaatata 360

taatatatat actttatata ttatatttta 390

<210> 101

<211> 582

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(582)

<223> MAR of chromosome 2 genomic contig; 1157405..1157986

<400> 101

tgtatatgta tatatacaca tacgcacata tatgtatatg tatatatata catacgcaca 60

SEL PCT 012.ST25

tatatgtata tgtatatata cacatacgca catatatgta tatgtatatg tatatgtata 120
 tatacacata tacacatata tgtatatgta tatatacaca tatacacata tatgtatatg 180
 tatatatata catatacaca tatatgtata tgtatatata cacatacaca tatatgtata 240
 tgtatatgta tatatacaca tacacatata tgtatatgta tatgtatatata tacacatata 300
 cacatatata catatatgta tacatatatg tgtatatata tacacatata tatacatata 360
 tgtatacata tatgtgtata tatacacata tatatacata tatacatata catatatatg 420
 tgtatgtata tatacacata tacatatata tgtatatgtg tatatatatt agacagatat 480
 atatgtacat atacatatat atgtatatgt atatgtatat gtatatgtat atgtatatgt 540
 atatgcatat ataatatata tatacatata tgtatatgta ta 582

<210> 102

<211> 322

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(322)

<223> MAR of chromosome 2 genomic contig; 1858638..1858959

<400> 102

acacatatata tacacatatat atatacatat catatatata ccatatatat acataccata 60
 tatataccat atatatacat accatatata cccatatat atacatacca tatatatata 120
 ccatatatat acataccata tatataccat atatatacat accatatata taccatatat 180
 atacatacca tatatatata ccatatatat acataccata tatatacacc atatatatat 240
 ataccatatata tataccatat atataccata tatatacacc atatatacac accatatata 300

SEL PCT 012.ST25

ccatatatat acaccatata ta

322

<210> 103

<211> 914

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(914)

<223> MAR of chromosome 2 genomic contig; 5712196..5713109

<400> 103

aaaatatat tctatatata gaaaatatat attctatatata tatagaatat atatagaata 60
 tatattctat atatattcta tatatataga atatatatat aaaacatata ttctatatat 120
 aaaatatata ttctatatat ataaaatatata tattctatat atatagaatg tatataaaat 180
 atatattcta tatatataga atgtatatata aatatatatatt ctatatatat agaattgtata 240
 taaaatatat attctatatata tatagaatgt atataaaata tatattctat atatatagaa 300
 tatatataac atatatatga aatatatatata aaatatatat aaatacatat ttctatatat 360
 aaatatatat aaatacatat ttctatatat aaatatatat caatacatat ttctatatat 420
 aaatatatat aaatatatat tcatatatat aaaaatatat aaatatatat tcatatatat 480
 aaaaatatata tgaatatata ttctctatat ataaaatatata tataatatat attatatata 540
 taaaatatat ataatatata ttatatatat aaaaatatata taatatatat tcatatatat 600
 aaattatata taaatatata ttcatatatata taatatatat aaatatttat ttcatatatata 660
 aaatatattt aaatatatat ttctatatag aatatatatatt ctatatataa aatatatatata 720
 taaatatatt ttctatatag aaatatatat gaaatatata gaatatatat aaatatatat 780

SEL PCT 012.ST25

tatatatact atatatacaa tatataattat atataaaata tatatacaat atatattcta 840
 tatattaata tatagaatat atattaacat atatttcaat atattaatat atgaaatata 900
 tataaatatt tcat 914

<210> 104

<211> 370

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(370)

<223> MAR of chromosome 2 genomic contig; 5713613..5713982

<400> 104

tatttcatat ataatatata tataaaatat atatttcata tacataatat atataatata 60
 aataaaatat atatttcata tatataatat atataatata tataaaacat atatttcata 120
 tataatatat ataaactata tatttcatat ataatatata taaactatat atttcatata 180
 cataatatat ataatatata ttccatttat attatatata taatatatat ttcatatata 240
 taatatataa aatagatata aatatatata aatatatatt tcatatataa tatatatataa 300
 atatatatata atatatatat tatatatata atatattt catatatataa tataaaaaaa 360
 tatatatattc 370

<210> 105

<211> 442

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(442)

<223> MAR of chromosome 2 genomic contig; 7481647..7482088

<400> 105

atataaatta tataatatgt tatataatat ataaatatat tatataacat gttatataat 60

atataacatg ttatataata tataacatgt tatataatat ataacatgtt atataatata 120

taacatgtta tataatatat tatgtaatat gttatataat atataatata ttatataaca 180

tgttatataa tatataacat gttatataat atgttatata atatataaat atattatatt 240

atatgttata taatatataa atatattata ttatatgtta tataatatat aaatatatta 300

tattatatgt tatataatat ataaatatat tatattgtat gttatataat atataaatat 360

attatatgt atgttatata atatataaat atattatatt gtatgttata taatatataa 420

atatattata ttatatatgt ta 442

<210> 106

<211> 338

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(338)

<223> MAR of chromosome 2 genomic contig; 9594557..9594894

SEL PCT 012.ST25

<400> 106

tatataaata tataccatat atataaatat atatattcca tatataaata tatatattcc 60

atatatataa atatatatat tccatatata aatatatata ttccatatat ataaatatat 120

atataaatat atatattcca tatatatataa tatatatata aatatatata ttccatatata 180

aatatatata tattccatat ataaaaatat atatatttc catatatataa aatatatata 240

tattccatat atataaatat atatatttc catatatata aatatatata tattccatat 300

atataaatat atatatttc catatatata aatatata 338

<210> 107

<211> 364

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(364)

<223> MAR of chromosome 2 genomic contig; 10519720..10520083

<400> 107

ttatatatat ttataaat atataaagc tatatatatt tatatataat atattatata 60

tattagctat atatatttat ataataatat attatatatt agctatatat atttatatat 120

aataatatat ataagctata tatttatata tattatatat tagctatata tatttatata 180

taatatatta tatattagct atatatttat atataataaa taatatatat attagctata 240

tatatttata tataataata tatataagct atatatttat atataatata ttatatatta 300

gctatatata ttatatata ataatatatt atatattagc tatatatatt tatatataat 360

atat 364

SEL PCT 012.ST25

<210> 108

<211> 342

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(342)

<223> MAR of chromosome 2 genomic contig; 11481943..11482284

<400> 108

tacatataat atataattat atataatata tattatatat tacatatata atatatatat 60

tacatatgta atatatatat tatatatgta atatatatta tataatgtaat atatatatta 120

tatatgtaat atatattata tatatgtaat atatatatta tatatgtaat atatatatta 180

tatgtaatat atatatgtaa tatatatata atatatatgt aatatatata taatatatat 240

gtaatatata tataatatat atgtaatata tatattatat atatgtaata tatatcatat 300

atatgtaata tatatcatat atatgtaata tatatcatat at 342

<210> 109

<211> 415

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(415)

SEL PCT 012.ST25

<223> MAR of chromosome 2 genomic contig; 13499598..13500012

<400> 109

tatatatata tatatatata atataatata atatatatat aaatatatat aatataaatt 60
 tatatatata tatttatata tacatatata aatatatatt tatatttata tataaatata 120
 tataaatata tataaatata tatttatata tacatatata aatatatatg ttcatataaa 180
 tatatatgta tatatacata tataaatata tatttatatat gtatatatat aatataatat 240
 ataataataa tataatatat attatataaa tataatatat tatatataat atatataata 300
 tataatatat aatatataat atataatata tatttatatat tatataatat ataaaatata 360
 tattatataa tatatataca taatatatat aaataaatat atataaagat ataaa 415

<210> 110

<211> 330

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(330)

<223> MAR of chromosome 2 genomic contig; 16370976..16371305

<400> 110

cattacata tgtatgtata agtatgtata ttacatactt atacatacat acttataaat 60
 atataagtat aatacatata tacttataaa tatataagta taatacatat atacttatac 120
 atatataagt ataatacata catacttata catatataag tataatacat acatacttat 180
 acatatataa gtataatata tacatactta tacatataag tataatacat acatacttat 240

SEL PCT 012.ST25

acatatataa gtataataca tacataactta tacatatata agtataatac ata cttatta 300

catatgtata taagtatat acataacttat 330

<210> 111

<211> 702

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(702)

<223> MAR of chromosome 2 genomic contig; 626641..627342

<400> 111

tatatataca catatacata tataatatat atacatatac atatataatta tatatacata 60

tatattacat atatacata tacatatata ttatatatac atatataatta tatatatcat 120

atatacatat atatattata tattatatat atcatatata catatatatt atatataatta 180

tatatatcat atatacatat atattatata tattatatat acatatatat tatatatatc 240

atataaacat atatattata tatatcatat atacatatat attatatata ttatatatat 300

catatatata tatatattat atatacata tataatatat attatatata ttatatataa 360

tatatattat atatacatat atattatata tacatatata ttatatatac atatataatta 420

tatatacata tatattatat atacatatat attatatata tacatatata ttatatatac 480

atatataatta tatatacata tattatatat acatatatat tatatatata tatattatat 540

atatacatat atattatata tacatatatt atatatatac atatataatta tatatacata 600

tattatatat atacatatat attatatata catatattat atatacatat atattatata 660

tacatatata tttatatat atataatata tttttatat at 702

SEL PCT 012.ST25

<210> 112

<211> 679

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(679)

<223> MAR of chromosome 2 genomic contig; 3196047..3196725

<400> 112

atatatata tattcatata tcataaatat atatattata tattcatata ttatatatct 60
 atatatttat atattcatat attatatatc tatatatffa tatattcata tatttatatat 120
 ctatttatat attcatatat tatatatcta tatattttat atattcgtat attatatatc 180
 tatatattat atattcgtat attatatatc tatatattat gtattcatat atatctatat 240
 attatatata ttcatatata ttataaatta tattcatata gtatatatct attataaatg 300
 tatattcata tagtatatat ctatatatta taaatataca tatattatat atttatatat 360
 tatatattca tatagatcta tatattatat atattcatat atgaatatat atatttatatg 420
 tatatatatt ataaatatat ttatatagta tagatattat atagtataig catatttata 480
 ttataaataa ttacatagt atagtatat ttataaatta tatatatffa catattacat 540
 gtatatttat atattataaa tacatattta catattataa atatatttat atattatgaa 600
 tataatttat atattattac atatttcat atatgcatag ttatatatta taaatatgca 660
 ttatgtaaa tatatattt 679

<210> 113

<211> 728

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(728)

<223> MAR of chromosome 2 genomic contig; 3196778..3197505

<400> 113

```

tacataaata tatatttaca atatgtaaat atctgatatg taaatatgta ttataaatat   60
ataaatatac atataaatg taaatatata aatatacata tactatgtaa atatatgta   120
tatatacata tactatataa atatagaata tataaatata catatactat ataaatatgt   180
aatatataaa tatatactat ataaatatac atatactata taaatgtatt tataaatatat   240
aaatatacat atactatata aattcatata tgaatatata atatataaat atatataata   300
tatgaatata tactcatata taaatatata tgaatatata ttataaatat atagatataa   360
tatgaatata tatttataat atatagatat atattatatg aatatatatt tataaatatat   420
agatatatac catatgaata tatatttatac actatatgaa tatatattta taatatataa   480
atagatatat actatatgaa tatataatat atatactcta tgaatatata atatatatac   540
tatatgaata tattatatac tgtatgaata tataaatatat agatgtatac tatatgaata   600
tataaatatat agatatatat actatatgaa tatatataat atatagatat atactatatg   660
aatatatatg atatatagat atatactata tgaatatata atatatagat atatatttat   720
gatatatg

```

728

<210> 114

<211> 413

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(413)

<223> MAR of chromosome 2 genomic contig; 2560638..2561050

<400> 114

atataaatat atatttatat attttatata aatatatata tttatatatt tttatataaa 60

tatatatatt tatatatatt tatataaata tatatattta tatatattta tataaatata 120

taaatatata tatttatata aatatataaa atatataaat atatttatat aaatatataa 180

aaatatataa tatatttata taaatatata aaatatataa atatatttat atataaatat 240

ataaaatata taaatatctt tatatataaa tatataaaat atataaatat ctttatatat 300

aaatatataa aatatataaa tatatttata tataaatata taaaatatat aaatatattt 360

atatacaaat atataaaata tataaatata tttatatata aatatataaa ata 413

<210> 115

<211> 361

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(361)

<223> MAR of chromosome 2 genomic contig; 4965309..4965669

SEL PCT 012.ST25

<400> 115

tatacgtata tatacatata tatacgtata tatatacata tatatacgta tatatacata 60

tgtatatatg tgtgtacatg tatatatata catatgtaca tatatatgta cacatatata 120

tatacatata tatgtacaca tatacatata tatgtacaca tatacatata catatatatg 180

tacacatata tatacatata tatgtacaca catatatata catatatatg tacacacata 240

tatacgtata tatgtacaca catatatatg tatatatatg tacacacata tatacgtata 300

tatatgtaca cacatatata tacgtatata tatgtacaca tatatatata cgtatatata 360

t 361

<210> 116

<211> 325

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(325)

<223> MAR of chromosome 2 genomic contig; 5258150..5258474

<400> 116

tacacacaca tatacatata tacatatata cgtgtatacg tatacgtata tacgtatata 60

tacatatatg tatacgtata cgtatatatg tatatatata tatatgtata cgtatacgta 120

tatacgtata tatacatata tgtatacgta tacgtatata cgtatatata catatatgta 180

tacgtatacg tatatacgta tatatacata catatgtata cgtatacgta tatatgtata 240

tatacgtata tgtatacgta tacatatata cgtatatata cgtatatgta tatgtatata 300

cgtatatgta tatatgtaca tatac 325

SEL PCT 012.ST25

<210> 117

<211> 1508

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1508)

<223> MAR of chromosome 2 genomic contig; 6057499..6059006

<400> 117

atataatata tataaattat ataatatata aaaattaata tataatatat ataaattata 60
 taatatataa attaattata taatatatat aaattatata atatatataat taattatata 120
 atatatataa attatataat acatatataat taattatata atatatataat tatataatat 180
 atacaaatta tatactatat taattatata ttatataatt aattatataa tatatataaa 240
 ttatatatta tttaaattaat tatataatat ataaattata taatatataa attaattata 300
 taatatataa attatataat atataaatta attatataat atataaatta tataatatat 360
 aaattaattg tataatatat aaattaatta tataatatat aatatataat taataaataa 420
 ttatatatta attatataat taataaataa ataataaata tatataatta atatatataa 480
 tacatcatat atatcacata tagattatat aatagttata tattatataa taaattatat 540
 ataatatata ataaacatat ataacatatg ttatatatta cataatatag tataatatat 600
 aacatatgtt atatattaca taatatagta taatatataa catgttatat attacataat 660
 atagtataat atataacata tgttatatat tacataatat agtataatat ataacatatg 720
 ttatatatta cataatatag tataatatat aacatatgtt atatattaca taatatagta 780
 taatatataa catatgttat atattacata atatagtata atatatataa tatgttatat 840

SEL PCT 012.ST25

attacataat atagtataat atataacata tgttatatat tacataatat agtataatat 900
 ataacatatg ttatatatta cataatatag tataatatat aacatatgtt atatattaca 960
 taatatagta taatatataa catatgttat atattacata atatagtata atatataaca 1020
 tatgttatat attacataat atagtataat atataacata tgttatatat tacataatat 1080
 agtataatat ataacatatg ttatatatta cataatatag tataatatat aacatatgtt 1140
 atatattaca taatatagta taatatataa catgttatat attacataat atagtataat 1200
 atataacata tgctatatat tacataatat agtataatat atatgttata tattacataa 1260
 tatagtataa tatataacat atgttatata ttacatatta tagtataata tatatgttat 1320
 atattatata atatagtata atatataatg tatgttatat attatataat atagtataat 1380
 atataacatg ttatatatta tataatatag tataatatat atgttatata ttatataata 1440
 tagtataata tataatatat gttatatatt atataatata gtataatata tatgttatat 1500
 attatata 1508

<210> 118

<211> 415

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(415)

<223> MAR of chromosome 2 genomic contig; 7996866..7997280

<400> 118

caattatata atatacatat tatataattg tataaattat acaatcatat aattatatta 60

tatataatat acatataata taattatata taattatata attttataat ataattatat 120

SEL PCT 012.ST25

ataattatat aattatatat aatatatatt ataattatat atataatata tatattatat 180
 atattatata taatatataa ataatatata taatatatat ataattatat ataataatat 240
 atgtaatata tataatatat atataatata ttatttataa ttatatatta tatatatatt 300
 ataatatata taattataaa taatatatat tataatatat ataataatat atatataatt 360
 atatataata atatatatta taattatata taataatata tataatttat ataatt 415

<210> 119

<211> 526

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(526)

<223> MAR of chromosome 2 genomic contig; 8300930..8301455

<400> 119

tatatcatat gatataattat acaatatatc atataatatg atatattata tgatatattg 60
 tacaatatat catatgatat atgatataatt atacaatata tcatataagg tatataattat 120
 atcatatata atatataata taatatatga tataatatat gatatatgat atataatata 180
 tgatatatga tatatgatat ataatatatg atatatgata tatgatatat aatatatgat 240
 atatgatata tgatatataa tatatgatat atgatatatg atatgatata tgatatatga 300
 tataatatat gatataatat atgatatata ttatatgata tataatatat gatataattt 360
 atatgatata taatatatga tatataatat ataatatatg atatgatata tattatatca 420
 tatataatat ataatatatat atatgatata tattatatat tttatatacat tatatatata 480
 aactatataa caatatataca tattatgtgt ataatatata ttacatt 526

SEL PCT 012.ST25

<210> 120

<211> 402

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(402)

<223> MAR of chromosome 2 genomic contig; 8576553..8576954

<400> 120

```

atgtatatta tatacaatat agtatatcat atatagtata tattatatag taatgtatta    60
tatataatgt ataatgtata aatatataat atatactaca tactatacta ttatatatac    120
tatatattat atagtataca tatactatat aatatgctat atattatact atataaatg    180
ctatatatta tactatataa tatgctatat attatactat ataatatgct atatattata    240
ctatataata tgctatatat tatactatat aatatactat ataatatgct atatattata    300
ctatataata tactatatat tatactatat aatatactat ataacatact atatattata    360
tatgatacat atactatatt acatatataa tatatatata ta                        402

```

<210> 121

<211> 477

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(477)

<223> MAR of chromosome 2 genomic contig; 8785649..8786125

<400> 121

tatttatata tatatttata tatatatatta tatatatatta tatatatatt tatatatata 60

tttatatata tatttatata ttatatata tatattttta tatatttata tatatatatta 120

tatatttata tatatttata ttatatata tatttatata tatttatata tatttatata 180

tatatatatta tatatatatta tatatatata ttatatata ttatatata ttatatata 240

tatttatata tatatttata tatatatatta tatatatatta tatatatatt catatatatt 300

tatatatata ttcatatata ttatatata tatttatata tatatttata tatatttata 360

tatatttata tatatatatta tatatatatt tatatatata tatttatata tatatttata 420

tatatatatt tatatatata ttatatata tatatttata tatatatatta tatatat 477

<210> 122

<211> 773

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(773)

<223> MAR of chromosome 2 genomic contig; 10064737..10065509

<400> 122

atattatata tattacatat atattatatt gtatataata tatatatatt attgtatata 60

atatatatat tatattgtat ataatatata tatttatattg tatataatat atatattata 120

SEL PCT 012.ST25

ttgtatataa tatattatat tgtatatatt atattgtata tattatattg tataacaatat 180
 atatttatatt gtatacaata tatattatat tgtatataat atatttatatt gtatataata 240
 tatttatattg tatatattat attgtatata atatattata ttgtatataa tatattatat 300
 tgtatatatt atattgtata taatatatta tatgtatata atatagtgtata tactatatta 360
 tataatatat attatatata atatataata tattgtatat catatatgat atattgtata 420
 taatatataa tatatgatat attgtatata atatattata tatgatatat tgtatattat 480
 atatttatata tgatatattg tatattatat attatatatt gtatattgta tattatatat 540
 tatatattgt atataaatg ttatatattg tatataatat gttatatatt atatattgta 600
 tatatggtat atattatgta ttgtatataa tatggtatat attatatatt gtatataatg 660
 tatttatatat tatatatatt atatattgta tataatgtat tatatattgt atattatata 720
 ttatatattg tatataatat attatatata ttatatata tattatatat tgt 773

<210> 123

<211> 1554

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1554)

<223> MAR of chromosome 2 genomic contig; 1039775..1041328

<400> 123

ataatatatt aaatgtatat ataatatatt aaatataaat atatttataa tatataaata 60
 ttatatataa tataaaatat atattaaata taaatatata taaaatatat attaaatata 120
 taaaatataa atatatatta aatatatatt aaatatataa aatataaata tatattaaat 180

SEL PCT 012.ST25

atattttaa tatataaat ataaatatat attaaatata ttttaatat attaaatata	240
aatatatatt aatatatatt taaatatatt aatatataat acatatatta aatatatatt	300
atatatataa aatatataaa atataaatat atattaaata tatataaaat atatatgtta	360
aatatataaa agatatataa aatataaata tatattaaat atatatataaa tatatatata	420
ttaaatatat atattaaata taaatatata taaaatataa atatatgtat taaatatata	480
tattaaatat aatatatgt attaaatata tattaaatat gaatatatgt attaaatata	540
tattaaatat aatatatgt attatatata tagaatataa atatatgtat taaatatagt	600
atattaaata taaatatata taaaatatat attaaatatg aatatatata aatatatat	660
attaaaaata tatataatat aatatatatat aaaatatata tattaaaaat atatatata	720
taaatatata taaaatatat atattaaaaa tatatatataa atatatatat taaaatatata	780
tataaaatat atattttaa aatatatata aatatatatat attaaaaata tatattaaat	840
ataaatatat atattaaaaa tatatatata atataactat atattaaata tatattaaat	900
ataactatat attaaatata tattaaatat aactatatat taaatatata ttaaataata	960
ctatatatta aatatatatt aaatataact atattttaa tatatatata atataactat	1020
atattaaata tatattaaat ataactatat attaaatata tattaaatat aactatatat	1080
taaatatata tgaaatataa ctatatatta aatatatatt aaatataact atagtatta	1140
aatataaata tatgtcttaa atatatatta aatataaata tatgtattaa atatatatta	1200
aatataaata tgtgtattaa atatatatta aatataaata tgtgtattaa atatatatta	1260
aatataaata tgtgtattaa atatatatta aatataaata tgtgtattaa atactatat	1320
taaatatataa tatatgtatt aatatatatat taaatatataa tatatatataa atatatatat	1380
taaatatataa tatatatata atataaatat atattttaa tatatatatt aaatataaat	1440
atatatataa tatatatatt aaatataaat ataaatataa aatatatatt aaatataaat	1500
acatatatta aatatatgta ttaaatatat atataaaata tatgtattaa atat	1554

<210> 124

<211> 650

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(650)

<223> MAR of chromosome 2 genomic contig; 3944813..3945462

<400> 124

catgatatat tatgtataat atatattata gattacatat aaattatata tataatatat 60

aattatataa tatataatat tatataatat attatatata ttatacaatt atataatata 120

tataatatac aattatataa tatataatat acaattatat aatatataat acaatataat 180

atatatttaa tatattatat aatacatatt taatatatta tatattatat gttatatact 240

aaatatataa tatgtattta atatatacta ttatatatgt aatatattat ataatttatg 300

taacatatta tatattatat atgcaatata ttacatgtta catatatatt acatataata 360

tatgtaatat ataatataca ctatattatt atagtatata atatactata ttatgtaatt 420

atataatata gtatattata cactatatta tattatcata taattatata ttatatacta 480

tattacatat atattatgta atataaatg caatatgtta catatataat atatatgtat 540

tatatagtat atatactata gtatatataa aatatatgct ataatatata ttttatatat 600

tatataatac atataatgta tcatatatta tatataatat attttataat 650

<210> 125

<211> 441

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(441)

<223> MAR of chromosome 2 genomic contig; 5314265..5314705

<400> 125

tataaatata tatgaaatat atataaatta tatataattt atatatacat atataaatta 60

tatataaatt atatataaat tatatatata tatataaatt atatattata tataaaattg 120

tatatattta tatataaatt gtatatataa ttatatata aattgtatat ataatttata 180

tatacaatgt atatattaat ttatatatac attgtatata taatttatat atacattgta 240

tatacaattt atatatacat tgtatatata atttatatat acattgtata tacaatttat 300

atataaatta tattatttat atagtagata tataaatata tatactatat ataaattata 360

tatttattta tatatttat ttttatata taaattatat attatttata tatacattat 420

atataaatta tatattattt a 441

<210> 126

<211> 1169

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1169)

<223> MAR of chromosome 2 genomic contig; 5953971..5955139

<400> 126

SEL PCT 012.ST25

atgtattcat attatatatt tatatatataa taatatatacat tcatattata tatttatata	60
taaataatat atattcatat tatatatatta tatataaata tataatatat ttatgtataa	120
ataatatata tattcatatt atatatttct atataaataa tatatatatt catattatat	180
atttatatat aaatatataa tatatttata tataaatata taatatattt atatataata	240
tatatattca tatttatatat ttatatataa atatataata tatttatata taaataatat	300
atataattcat attatatatt tatatatataa taatatatat tcatattata tatttatata	360
taaataatat atattcatat tatatactta tatataaata atatatattc atatttatata	420
cttatatata aatatatat attcatatta tatatttata taaaaataat atatattcat	480
attatatatt tatatatatt atatatattc atatttatata ttatatatt ctatatattc	540
atatttatata ttatatata aataatgtat attcatatta tatatttata tataaataat	600
gtatatattcat attatatatt tatatatataa tatatatcca tatttatatat ttatatataa	660
atatatattc atatttatata ttatatata aatatatat catatttatat atttatataa	720
aatatatata ttcatattat atttatatat aaatatatat attcatatat atatttatat	780
ataatatata tattcatatt atatatttat atataatata tatattcata ttatatattt	840
atatataaat aatatatata ttcatattat atatttatat ataaataatg tatattcata	900
ttatatattt atatataaat aatgtatatt catatttatat atttatatat aaatatatat	960
attcatatta tatatttgta tataaatata tattcatatt atatatttgt atatatattc	1020
atatatatatt atatataaat atataatatt catatttatat ataaatatat atattcatat	1080
tatatattta tatatatataa taatatatat tcatattatt tatatatata aatatatat	1140
attcatatta ttatatata taaataata	1169

<210> 127

<211> 653

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(653)

<223> MAR of chromosome 2 genomic contig; 6427669..6428321

<400> 127

tatatatgta tacatatatg tatatatgtg tatatatgta tacatatatg tatatatgtg 60
tatatatgta tacatatatg tatatatgtg tatatatgta tacatatatg tatatatgtg 120
tatatatgta tacatatatg tatatatgtg tatatatgta tacatatatg tatatatgtg 180
tatatatgta tacatatatg tatatatgtg tatatatgta tacatatatg tatacatgtg 240
tacatgtgta tacatatatg tatacatgtg tacatgtgta tacatatatg tatacatgtg 300
tacatgtgta tacatatatg tatatatgtg tatacatata tgtatatatg tgtatatatg 360
tatacatata tgtatataag tgtatatatg tgtatatgta tataagtgtg tatatgtgta 420
tatgtatata agtgtatata tgtgtatatg tatataagtg tatatatgtg tatatatgta 480
tacatatatg tatatatgtg tatatatgtg tatatgtata taagtgtata tatgtgtata 540
tatgtataca tatatatgtg tatatatgta tacatatatg tatatatgtg tatatatgta 600
tacatatatg taaatatgtg tatatatgtg tatatgtata taagtgtata tat 653

<210> 128

<211> 414

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(414)

SEL PCT 012.ST25

<223> MAR of chromosome 2 genomic contig; 10890453..10890866

<400> 128

tatatattgt aatatatat atagtaaata tatgtaaata tatatatatt gtaaataatat 60
 atatatttg taaatatatg taaatatata tattttgtaa atatatgtaa atatatatat 120
 ttgtaaata tatgtaaata tatatatatt gtaaataatat gtaaataatat atattttgta 180
 aatatatgta aatatatata tttgtaaat ttatgtaaat atatatatt tgtaaataata 240
 tgtaaataata tatatatatt gtaaataatat atacatatat attttgtaaa tatataaaca 300
 tatatatatt ataaatatat ttataaatat atatatgtta aatatattta taaatatatt 360
 tataatatat atattgtaaa tatgtttata aatatatata ttgtatatat aaat 414

<210> 129

<211> 496

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(496)

<223> MAR of chromosome 2 genomic contig; 13952568..13953063

<400> 129

taatatatcat attatatatt atatatgtta tatataatat acatatattata tattatatat 60
 tgtatatata atatacatat tatatatatt atattgtata tataatatat acatatatata 120
 ttatatattg tatatatat atacatatata tatattatat attgtatatata taatatatcat 180
 attatatatt atatatgtta tatataatat acatatattata tattatatat tgtatatata 240

SEL PCT 012.ST25

atatacatat tatatatatt atattgtata tataatatat atattatata ttatatattg 300

tatatataat atacatatta tatattatat attgtatata taatatatat attatatatt 360

atattattga tatataatat acatattata tattatatat tgtatatata atatacatat 420

tatatattat atattgtata tataatatat atattatata ttatatattg tatatatatt 480

atacatatta tatatt 496

<210> 130

<211> 317

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(317)

<223> MAR of chromosome 2 genomic contig; 16942865..16943181

<400> 130

tctcctagta gttatatata tatatatgtg tatatatata tctcctagta gatatatata 60

tatatatatt ctagtagata tatatatata tatatcctag tagatatata tatatatata 120

tcttagtagt tatatatata tatatatcct aacagttata tatatatata tcttagtagt 180

tatatatata tatatcctag tagttatata tatatatata tcttagtagt tatatatata 240

tatatcctag tagttatata tatatatatt ctagtagtta tatatatata ttatatatta 300

tataatatat atataat 317

<210> 131

<211> 464

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(464)

<223> MAR of chromosome 2 genomic contig; 17217049..17217512

<400> 131

```
acatactata tatatacaca tactatatat actatataca gtatatagta tacatatact   60
atacatatac atatactata catatacata tacatatact aagtatacgt atatacagta  120
catagtatat gtatactata tagtatgtat atatagcata tagtatgcgt atactctata  180
tagcatatag tatgcatata cgctatatag catatagtat gcatatacta tatatagtat  240
agagtatgcg tatactatat atatagtata gagtatgcgt atactatata tatagtatag  300
agtatgcgta tactatatat atagtataga gtatgcgtat actatatata tagtatagag  360
tatgcgtata ctatatatat agtatagagt atgcgtatac tatatatata gtatagagta  420
tgcgtatact atatatatag tatagagtat gtatatatat agta                    464
```

<210> 132

<211> 430

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(430)

<223> MAR of chromosome 2 genomic contig; 19647266..19647695

SEL PCT 012.ST25

<400> 132

tgtaaata tgtaaata ttttatatt atatattata taaaaata atatataata 60
 tataatata aaactatata ttaatataat atatataaac tattatataa atacatatta 120
 aatatattat attttaata tttatatatt aaatataata tatatttaatt atttatatat 180
 taaatatata atatatttaa ttttatata atatatagca tttttatat ttatattata 240
 tataacattt tatatttata tttatattta tatatattta atttatattt atattatatt 300
 tatatttata ttatatataa cataattata tatattttca tattgtatat aataaagaaa 360
 tgtatatttg ttatatataa tatatattat ataatttatt atatattata taatatatat 420
 tatataatat 430

<210> 133

<211> 2131

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(2131)

<223> MAR of chromosome 2 genomic contig; 20481223..20483353

<400> 133

tatatataaa tatatttata tttatatat atttatataa atatattttt atataaatat 60
 atatttaata taaatatctt tatatttaatt atatatttaa tataaatatc tttatattta 120
 atatataattt atatataaat atatatttat atttaataata tattaatatt taatatacgt 180
 ttatatttaa tatatatttc tatataaata ttttatatt aacatatatt tatatataaa 240

SEL PCT 012.ST25

tatatattata tttaatatat ttacatatata atatatattat atgtaatat tttacatat	300
aatatatatta tatttaatat atagcatat gtaatatat ttatatatta taatatattat	360
atataaatat atttatattt aataatattt atatataaat atatttatat ttaatatata	420
ttaaatatat atttatattt aatatatatt aatatattaat atatatattat atttaatat	480
tatttatatat aaacatatat ttatatatta tatatatatt atataaacat atatttatat	540
ttaatatata ttatatataa acatatattt atatttaata tatatttata tttaatatat	600
tatatataaa catatatatta tatttaatat atatttatat taaatatata ttatatataa	660
acatatattt atatttaata tatatttata ttaaatatat atttatattt aatatatata	720
tattaaatat atatttatat ttaatatata ttatatatta atatatattt atattaaata	780
tatttatatt taatatatat ttatatataa tatatatata atatttaata tatatttata	840
ttaatatat acatatatat ttatatatta tatatacata tatatttata tttaatatat	900
acatatatat ttatatatta tatatacata tatatttata tttaatatat aaatttatat	960
ttatatata taaaaatata tatttatatt taatatatat aaatatatat ttatatatta	1020
tatatatatt tatattgaat atatacataa atatatattt atatttaata tataaacata	1080
tatttatatt tatatatata atatatattt atatttaata tataaatata tatttatatt	1140
taatatattt atatatata atatatattt atttaatat ttatatata gatatatatta	1200
tatttaatat atttatgtgt attaatatat ttatatatta tatatttata tattaatat	1260
ttatatattt atatttatat attaatatat ttatatatta tatttatatt ttatatattt	1320
atatattaat atatttatat ttatatatat ttatatatat taataaattt atattttata	1380
tatttatata ttaataaatt tatattttat acagttatat aaatatattt atattttata	1440
cagttatat aatatattta tttttatag ttatatataat atatttatat ttatacagt	1500
tatatataaata tatttatatt ttatacagtt atataaatat atttatattt tatacagtta	1560
tataaatata ttatatattt atacagttat ataaatatat ttatatatta tacagttata	1620
taaatatatt tatattttat acagttatat aaatatattt atattttata cagttatat	1680
aatatatatta tttttatac agttatatata atatatattt attttatata gttatatata	1740

SEL PCT 012.ST25

tatatttatg tttatacat ttatataaat atatttatat tttatacatt tgtatttaaat 1800
 atatatttat atataaatat attttatatt taatatattt atatataaat atatattgat 1860
 atttaatatata tatttatata taaatatata ttgatattta atatgtttat atataaatat 1920
 atatttatat ttaatatata tgtttatata tcaatatata tttatattta atatatattt 1980
 acatataaat atatatttat attgatata tatttatatt tgatatatat ttatatata 2040
 ttaatatatt tacatttgat atatatttta tatatattaa tatatttaca ttgatatat 2100
 attttatata tattaatat tttacatttg a 2131

<210> 134

<211> 842

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(842)

<223> MAR of chromosome 2 genomic contig; 20483478..20484319

<400> 134

tatatattta tgtttaatat atatttatag ataaatatat atttacgttt aatatatatt 60
 tatagataaa tatatattta cgtttaatat atatttatct ataaatatat ttacgttttaa 120
 tatatattta tatattaata tatttatggt taatatatat ttatatatat taatatattt 180
 atgtttaata tatatttata tattaatat tttatgttta atatatttat atatattaat 240
 atatttatgt ttaatatata tttatatggt aatatattta ggtatatata tatttatatg 300
 ttaatatata tttatatata tatatttat ttatatataa aagtatatat aatatataaa 360
 tatttatataa attatttat agtattttta tatatattta tatataaatt ttatatattt 420

SEL PCT 012.ST25

tatatatata aatatatatt tatatataca ttttatatat aaatatatat ttatatatac 480
 attatatata taaatatata ttttatatt ttatatataa atatatatat ttatatatac 540
 attttatata ttttatatat gtaaatatat atataaattt tatatatgtt atatatattt 600
 ataaatttta tatatatatt tatatatata atatatataa tatatatataa ttttatatat 660
 attatatata tttatatttt atatattata ttttatatta tatatatatta tatgttatat 720
 atatttatat ttatatattt ttttatatta tatattttat atatatattt atatatgtat 780
 attatatata ttatatatta tataatatat tatatatatt atattatata tttatattat 840
 at 842

<210> 135

<211> 645

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(645)

<223> MAR of chromosome 2 genomic contig; 20897566..20898210

<400> 135

gtatatattt attatatatt atataatata tatttatatat taataaatta tatataatat 60
 aatatatatg tatatttata ttatgttat aatatacata taattatata tgtatgtata 120
 catgtatata tatacgata tgtgtatatg tatacatata ggtatatgtg tacatgtata 180
 catataggta tatgtatatg tatacatgta tacatataat ataattacat atgtatgtat 240
 acatacatat gtaattatat tatatatgta tatgtatatt tatataatat ataatatgta 300
 ttatatatta tacatgcata ttttatatgta tatttatatat acacatatata tataattata 360

SEL PCT 012.ST25

tatgtatgta tatatacaca tatatatatta tattatatat gtatattata tacatatatt 420

tatattatat atgtatatat atttatcata ttatatgta atatgcatgt gtaataaata 480

atatacacat ttatatatgt atattatata catatatatta tattgtatat gtatatatat 540

ttatatatat ttgtatatca tatatttata tattgtatat ttatgtatat tatatatatta 600

tatattatat atgtattata taatatatat gtaaatatat attat 645

<210> 136

<211> 722

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(722)

<223> MAR of chromosome 2 genomic contig; 21664541..21665262

<400> 136

tataatatat attatattct atataaatg taaaatatat attatattct atataatgta 60

ttatatatag aatataatat attctatgta ttctataatc tatataatac atattatata 120

ttatatagaa tattataaat aatatattct atattatata tagaatatat tctatatgtt 180

tatattctat atattatata tgaaatagta tataaaatat atataatata tataaaatat 240

gatatataat atatataaaa taatatataa tgtataatat ataaaataat atataatgta 300

taatatataa aataatatat aatgtataat atataaaata atatataatg tatattatat 360

aaaataatat ataatgtata ttatatataa aataatatat aatgtatatt atatataaaa 420

taatatataa tgtatataaa ataatatata atatattata tataaaataa tatataatat 480

attatatata aaataatata tattatatat aaaataatat ataatatatt atatataaaa 540

SEL PCT 012.ST25

taatatatat tatatatataa ataatatata atatattata tataaaataa tatatattat 600
 atataaaata atatatatta tatataaaat aatataatat atattatata taaaataata 660
 tataatatat tatataaaaa tataaatata ttatataaaa atataaaata taaaatatta 720
 ca 722

<210> 137

<211> 305

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(305)

<223> MAR of chromosome 2 genomic contig; 22834991..22835295

<400> 137

aatataaaat atatgatata taatacgtat tatatatgta taatacgtat tatatattaa 60
 tatataatat ataatacata ttatatatgt atataatata tactaatata tataatgtat 120
 acattatata ttacataat atataataca taatatagaa ttataattat atataataca 180
 taatatataa ttatatatat tattatatat gtatttatat tatatataat atattatata 240
 taatatatat tatataatta tataagtata taattatgtt atatacataa taatatataa 300
 tatat 305

<210> 138

<211> 352

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(352)

<223> MAR of chromosome 2 genomic contig; 25277762..25278113

<400> 138

taatatatat aatatattat atattatata taatatattt tataatatat aaaatatatt 60

atatataata tataatatat ttataatat atataatata ttatatataa tatataatat 120

atattataat atatataata tattatatat attatatatt tatatttatt tatatatcca 180

taaatatata ttatatata atatatatta taatatatta tatataatat ataatatatt 240

ttataatata ttataatata taatatataa tatattttat aatatatata atatataata 300

tattatatat ttatatttat ttatatattc ataaatatat atatttatat ta 352

<210> 139

<211> 342

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(342)

<223> MAR of chromosome 2 genomic contig; 25378452..25378793

<400> 139

tatgtacata tatattttat atattatata taatatatat tatatgatat atataatata 60

SEL PCT 012.ST25

ttatataata taatatataa aatatatata atatatatta tattatataa attatattat 120
 atatatcata taatatattt tatatattat ataatatata ttatattata tatattttat 180
 atattatatt atatatata tatacatat aatatatatt atattatata ttttatatat 240
 tatataatat atattatata ttttatata ttatataata tatattatat attttatata 300
 ttatataata catatattat atataatata atatatatta ta 342

<210> 140

<211> 663

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(663)

<223> MAR of chromosome 2 genomic contig; 30209437..30210099

<400> 140

aatatatatt acatatgtta tatatagtat atgtaatgta tataatatag tatattctat 60
 attgtataat agtaatatat agtatatgat atactatata ttacttatca tatatacaat 120
 atatatata tcgtatattg tatattatat attgtatata tgtaatatat gatatgtaca 180
 tatgttatat atgtatataa tatactatat tatatattgt atattatata catatataac 240
 actattatac aatatataat atagcatatt atatacaata tagcatatac aatatataat 300
 atagcatatt atatatata tagtatatta tatacaatat ataatatagc atattatata 360
 taatatataa tagtatatta tatacaatat ataatatagc atatacaata tagtatataa 420
 tatataatat agcatatata atatagtata ttatatataa tatataatat agcatgtaca 480
 atatagtatg ttatatataa tatataatat agcatatata atatagtata ttatatataa 540

SEL PCT 012.ST25

tatataatat agcatatata atatattata ttatatataa tatataatat agcatatata 600
 atatagtata ttatatataa tatataatac agcatatata atatagtata ttacatacag 660
 tat 663

<210> 141

<211> 1200

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1200)

<223> MAR of chromosome 2 genomic contig; 31725089..31726288

<400> 141

tgtacttata tattataatg tatatatataa gtatatactt tatatatact tatatattat 60
 aatgtatatt attgtatata agtatatatc ataatatata cttacatatg ctcacatata 120
 ttataatgta tattgtatat attatatata tattatatat gtataatgta tatatacatt 180
 atatatgtat aatgtatata tacattatat atgtataatg tatatatata ttatatatgt 240
 ataatgtata taatatatac aatatatgta taatatataa tatatacaat atatgtataa 300
 tatacaatat atgtataata tacaatatat gtatagtata taatatatat tatatatgta 360
 tagtatatta tatattatat atgtatagta tataatatgt ataatgtata tattataata 420
 tattatatat aatatctata acaatatata atattgtata tattatatat aatatatatt 480
 tatataatat atattatata taatatatta tgtatttatt tatattatat ataatatataa 540
 tatatatata ataaataata ttattatat attaatataa atatttatat taatatatat 600
 ttattatata taaataatat ctatgatata aatatatat aatatacatg tatatgttat 660

SEL PCT 012.ST25

aatatataca tataatatac atgtgtatat atactataca tgtatatata acatgtatat 720
atatacatgt atatatatta tgtatacatg tatagtatat atacatgtat atatatatat 780
atatactata catgtatata tacatgtata tatatacata tatactatac atgtatatat 840
acatgtatat atacacatat atactataca tgtatatata catgtatata tatacatgta 900
tggtatatac attattataa tatacatata tagtatacat tatatacatt atataatatg 960
cattattata atataatata cattattata atatacatta ttataatata atatacatta 1020
ttataatata cattattata atatacatta taataatata cattattata atatacatta 1080
taatattgaa gtatatatac tataatatat gtatatatta taatgtatat aatatacatt 1140
attatatata agtatgtatt atataaagt atatattata atatatgtat atacatatat 1200

<210> 142

<211> 325

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(325)

<223> MAR of chromosome 2 genomic contig; 32147252..32147576

<400> 142

aaatacaa atttatttat atataatata taatataata tttttattta tatataatat 60
ataatttata attatataaa tatataatat atttatatat aatataatat tttattatat 120
attaattata tatataataa atatatataa tatataattt tatttatatat taattatata 180
tataataaat atataataa tataataata ttatatacat tatatatataa tataaatatt 240
tatataatat ataataaat atatttattt atatatataa atataatata taattatata 300

SEL PCT 012.ST25

aatatataat atatttatat ataac

325

<210> 143

<211> 507

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(507)

<223> MAR of chromosome 2 genomic contig; 32312662..32313168

<400> 143

attatttata taaatattat atttatatta ttatataaaa tatttatatt atattattta 60

tataaatatt atatttatat ttttatata aatattatat ttatattatt tatataaata 120

ttatatttat attatttata taaatattat atttatatta ttatataaaa tatttatatt 180

atattattta tataaatatt atatttatat ttttatata aatattatat ttatattatt 240

tatataaata ttatatttat attatttata taaatattta ttatattat ttatataaat 300

attatattta tattatttat ataaatattt atttatatta ttatataaaa tatttattta 360

tatttatata aataatatat aaataaatat ttatatgta tataaatatt atttatatta 420

ttattttaa taaataatat aaattaatat aaatattaat attatttatt ttattataaa 480

taatataaat attatattta tatttat

507

<210> 144

<211> 339

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(339)

<223> MAR of chromosome 2 genomic contig; 33651118..33651456

<400> 144

aaatataata tattatttat atataatata aatgatatat tatgtatata taaaatataa 60

ataatatatt atgtatatat aaaatataaa tattatttat atataaaata taaataatat 120

ttatatataa aatataaata ttatatatt tatatatataa atataaataa tatattattt 180

atatataaaa tataaataat atattattta tatataaata atatataaaa tataaatata 240

tattatatat aaataaaata tatatattat atatataaat ttatatataa tatataaaat 300

ataatatata tatttaatat ttattatata atatataat 339

<210> 145

<211> 461

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(461)

<223> MAR of chromosome 2 genomic contig; 45073053..45073513

<400> 145

tgtgtatata tatatacgtg tacatatata tatatacatg tgtatatata tacgtgtaca 60

SEL PCT 012.ST25

tatacatata tacatgtgta tatatatgta catatacata tatacatgtg tatacataca 120
 tatatacatg tacatataca tatatacatg tgtatacata catatataca tgtacatata 180
 catatataca tgtgtatact tacatatata catgtacata tacatatata catgtgtata 240
 tatacatata tacacgtaca tatatacata tacatgtaca tatatacatg tatacatata 300
 tacatgtaca tatgtacata tatatacatgta tacatatata catgtacata tgtacatata 360
 tacatgtata catatataca tgtacatatg tacatatata catgtataca tatatacata 420
 tgtacatacg cacagataga catatataca tatgtacata c 461

<210> 146

<211> 1162

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1162)

<223> MAR of chromosome 2 genomic contig; 45487691..45488852

<400> 146

attatattat ctatataaat ctattatata tattatatta tctatataat atctattata 60
 tatattatat tatctatata aatctattat atatattata ttatctatat aaatctatta 120
 tatatattat attatctata tatctattat atattatatt atattatatt atatataata 180
 tctattatat atattatatt atattatatt atatataata tctattatat atattatatt 240
 atctatataa tatctattat atattatata ttatattata tataatatct attatataata 300
 ttatattata ttatataata tatctattat atctattata tatattatat atatctatta 360
 tatctattat atatattata tataatatct attatatcta ttatatatat tatatataat 420

SEL PCT 012.ST25

atctattata tctattatat tatattatat ataatatcta ttatatctat tatatatatt 480
 atatatatct attatatcta ttatatatat tatatataat atctattata tctattatat 540
 atattatata taatatctat tataatctatt atatattata tatataatat ctattatatac 600
 tattatatat tatatatata atatctatta tatctattat atctattata tatatatcta 660
 ttatatctat tatatatatt atatacataa tatctattat atctattata tatattatat 720
 atataatatc tattatatct attatatata tactatctat tataatctatt atatataatta 780
 tatatgtact atctattata tctattatat ctattatata tataactatct attatatcta 840
 ttatatatat tatatatata ctatctatta tatctattat atatattata tatataactat 900
 ctattatata tctattatat atattatfff atattatata tagtatctat tacatatatt 960
 atattatatt atatataata tctattatat atattatatt atattataaa taatatatat 1020
 aatatctgtt atatataata gatattatat ataatatata atatataata tagatattat 1080
 atatattata ttataataa tataatatat aatataatta atataaaata tatataatat 1140
 ataattaata taatatgtaa ta 1162

<210> 147

<211> 562

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(562)

<223> MAR of chromosome 2 genomic contig; 45516233..45516794

<400> 147

acatattata tatattatat ataatatata ttatatatac atattatata tattatatat 60

SEL PCT 012.ST25

aatatatatt atatatacat attatatata ttatatatac atatatatat tgtatataat 120
atatacatat tatatatatt atatatacat attatatatt atatataata tatacatatt 180
atatattata tataaatatt atatattata tataaatatt atatatataa atattatata 240
ttatatataa atattatata tcttatatat aaatataata tataatatat ataatttta 300
tatattatat ataaatatta tatatattat ataattatt atataatata taaatata 360
tattatataa atattgtata tattatataa atattatata tattatatat aaatattata 420
tatattatat aaatatatat aaatataata aatatataaa tatgtaaaat ttattttat 480
aaatataata tataaatata taaatataaa tataaattat atataatata taatatatta 540
tacataatat atactatata ta 562

<210> 148

<211> 801

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(801)

<223> MAR of chromosome 2 genomic contig; 45727251..45728051

<400> 148

atatatatat ataatatata catatataga atatatatat tattatattc tatatataga 60
atatatatat agaatatata tatatagaat atatatatag aatatatata tagaatatat 120
atatagaata tatatataga atatatatat atagaatata tatatagaat atatatatat 180
agaatatata tatatagaat atatatatat agaatatata tatatagaat atatatatag 240
aatatatata tagaatatat atatatagaa tatatatata gaatatatat atatagaata 300

SEL PCT 012.ST25

tatatataga atatatatat atagaatata tatatagaat atatatatat agaatatata 360
 tatagaatat atatatatag aatatatata tagaatatat atatatagaa tatatatata 420
 gaatatatat atatagaata tatatataga atatatatat atagaatata tatatagaat 480
 atatatatat agaatatata tatagaatat atatatatag aatatatata tagaatatat 540
 atatatagaa tatatatata gaatatatat atatagaata tatatataga atatatatat 600
 atagaatata tatatagaat atatatatat agtatatata gaatatatat atatagtata 660
 tatagaatat atatatatag aatatatata tagaatatat atatatagaa tatatatata 720
 gaatatatat atatagaata tatatataga atatatatat atatagaata tatatataga 780
 atatatatat atatatagaa t 801

<210> 149

<211> 346

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(346)

<223> MAR of chromosome 2 genomic contig; 50937238..50937583

<400> 149

taaaattata tatattatat ataatatata atatattata tataatatat attataatat 60
 atataatat tattatataa aatatattct atagaatata tattctattt tataatatat 120
 attctattat aatatatatt atatataata tatattctat tataatatat attatatata 180
 atatattcta ttatgatata tattatatat aataacatat attatatata atatatttc 240
 tattatataa aatatatatt atataaaata tatattctat tatataaaat atatattata 300

SEL PCT 012.ST25

taaaatatat attatattat ataaaaatata tattatacta tatata 346

<210> 150

<211> 462

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(462)

<223> MAR of chromosome 2 genomic contig; 55672627..55673088

<400> 150

taaatatata ttatatatta tattatatat aatatattta tatttatata tactataatt 60

tatatataat atatattata tatataatat atttataata tatatcatat aaataatata 120

tatttataat atatatcata taaataatat atatttataa tagatatcat ataaataata 180

tatatttata atagatatca tataaataat atatatttat aatagatatc atataaataa 240

tatatattta taatagatat catataaata atatatatatt ataatatata tcatataaat 300

aatatatatt tataatatat atcatataaa taatatatat ttataatata tatcatataa 360

ataatatata ttataatag atatcatata aataatatat atttataata gatatcatat 420

aaataatata tatttataat agatatcata taaataatat at 462

<210> 151

<211> 401

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(401)

<223> MAR of chromosome 2 genomic contig; 56081352..56081752

<400> 151

tatacatgta tgtattcgta tatgtatggt atatatgtat atgtgttata tacatatata 60

tatatacatg tatatgtggt atatacatat acatatatac atgtatatgt gttatatata 120

tatacatata tacatgtata tgtgttatat acatatatac atatacatgt atatgtgtta 180

tatacatgtg tatgtgtata tgtatatata catatatgtg tatgtgcatg tgtatatata 240

catatatgta tatgtgtata tgtatatata catatatgta tatgtgtatg tgtatacgta 300

tatatacata tatgtgtatg tgtatgtgta tacgtatata tatacatata tgtgtatgtg 360

tatacgtaca tatacatata tgtgtatgtg tatacgtaca t 401

<210> 152

<211> 765

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(765)

<223> MAR of chromosome 2 genomic contig; 56404208..56404972

<400> 152

tatattatat aaagaatata tattatataa tatgtaaaga atatatatta tataatatgt 60

SEL PCT 012.ST25

aaagaatata tattatatat tatgtaaaga atatatatta tatataatat ataaagaata 120
 tatattatat aatatataaa gaatatatat tatatattat ataaagaata tatattatat 180
 ataatatata aagaatatat aatatataat atataaagaa tatatattat atataatata 240
 taaagaatat atattatata taatatataa agaatatata ttatatatta tataaagaat 300
 acatatatat aatatataaa gaatatatat tatatataat atataaagaa tatatattat 360
 atataatata taaagaatat atattatata taatatataa agaatatata ttatatataa 420
 tatataaaga atatatatta tatataatat ataaagaata tatattatat atattatata 480
 aagaatatta tatattatat aaagaatata tattatatat aatatataaa gaataaacat 540
 atatactata tataaagaat atacattata tatactatat ataaagaata tacattatat 600
 atactatata taaagaatat atataatata taaagaatat acattatata taatatataa 660
 agaatatatt atattatata taaagaatac attataatat aaagaatata ttatatataa 720
 tataaagaat acattataat atataaagaa tatatataat atata 765

<210> 153

<211> 443

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(443)

<223> MAR of chromosome 2 genomic contig; 61953416..61953858

<400> 153

ttatatatt atagataaaa ttatattata ttacatgtaa tatataatat gtaaaatata 60
 ttatattaca tatataatat ataatatgta aaatatatta tattacatat ataatatata 120

SEL PCT 012.ST25

atatgtaaaa tatattatat tacatatata atataaaata ttacatatata tatattttac 180
 ataaatatat attatctatt acatatttat tataatgtaaat aatatgtaca tatgtataaaa 240
 tatgtatata ttatatacata tgtatatatt atatatacat atatatgtat atatttatata 300
 tacatatata tgtatatatt atatttatata tacatatata tgtatatatt atatttatata 360
 tacatatata tgtatatatt atatttatata tacatatata tgtatatata ttataaatat 420
 gtataataaaa gatttatatg taa 443

<210> 154

<211> 372

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(372)

<223> MAR of chromosome 2 genomic contig; 62076211..62076582

<400> 154

tatataaat tatatatgta attatatatc agtatatata attatatata attatcaata 60
 tatataatta tatataatta tcaatatata taattatcaa tagatatata taattatata 120
 tataattata tataattata tatcagtata tataactata taattatata tatgtatata 180
 taattatatg tataaattat ctataagtat atataactat aatatatatc aattatatat 240
 acttatgtat aattatatat actgatatat aattatacat aattatatat atcaattata 300
 tataaattatg tataaattata tatacatata tataaattata tatataaatt atatgtaatt 360
 atataaattac ac 372

<210> 155

SEL PCT 012.ST25

<211> 484

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(484)

<223> MAR of chromosome 2 genomic contig; 62158581..62159064

<400> 155

attatatata atataaaaat tatacatatt attttattat atattatata cataatatat 60

atatttcata tataatatat attatatata atataaaaata tatattatgt ataattatat 120

ataaaaatata ttatataatt atatataaca taaaatatat ataatatata attatatata 180

atataaaaata tatatataat ataaaatata tattatatgt aattatatat aatatataaat 240

atatatatata tataaaaat atattatata taattataat ataaaatata tattatatag 300

tatatattat ataaaatata tattatatat aattatatat tatataaaaat atatattata 360

tataattata taatatataaa tatatatgt atataattat atataatata aatatatat 420

aatatatgaa ataagatata tactatatat aatatatata attacatat aagatatata 480

tcat 484

<210> 156

<211> 644

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(644)

<223> MAR of chromosome 2 genomic contig; 68145036..68145679

<400> 156

tatatatatg ctaatatatg taatatatat tatatatatg ctaatatata tatgctaata 60

tataatatat attatatata aatatataat atatatattat ataaatatat aatatattat 120

atataaatat ataatatataa tatatatataat atatactata ttatatatta tgtataacat 180

ataatacata ttgttatat ataatatata tattatatgt tatatattat atattatata 240

taatatataca atatatattta tatattatat gttatatatt atatatata tataatataa 300

cataatatat aatatatatt atattatata ttacatatat tagcaatatt atatatataa 360

tatatatataat atatatataa tatatatataa aatatataat atatatcaaa atataaacta 420

tataatatat aaaaatatat tatatatataat atataaaaat ataaactata taatatataa 480

aaatatatta tatataatat ataaaaatat attatatatt atatatataa atatatata 540

tataatatat aaaaatatat ataaaatata aaaaatatat ataaaatata aaaaatatat 600

aaaataatat aaaaatatata atatatataa atataatata taat 644

<210> 157

<211> 530

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(530)

<223> MAR of chromosome 2 genomic contig; 71257289..71257818

SEL PCT 012.ST25

<400> 157

atatctatta tatttatata ctttatataa attatatcta ttatatattat atactttata 60

taaattatat ctattatatt tatatacttt atataaatta tatctattat atttatatac 120

tttatataaa ttatatctat tatatttata tactttatat aaattatatac tatttatatt 180

atatacttta tataaatata taattatatt tatatacttt atataaatat aattataaat 240

atatttatat actttatata aatataatta taaatatatt tatatacttt atataaatat 300

aattataaat atatttatat actttatata aatataatta taaatatatt tatatacttt 360

ataattatat gttatattha taattatatt tatataattc ataattatat acattatggt 420

tatagttata taatttataa ttatatatac tatatttata ttatatataat ttataattat 480

ataaattata taaattatat aaattatctt taatttatat tatataatct 530

<210> 158

<211> 337

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(337)

<223> MAR of chromosome 2 genomic contig; 73413615..73413951

<400> 158

acttatatta tatataacta tattattgta tattaatata aattaatgat atataatata 60

ttaattatat attattatat gtgatataaa atacttatat ttatactgta tatatgtata 120

tacacacata tatgtatata tgtatatata cacatatgta tatatgtata tgtatatatg 180

SEL PCT 012.ST25

tatactgtat atatgtatat acacacatat atgtatatat gtatatgtat atatgtatac 240

tgtatatatg tatatacata tatacatata tgatatatat cacatatatg tgatatataa 300

atatatttat ataaatataa tattaatatt tatatta 337

<210> 159

<211> 1340

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1340)

<223> MAR of chromosome 2 genomic contig; 77011049..77012388

<400> 159

atgtatttta tatagtatat attatgtatt atattgatat aattatataa caattattta 60

tatataaaat aacaaataaa tatataaaat aataaatata tatttattat taaataataa 120

atatatattt attattaaat aataaatata taaagtaata aatatatatt tatatattaa 180

ataattcata tatatttata tattaaataa ttcatatata tttaaataat taatacatat 240

ttaaataatt aatatatatt tatataatat atatttatat attaaataat taatatatat 300

ttatagatta aattaatata tatttatata ttaaattaaa tttaatatat tatatattta 360

tataatttaa attaataat ttatataatt taatttaatt taatataatt aaaatatatt 420

aaacattata taatatataa tatatttaat atataatata tatttaatat ataatatatt 480

taatataata tatatttaat atataatata tatttaatat ataatatatt taatatataa 540

tatatttaat atataatata tatttaatat ataatatatt taatatataa tatatattta 600

atatataata tatttaatat ataatatata tttaatatat aatatattta atatataata 660

SEL PCT 012.ST25

tatatttaaat gtataatata ttaatatat aatatatatt taatgtataa tatatttaaat 720
 atataatata tatttgatgt ataatatatt taatatatat ttgatgtata atatatttaa 780
 tatataaatat atatttgatg tataatatat ttaatatata atatatattt gatgtataat 840
 atatttaata tataatatat atttgatgta taatatattt aatatataat atatatttga 900
 tgtataaatat atttaatat taatatatat ttgatgtata atatatttaa tatataaat 960
 atatttgatg tataatatat ttaatatata atatatattt gatgtataat atatttaata 1020
 tataatatat atttgatata taatatattt aatatataat atatatttga tatatattta 1080
 atatataata tatatttgat atataatata ttaatatat aatatatatt tgatatataa 1140
 tatatttaaat atataatata tatttgatat ataatatatt taatatataa tatatatttg 1200
 atatataata tatttaatat ataatatata ttgatatat aatatattta atatataata 1260
 tatatttgat atataatata ttttcttatt aattatttat atataatata taaatatata 1320
 ttaattaatt atatattaaa 1340

<210> 160

<211> 937

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(937)

<223> MAR of chromosome 2 genomic contig; 78226855..78227791

<400> 160

tgtgtatata tacatatatg tgtatctatg tgtatatata catatgtgta tatatacata 60

tatgtgtata tatacatatg tgtatatatg tgtatatatg tgtatatata catatatgtg 120

SEL PCT 012.ST25

tatatatgtg tatatatgtg tatatataca tatatgtgta tatatgtgta tatatacata 180
 tgtgtatata tgtgtatata tacatatatg tgtatatatg tgtatatata catatatgtg 240
 tatatatgtg tatatataca tatatgtgta tatatgtgta tatatgtgta tatatacata 300
 tatgtgtata tatgtgtata tatacatata tgtgtatata tgtgtatata tacatatatg 360
 tgtatatatg tgtatatgtg tgtatatata catatatgtg tatatacaca catatatgtg 420
 tatatatgtg tatatataca tatatgtata tatacatata tgtgtatata tgtgtatata 480
 tacatatatg tgtatatatg tgtatatata catatatgtg tatacataca tatatgtgta 540
 tatatgtgta tatatacata tatgtgtata catacatata tgtgtatata tgtgtatata 600
 tacatatatg tgtatacata catatatgtg tgtatatgtg tatacataca tatatgtgtg 660
 tatatatgtg tatacatatg tgtgtatatg tgtatatata catatatgtg tgtatatatg 720
 tgtatatata catatatgtg tgtatatatg tgtatatata catatatgtg tgtatatatg 780
 tgtatatata catatatgtg tgtatatatg tgtatatata catatatgtg tgtatatatg 840
 tgtatatata catatatgtg tgtatatatg tgtatatata catatatgtg tgtatatatg 900
 tgtatatata catatatgtg tgtatatatg tgtatat 937

<210> 161

<211> 1350

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1350)

<223> MAR of chromosome 2 genomic contig; 79287748..79289097

<400> 161

SEL PCT 012.ST25

tatatatatt atatatatag taactgttct attatatata tattatatat atttctgttc	60
tattatatat tatatatatt atattatata ttatatgtaa tatattatat atattataag	120
taatatatta tatatattat atgtaatata ttatatatat tatatgtaat atattatata	180
tattatatgc aatatgttat atatattata tgcaatatgt tatatatatt atagcaata	240
tattatatat attatatgca atatattata tataatatat gtaatattat atattatata	300
ttatatgtaa tatcttatat attatatgta atatattata tatattatat gtaatattt	360
atatatatta tatgtaatat attatatatt atagtaata tattatctta tatatattat	420
atgtaatata ttatatatta tattatatgt aatatatatt atagtaata tattacatat	480
tatatgtaat atatattata tgtaatatat tacatattat atgtaatata tattatatgt	540
aatatattac atattatattg taatatatta catattatat gtaatattat atagtatta	600
tatgtaatat attatatgta ttatatgtaa tatattatat gtattatatg tattatatgt	660
aatatattat atgtattata tgtaatatat tatatattat atgtaattat attatatgta	720
atatattata ttatatatta tatatattat atgtaatata ttatatatta tattatatat	780
attatatgta atatattata ttatatatta tatatattat atgtaatata ttatatatta	840
tattatatat attatatgta atatattata ttatatatta tatatattat atgtaatata	900
ttatatatta tattatatat attatatgta atatattata ttatatatta tatatattat	960
atgtaatata ttatatatta tattatatat attatatgta atatattata ttatatatta	1020
tatatattat atgtaatata ttatatatta tattatatat attatatgta atatattata	1080
ttatatatta tatatattat atgtaatata ttatatatta tattatatat attatatgta	1140
atatattata ttatatatta tatatattat atgtaatata tttatatatta tatatattat	1200
attatatatt atagtaata tattatatatta tttattatat attatatatt atagtaata	1260
tattatatatta tttattatat atattatatt attattatata tataatatat tatattatat	1320
atattatatt atatatatatt ctgttcta	1350

<210> 162

<211> 332

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_difference

<222> (1)..(332)

<223> MAR of chromosome 2 genomic contig; 81142998..81143329

<400> 162

ctatgtatat aactatatat aactattata taacttaata agatatataa ctattatata 60

acttaataag ttatatataa ctattatata taacttaata agttatatat aactattata 120

taacttaata agttatatat aactattata taacttatta agttatatat aactatatat 180

aacttaataa gttatatata actattatat aacttaataa gttatatata actattatat 240

aacttaataa gttatatata actattatat aacttaataa gttatatata actatatata 300

acttatatac aacttattaa gctatatata ta 332

<210> 163

<211> 327

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(327)

<223> MAR of chromosome 2 genomic contig; 84019536..84019862

SEL PCT 012.ST25

<400> 163

actgacagta tacatactgt atatataac agtatgtata catatacagt atgtatacta 60

tatacagtat gtatactgta tatatatata cagtatgtat actgtatata tatacagtat 120

gtatacgtat gtatactgta tatatgtatt atagtgtata tatgtattat agtgtatata 180

tgtattatat atattatagt gtatgtatta tatgtgtata tacatataat atattataca 240

tatacatatg cacaatatgt atatgtatta tatgtattca tatacatata tgtatatgta 300

taatatatgt atacatataa tacacat 327

<210> 164

<211> 407

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(407)

<223> MAR of chromosome 2 genomic contig; 1448030..1448436

<400> 164

tatataatat atattacata tatattatat ctatattatt tatattacat atgtaatata 60

tattatattt atattattta tataatatat tatatatatt atattattta tatgtaatat 120

atttatattg ttatatata ttatattat attatttata tataatacat attatattta 180

tattatttat atataatata tataataaat atataatata tataaaaaata tatatattta 240

atatacttat aatatatatt atatatatta tatataatat atataattgt acatatattt 300

attatatata ttatatatat aatatatatt ataaatataa tatataaata tatttataaa 360

tatatataaa tatttatattt atacattata ttatatata tattata 407

SEL PCT 012.ST25

<210> 165

<211> 1959

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1959)

<223> MAR of chromosome 2 genomic contig; 2117630..2119588

<400> 165

tatacatgtt atagtgata tagtatacta atatataatg tatgtatgtg tatacatata 60
cacatataat atacacatat ataatatata tagtatataa taatgtataa tatataatat 120
ataatataaa atgtatagta tactacatat ttatatatag tatatagtat gcatagtaca 180
tatatactat atatgtagta tactatagtg tatatatagt acaccatata tagtataaat 240
atactatata gtatatgtac tatatatata ctatatagta tatacagtat acatatatag 300
tataacctata ctatatagta tatatagtggt gcgtatacta tatagtatat atagtgtgcg 360
tatactatat agtatatata gtgtgcgtat actatatagt atatatagtg tgcgtatact 420
atatagtata tatagtggtgc gtatactata tagtatatat agtatacata tatagtggtgc 480
gtatactata tagtatatat agtatacata tatagtatgc gtatactata tatagtatac 540
atatatagta tatctagagt atatgtagta tgtatagtat atatagtcta catactgtat 600
atacagtata tatatactct atagtatact atacagtata gtatactata tagtatataa 660
tatatgtata ctatagaaac acactatata tagtatacta tatatactat atactatata 720
ctatatatag tatactatat atactacata ctatatatag tgtatgtata gtatatataa 780
actatatata gtgtatatag tatatatatt atatataata tatattatat tatattatac 840

SEL PCT 012.ST25

tatatattat atgtatatta tagtatatta tactattata tattatatat tatattatat 900
 attatataat ataataaat tatatattat aaaatatata ttttatatt atatatattt 960
 aaatatattta taatatatat ttataaat atatatatta attattttat atataatata 1020
 aaatataata aatatattat aatatatatt tttaaaatat aatatattata tattataaaa 1080
 atataaatat ataataat atatatatta tatagtataa tatataaat gttatatagt 1140
 atcttatact attatactat atatatatta tagtgtatat atagtatact atatatagtg 1200
 tatatagtg atactatagt gtatatagtg tatactatag tgtatatagt gtatactata 1260
 tacactgtat atagtagtgt atactatata cactgtatat agtagtgtat actatataca 1320
 ctgtatatag tagtgtatac tatatacact gtatatagta gtgtatacta tatacactgt 1380
 atatagtagt gtatactata tacactgtat atagtagtgt atactatata cactgtatat 1440
 agtagtgtat actatataca ctgtatatag tagtgtatac tatatacact gtatatagta 1500
 gtgtatacta tatacactgt atatatagta tattatatat actatatatg tatatatagt 1560
 atacatatat attatatata cagtatatat agtatatata ctatgtagta tatatagtat 1620
 atatactata tagtatgtat agtatactat atagtatata tagtatatta tatagtatat 1680
 atactatata gtatatatag tatattgtat atatagtata tatactatat agtatatata 1740
 gtatatgtta tatatagtat attgtatata tagtatacat agtatgtata tatagtatat 1800
 atagtataca tatatagtat gtacacagta tatatagtct atagtatac tacatatagt 1860
 atacatgtat actatactac atatagtata catgtatact atactacata tagtatacat 1920
 gtatagtata ctacatatata tatacatgta tagaatact 1959

<210> 166

<211> 520

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(520)

<223> MAR of chromosome 2 genomic contig; 2119984..2120503

<400> 166

tatgtatgca tcgtatacat atatagtata tatatgtatg catcgtatac atatatacag 60

tatatatagt atgcatcgta tacatacagt atactatata tacagtatat acagtatact 120

atatatacag tatatacagt atactatata tacagtatat acagtatact gtatatacag 180

tatatacagt atatatagta tactatatat acagtatata tactatgtat tctatatata 240

gtatagtgtg catagtatac atatagtata cactatacta tatatagtat actatatata 300

ctctatatag tatatatagt atactatata tagtatatat gtatactata tatagtgtat 360

atatatacta tatatagtgt atatatatac tatatatagt atatatatac actatatatt 420

gtatagtata gtgtatatat agtatagtat atgtatatat acacatgtat acatgtatat 480

atgtatacta atatatacta atatatgtat aaatatatat 520

<210> 167

<211> 954

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(954)

<223> MAR of chromosome 2 genomic contig; 2578285..2579238

<400> 167

SEL PCT 012.ST25

tattatatat aactttataa tatataatat atattatata taactttata atatataata 60
 tatattatat ataactttat aatatataat atatattata tataacttta taatatataa 120
 tatatattat atataacttt ataatatata atatatatta tatataactt tataatatat 180
 aatatatatt atatataact ttataatata taatatatat tatatataac ttataatat 240
 ataatatata ttatatataa ctttataata tataatatat attatatata actttataat 300
 atataatata tattatatat aactttataa tatataatat atattatata taactttata 360
 atatataata tatattatat actatatata atatataact ttataatata taatatatat 420
 tatatactat atataacttt ataatatata atatatatta tatattatat ataactttat 480
 aatatataat atatattata tataacttta taatatataa tgtatattat atattatata 540
 ttatatatta tatataactt tataatatat aatgtatatt atatattata tataacttta 600
 taatatataa tatataatat aatatataac ttataatat atatatcata tattatatat 660
 aactttataa tatatatcat atattatata taactataat atatatatca tatattatat 720
 ataactataa tatatatatc atatattata tataacttta taatatatat atcatatatt 780
 atatataact ttataatata tatcatatat tatatataac ttataatat atatcatata 840
 ttatatataa ctttataata tatattatat ataactttat aatatatatc atatattata 900
 tataacttta taatatatat catatattat atataacttt ataatatata tcat 954

<210> 168

<211> 452

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(452)

<223> MAR of chromosome 2 genomic contig; 3836217..3836668

SEL PCT 012.ST25

<400> 168

tttatatata aatatatatc ttatatatat ttatatataa tacatatata tcttatatat 60

ataaaatata tatacatatt tatatatataa atacatatgt attatataca tttatatata 120

atacatatgt attatataca attatataat acatatgtat tatatacaat tatataatac 180

atatttataa atatatatat ttatatattat atatatttat atataaataa atatatattt 240

atagatttat ttatatataat atatatttat ataaatatat atttatatat atttatataa 300

atatatatatt atatatattt ctatatatat atataaatat atgtataaat atatatattt 360

atacatatat tcatatataat atatatattt atacatgtat ttatatgaat atatatttat 420

acatgtaatt atatgaatat atatttatac at 452

<210> 169

<211> 417

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(417)

<223> MAR of chromosome 2 genomic contig; 3837666..3838082

<400> 169

gatatatata tttatataaa tatatatata aagagatata tttatatatt tatttatata 60

aatatatattc tttatataaa gatatatgta aatatattta tttatataaa tatatttata 120

tatgtaaata tatatttata tatttatata tttatatatt tatttatata aatatatata 180

tttatatatatt tatttatata tataaaaaata tataaatata aatatatata aatatatata 240

SEL PCT 012.ST25

attataaata tagaaataaa tataaatata aatatataaa tatatataaa tataaatata 300

tataaatata aatatatata aatataaata tataaatgta taaatatata aatataaata 360

tatataaata tgtataaata tataaatata taaatatata aaaatatata taaatac 417

<210> 170

<211> 1197

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1197)

<223> MAR of chromosome 2 genomic contig; 6294846..6296042

<400> 170

tatatactaa tatgtatata taaatatata aatatatata cacgtgtata tataaatata 60

tatgtatata taaatatata tacatatatg tatataaaaa tatatacgta tatacgtata 120

tacgtatata tagatatata cgtatatacg tatatacgta tatatagata tatacgtata 180

tacgtatata tagatatata cgtatatacg tatatacgta tacatgtgta tatacgtata 240

tacacatata cgtatacatg tgtatatacg tatatgtata cattatatat acgtatatat 300

acatatatgt atacatgtat atataaatat atacatatat gtatatatta tacatatatg 360

tatatataat atatatatta tatataaat atatatatta tataatatat atattatata 420

taatatatat attatatata atatatatta tattatatat aatatatata tatataaat 480

attatatatg tacatatgta cataatgtat atagtatat atataaatata tatgcacatg 540

tatatataat atagtatat tatatatata tatgtatata tgtacatatt atatatgtat 600

atatgtacct attatatata catatgtata tatgtacctt ttatatatac atagtatat 660

SEL PCT 012.ST25

atgtacatat tatatatata tatgtatata tgtacatatt atatatacat atgtatatat 720

gtgcatgcat atataatata taatatatta tagattataa tattatatac atatcatata 780

ttatatactt atatatacat gtatatatta tatacatatt atatattata tacatataat 840

atatgtatat aatatataca tatattatat attatatata atacattatg ttatatatta 900

tggtatataa tatatattat ataatatata tatattatat ataatatata catatataat 960

aaataatata taattatata tataatatat gcataataat atgtaatata ttttatatta 1020

tatatgatca tatataatat gacatattat atatgattat atatatgata tattatatat 1080

gattatatat attatatata aatatatgat tatatataat catatatata aatatatgat 1140

tatatgatta tatataaata tatatatatg attatatgat tatatataat tgattat 1197

<210> 171

<211> 362

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(362)

<223> MAR of chromosome 2 genomic contig; 6506971..6507332

<400> 171

tatatatagt gtatactata tatacgctat atgcacacat aaactatata tacagtatat 60

aatatgcgta tactatatac acagtatata ctacatgtat actatatata gtatataaga 120

tatatactat gtatataata tatatactag gtatatatat ccatatatat actatatact 180

atagtatata catatatatg tacgtatata tgtatatgta catatatatg tagtatgtat 240

atatatacat atatacacac tatagtatat acatatatat actatatata ccctatatag 300

SEL PCT 012.ST25

agtatattat atacagtata ctatatatac tatatatacc ctatatagag catgtctatg 360

ct 362

<210> 172

<211> 2578

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(2578)

<223> MAR of chromosome 2 genomic contig; 6507395..6509972

<400> 172

ggtatactat atatactata gagtatactt tatagtatat atacctatat tatatatata 60

tacatacact gtatagtata tatggtatat atactatata tggcatatat agtttatata 120

tatactatat atggtatata tagtttatat atatactata tatggtatat atagtttata 180

tataccatat atggtatata tagtttatat agtacatata gtatatatac acactgtata 240

gtatatatta ttagtatat atactatata tactgtatat atagtataaa tactatatat 300

agtatacact atatactata cactatatat actatatact atatactata tatagtatac 360

tatatagtat atagtatact ctatatgtac tatagagtat actatatata ctatacataa 420

aataatttta tatatagtac agcgatact atatactata tatagtatac tctatatgta 480

ctatagagtg tagtatatac tatacagtat actctatata tactatacag tacactatat 540

atactatata tagtatattt tatatatagt acagtatata cagtatatat attatactat 600

atgtagtaca tatatagttt agtatatata gtatatatac tatactatat gtactacata 660

tataatagta tatatagtat atatactata ctatatgtag tacatatata gtttagtata 720

SEL PCT 012.ST25

tatactagta tatagatata tagttatata gatataaat agtatatata gtatatatag 780
 catatatagt atatatgcta tatatactat atagcatata ctatatacta tatatacagt 840
 atatatagca tatatagcat atataatata tatacttttg atatacatat tatatacagt 900
 atatatagta tatatactgt ataaatatac tatatatacc gtatatgcac actatatgct 960
 atatatacta tatacactat atacagtata tatagtacac tatactatat aaagtatata 1020
 tagtatacag tacactatata tatatacatt atatatagta tatattatac atagtatata 1080
 gtatatataat agtatatata gtatatacag tatatatata gcatacttta tatagtatac 1140
 acagtatata gatactatat atgctatata tagtatctat atactgtata ttatatatac 1200
 taatatagta tataigtata tatatactgt atataata tatacatata tagtatatat 1260
 actatacata cacactatac atagtatat atactatata tactatatac tatatatacct 1320
 atatatacta tatagtatat tatatatacct atataatac tatagtatat tatatatacct 1380
 atataatac tatagtatat tatatatacct atataatac tatagtatat tatatatacct 1440
 gtatactata tatactatat agtatactgt atatactata tagtatactg tatatactat 1500
 atagtatact gtataatac tatagtatac tgtatatact atatatagta ctgtatatac 1560
 tatatagtat actgtatata ctatatatac tatatagtat actgtatata ctatatagta 1620
 tactatatat actatatacc atataatac tgtatatact atatatagta tatactatgt 1680
 atagtctata tatagtatat atagtatata tgctatatat agtatatata gtatatatgc 1740
 tatatatac gtctatatat agtatatata ctatatagac tatatatata gcatatatac 1800
 tatatatac atataata tatatggat atacatagta tctatatgta gtatctatat 1860
 atagtaccta tatatactat atagaggtac tatatatagt atataactt tatatagata 1920
 ctatatatag tatatactat ttatatagta tatatagtat atgtagcata tatagtatat 1980
 atagtatata tagtatatag tatgtatagt atatatagat tatattgtat atacagtata 2040
 tatactgtat atactatata aatagtacat acagtatata cagtatatat gtactatata 2100
 tagtatatac agtatatata gtatatatgt accatatata gtatatacag tatatacagt 2160
 atatatgcac tatatgttat atacagtata tacagtatat atgtactata taaatagaat 2220

SEL PCT 012.ST25

atactctata tacagtatat atgtactata taaatatata cactatgtac agtatatatg 2280
 tactatataa atagtatata cactatatac agtatatatg tactatatag tgtatacagt 2340
 atatacagta tataggtact atatatggta tatacagtat atatgcacta tatggtatat 2400
 acagtatata tgcactatat atggtatata cagtatatat gtactatata tggatatatac 2460
 agtatatatg tactatatat ggtatataca gtatatatgt actatatatg gtatatacag 2520
 ttatatacagt atatatgcac tatatatggt atatacagta tacaigtact atatatgg 2578

<210> 173

<211> 598

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(598)

<223> MAR of chromosome 2 genomic contig; 7770400..7770997

<400> 173

gtgtattgta tatacatata cgtatctacg tatatacata tatgtattgt atatacatat 60
 atgtattgta tatacatata tgtatatacg tatatacata tatgtattgt atatacatat 120
 atgtatatac gtatatacat atatgtatat acgtatatag atatacatat atatgtattg 180
 tatatacata tatgtatata catatatata tatatatga tatacatata tatgtattgt 240
 atatacatat acaatatatg tatatatata tatacatata caatatatgt atatacatat 300
 atatgtattg tatatacata tatatgtatt gtatatacat atattgatat acatatatgt 360
 atatatatat atatgcatat atgtatatat acatatatgc atatatgtat atatacatat 420
 atacatatgt acatatatac atatatatat atatgtatat atacatatat acatatgtac 480

SEL PCT 012.ST25

atatatacat atatacatat gtacatatat acatatatac atatgtacat atatacatat 540

atagatatat atacacatat atagatatatac ttatatgtat atatacatac atacatat 598

<210> 174

<211> 1048

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1048)

<223> MAR of chromosome 2 genomic contig; 8332422..8333469

<400> 174

cattatatat aatatataat atattattat atataatata tataacatta tatatagtat 60

atatgacata tataacatat attatatata acatatataa aatataacat attatatata 120

acatatataa aatataacat atattatata taacatatat aaaatataac atatattata 180

tataacatgt ataaaatata acatatatta tatataacat gtataaaata taacatatat 240

tatataacat gtataaacta taacatatat tatatatataa atatattata tgttatatat 300

tataaataaa atatattata tgttatatat tataacatat tatataaata atatataata 360

tataacatat attatatataa taatatataa catatattat ataaataata tataacataa 420

catatattat ataacatata acatataaca tatattatat ataacatata acatataaca 480

tatattatat ataacatata acatataaca tatattatat ataacatata acatatatta 540

tattatatat aacatatataac atatattata ttatatataa catataacat atattatatt 600

atatataaca tataacatat attatatatt atataacata taacatatat tatattatat 660

ataatatata acatatatat tatatatata atataacata taacatatat tatatatata 720

SEL PCT 012.ST25

ataatatata acatatatta tatataatat aatatataac atatattata tataatataa 780
 tatataacat atattatata taatataata tataacatat attatatata atataatata 840
 taacatatat tatatataat ataatatata acatatatta tatataatat aatatataac 900
 atatattata tataatataa tatataacat atataatata taacatatag catatataat 960
 atataacata taacatatat tatatataac atataacata tattatatat aacatataac 1020
 atatataata tgtaacatta tatataac 1048

<210> 175

<211> 375

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(375)

<223> MAR of chromosome 2 genomic contig; 8909678..8910052

<400> 175

tatatacaca tatatacgta tgaatatata tacacatata cgtatgaata tatataccca 60
 tatacgtagt aatatacaca tatatatacg tacgtatata tatacacata tatacgtacg 120
 tatatatata cacatatata cgtacgtata tatatacaca tatatacgta cgaatatata 180
 tacacatata tacgtacgaa tatatataca catatatacg tacgaatatata tatacacata 240
 tatacgtacg aatatatata cacatatata cgtacgaata tatatacaca tatatacgta 300
 cgaatatata tacacatata tacgtacgaa tatatataca catatatacg tacgaatatata 360
 tatacacata tatac 375

<210> 176

SEL PCT 012.ST25

<211> 563

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(563)

<223> MAR of chromosome 2 genomic contig; 10572503..10573065

<400> 176

atttataata tatatgtata aatatatgta tatatttata tttaaataata tgtatatata 60

tttatattta aatatacgtata tatatattta tatttaaata tacgtgtata ttttatatt 120

taaataatagc tgtatatatt tatattttaa tatacgtgta tatatttata tttaaataata 180

cgtgtatata ttatatattta aatatacgtg tatatattta tatttaaata tacgtgtata 240

tatttatatt taaataatagc tgtatatatt tatattttaa tatacgtgta ttttatatt 300

taaataatagc tgtatatatt tatttaaata tatgtatgta ttataaata tatattttaa 360

gtatatattt ataaatgtat acatgtatat ataaatatat atattttaa tatatatatt 420

tatatatatt tatatatatt tataagtata tatatatatt aatatatgta tatatttata 480

tatttatata agtatatata tttaaataata tgtatatatt tataatatat attttaaata 540

tatatttata tatttatatt ata 563

<210> 177

<211> 595

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(595)

<223> MAR of chromosome 2 genomic contig; 11609694..11610288

<400> 177

tataaatact atatatagta tatataatat tatatatact atatataaat atagttagta 60

taaataatat ataatataga tatataatat aatataatat gttataaata taaatatatt 120

tatataattt aattttataat atataatata taatatataa tttaatttta taatatataa 180

tatataattt aattttataa tatataatat ataatatgta aattatatat aatttaatat 240

atctaaatta tataatttaa atataaatat aatataaata tatctaacat aatatacata 300

acataaatat atatagtata tatagtacat ataaatatat atagtacata tagtatatat 360

aaatatatag tatatataaa tatagtatat ataaatatat agtatatata tagtatatat 420

aaatatatag tatatataaa tatatatagt atatataaat aatatatagt atataaataa 480

tatatattat taaatataat aataatttat tatatatact atatatatt atgtattata 540

ttatatatat tattttatat ttaatatata ttattttata tattatattt aatat 595

<210> 178

<211> 662

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(662)

<223> MAR of chromosome 2 genomic contig; 12699804..12700465

SEL PCT 012.ST25

<400> 178

gtatatatat atatatatat atggtgtata tatatatata tatatatggt gtatatatat 60
 atatatatat atggtgtata tatatatata tgggttatat atatatatgc tgtatatata 120
 tatggtatat atatatggta tatatatatt tgctatatat atagcagatc tgctatatat 180
 atatatttgc tatatatata gcagatctgc tatatatatt tgctatatat atgctatata 240
 tatgctacat atagctata tatatgctat atatatgcta tatatatgct atatatatgc 300
 tatatatatg ctacatatat gctatatata tgctacatat atgctatata tatgctatat 360
 atatatgcta tatatatgct atatatatat gctatatata tgctatatat atagctata 420
 tatatgctat atatatatgc tatatatatg ctatatatat gctatatata tagcatatat 480
 atatagctat atatatgcta tatatatagc ttatatatat gctatatatg ctatatatat 540
 gctatatata tagctatata tatgctatat atagctatat atagctaca tatatgctat 600
 atatatgccca tatgtatgct atatatatgc tatatatata tgctatatat atgctatata 660
 ta 662

<210> 179

<211> 649

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(649)

<223> MAR of chromosome 2 genomic contig; 12821904..12822552

<400> 179

SEL PCT 012.ST25

tatgtaatat tatatatata aattatatat tatacatatg taatattata tatatataaa 60
 ttatatatta tacatatgta atattatata tatataaatt atatattata catatgtaat 120
 attatatata tataaattat atattatata tatgtaatat tatatatata taaattatat 180
 attatacata tgtattatat atataaatta tatattatac atatataata tatatataaa 240
 ttatatatta tacatgtata atatatatata attatatatt atacatatat aatatatata 300
 aattatatat tatacatata taatatatat aaattatata ttatacatat ataatatata 360
 taaattatat attatacata tataatatat ataaattata tattatacat atataatata 420
 tataaattat atattatata tatataatat atataaatta tatattatac atatataata 480
 tatataaatt atatattata catatataat atatataaat tatatattat acatatataa 540
 tatatataaa ttatatatta tacatatata atatatatata attatatatt atacatatat 600
 aatatatata aattatatat tatacatata taatatatat aaattatat 649

<210> 180

<211> 3191

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(3191)

<223> MAR of chromosome 2 genomic contig; 15356889..15360079

<400> 180

tacaattata tataactata aatataatat aatatatatt atctatatta catattaata 60
 tataatatat attacctatt aatatataat ataatatata taatatatat tacctattaa 120
 tatataatat aatatatata atatatatata cctattaata tataataaaa tatatatat 180

SEL PCT 012.ST25

atataffaca tattattata taatatatat tatataacat atataacata tactatatat 240
 tatataacat atataattgt atagtatta tatatattat atatacttat acataatata 300
 taaataatta aatatatgtt ataaatataa caaatatata acatatataa catatataac 360
 atatatataa ttacataaaa tatataatac ataatatata ttatgcaaca tattatataa 420
 tatataacat ataagtata ttatattata tcatatataa tacataatat ataatatatg 480
 atataatata atatattata tatgatataa tataatatat tatatatgtt ataataaat 540
 atatattata tataggatat attataacat attacatatg atataataaa tttatctta 600
 tatataggat atattataat atacacata tagcatatat taaaatatat tacatatagt 660
 atattatata tactatatgt atatatacat atagtatatt atagtatatt atacagtata 720
 tattatatat actatatata gtagtatata gtatatatta tatatactat atatagtagt 780
 atacagtata tattatacag tatatattat atacactata ttatatatta tgtataatat 840
 atactatata tagtatatta tgtagtatat attaaacata atagatatat agtatatact 900
 atagataata gatattatat agtatatagt atatattata tataatatat ataatatata 960
 ttatatacat atagtatata tgatatatta tatataatat atataatata taatatatgt 1020
 aatataatac atattatata taatatatgt aatataatat aatatataat atagtataa 1080
 taataatata tattatataa tataacatat ataaatataa taatatatat tatatgatat 1140
 aacatacata aatataataa catatataat atatattata tattatattg tatatatgat 1200
 atactatata ttacacatta tacattattt ataatatata attaatatat aacatatatt 1260
 agataacata taattatatt tgtaacatat ataagatata attacatata taacatatat 1320
 aattatatat atatttatct aattatatat gaaattatat atgacatatata aaattatata 1380
 ttatatatgt tatatgtatt atatattata tatgttatat atgttatata taacatatat 1440
 aacatatata acacacacat ataacatatata taacatatat tacatatata acatatataa 1500
 cacatatata attatctaac atagataata tatataatat ataataaac atatattata 1560
 tatattatac actctattat attatatata ttatacataa tatataatat atagtatata 1620
 atataatata ttgtatatat gatataatat atattgtata tagtataata tacatatata 1680

SEL PCT 012.ST25

gtatattatg tataacataa tatatagtat attatgtata acataatata tagtatatta 1740
 tgtataacat aatatatagt atattatgta taacataata tatagtatat tatgtataac 1800
 ataatatata gtatattatg tataacataa tatatagtat attatgtata acataatata 1860
 tagtatatta tgtataacat aatatatagt atattatgta taacataata tatagtatat 1920
 tatgtataac ataatatata gtatattatg tataacataa tatatagtat attatgtata 1980
 tataatatat atattatata gtatattatg tatatataat atacatatta tatagtatat 2040
 tatgtatata taatatatcat attatatagt atattatgta tatataatat acatattata 2100
 tagtatatta tgtatatata atatacatat tatatagtat attatgtata tataatatat 2160
 atattatata gtatattatg tatatataat atacatatta tatagtatat tatgtatata 2220
 taatatatcat attatatagt atattatgta tatataatat acatattata tagtatatta 2280
 tgtatatata atatacatat tatatagtat attatgtata tataatatat atattatata 2340
 gtatattatg tatatataat atacatgta ttagtatat tatgtatata taatatatcat 2400
 gttatgtagt atattatgta tatataatat acatggtatg tagtatatta tgtatatata 2460
 atatatataa ggtgtatata tattatgtat atataatat taaggatat atattatgta 2520
 tatataatat atataagggt ttatatataat gtatatataa tatataagggt atgtatatta 2580
 tgtatatata atagtatat tatatataat atatatatt tatatacatat atgtatctat 2640
 ataatatata ttatgtatat attaggtatc tatataatat atattatgta tatatattat 2700
 gtatctatat aatatatata ttatgtatat atattatgta tctatatat atatatatta 2760
 tatgtatatt atgtatctat ataatatata taatgtatat agatatatta tatattatgt 2820
 atatatatta tgtatctatt ttatatataa tgtatataga tatacaatat atattatgta 2880
 tatattatgt atctatatata tatatattat ttatatagat atatatatta tgtatatata 2940
 cataatatat facatattat gtatatatat ataatatata atatatattg tatatatata 3000
 taatatataa tatattatat attacatata ttatatataa tatattatat tatgtatata 3060
 tattatgtat atataatgta tatataatat ataaagtga tatatattgt gtatatataa 3120
 tgtatatata ttacatatat tatgtgtata tatattatata tactacatta 3180

SEL PCT 012.ST25
3191

tacataatat g

<210> 181

<211> 314

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(314)

<223> MAR of chromosome 2 genomic contig; 728676..728989

<400> 181

tgtgtatata tgtatatata atatatatta tataatatgc atagtataa aatatgtata 60

ttatatatgt atatffffata tataatgtata tattatatgt atatffffata tatgtatatt 120

ttatatatat gtatatatta tataatgtata tttatatat atgtatatat tataatgtata 180

ttttatatat atgtatatat tatatatgta tttttatat atagtatat attatatatg 240

tatatffffat atatatgtat attttatata tatgtatatc atatatatgt atatatattata 300

tatatgtata tctt 314

<210> 182

<211> 423

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(423)

<223> MAR of chromosome 2 genomic contig; 737493..737915

<400> 182

ataatatata gtgtctttta tattatctaa tatgtaatat aatgtatfff atattatgta 60

ttttatatta tataatatat aatataatgt attttatatt atagtgtata taatatatag 120

tgcattatat attatgttat attatatata ttttatttat ataaattata tattatatgt 180

tattttatat atattatata acatataata taacaatgca ttatatatta taaaatatat 240

aatacattac atatattata taatatataa tacattacat atattatata atatataata 300

cattatcata tattacaaat attacattag tataatagta attataatat aatatattat 360

atattacata tattatatta atgtaatagt aattataata taatatatat tatattttat 420

att 423

<210> 183

<211> 724

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(724)

<223> MAR of chromosome 2 genomic contig; 1069556..1070279

<400> 183

tattataata tattatatac atttatattgt atatatacta tatatggtat atatagtata 60

cataatataa aatgtataff gtaatataca ttatatatat acatagtgtta cattatataa 120

SEL PCT 012.ST25

tataatataa tgtatattat aatatacatt ataataaat agtgtactat gtatatagta 180
 tatataatgt atattataat gtattatata gtataatata atataatata cattatatag 240
 tattgcatta tatatgctat ataatatata atatattatg tatatatata ttatatatac 300
 tatattatat agtacatata atgtatatta tatagtatat ataataaat acattatata 360
 tacaatatat aatgtatatt atatagtatg tataatgtaa tacattatac atagtacata 420
 aagtatatta taatatatta taatatataa tatacattat atattataat gtatataata 480
 tattgtatat atactatata taatgtatat acaattatat ataattgtat atatacatgt 540
 atatgtatat gtatatatac atgtatatgt atgtgtatat atacatatat gtatatgtat 600
 gtgtatatat gtatatgtat atatgtatat gtatacgtat atatgtatat acaatgtata 660
 tataatgtat ataaaaatat ataatatata caatatgtat ataatgtata taattatata 720
 atat 724

<210> 184

<211> 383

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(383)

<223> MAR of chromosome 2 genomic contig; 2719918..2720300

<400> 184

atatttatat ttatatatt atttatatat aaatatatat ttatatatta tatattattt 60

atatataaat atatatttat attttatata ttatttatat ataaatatat atttatattt 120

tatatattat ttatatataa atatatatatt atattttata tattatttat atataaatat 180

SEL PCT 012.ST25

atatttatat ttatatatt atttatatat aaatatatat ttatatatta tatattatt 240

atatataaat atatatttat atttatata ttatatatt atatattata tatattata 300

ttaattgtg tataatatat attataaat ataataaata tattatttt tatatattat 360

ataaaaatat ataatatata aaa 383

<210> 185

<211> 309

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(309)

<223> MAR of chromosome 2 genomic contig; 4994249..4994557

<400> 185

tataatatat aattgttata acattataac aattatatat tatatataat acaattatat 60

aatatatatt atataattgt aatatataat ataattatat aatatatatt atataatata 120

atatataata tatcatatat gttatatatt ttattatata atatatatta tatataatat 180

tatatataat atatattata tataatatta tatataatat atattatata taatatatt 240

atatatatta tatataatat atattatata ttaaatatta tatatataat atatataaca 300

ttattgtta 309

<210> 186

<211> 740

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(740)

<223> MAR of chromosome 2 genomic contig; 5034916..5035655

<400> 186

```

tttatatata aaatattata tataatatta tatataatat ttctatata aaatgtgtat   60
ataattatat ataattatat aaaatataat atagaatatc taataatgta taatatataa  120
catataaaaa taatattatt taatatataa tattttatat ataatatatt tatatataat  180
ataatatata ttftatatat aattattaat tatataatta atatataata tatattttat  240
acataattat taattatata taattaatat ataatatatc ttatacataa ttatcaatta  300
tatataatta atatataata tatattttat acataattat taattatata taattaatat  360
ataatatatc ttatacataa tatatataaa tatattatat ataatatata ttatatataa  420
tattatatat aatatatatt atatataata aatttatata taatattata tataatatta  480
tatattttat atacaatatg atatataata taatttatat attatatata ttatatata  540
attattatat aaattatata aatataaatt atatatttat atataattat tatataaatc  600
attatataat tattataatt ataatatata atataatata atattatata taatatatag  660
tattctatat aaataatata acatatattt tatatagaat attatatata atataatata  720
tattttatat agaattattat

```

740

<210> 187

<211> 847

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(847)

<223> MAR of chromosome 2 genomic contig; 6074678..6075524

<400> 187

```

aatatagaca taaatatata tgcataaata tatatatgca taaatatata taaaaatata   60
tataaatata tacataaata tatataaata tatacaaaaa tatatataaa tatataaaaa   120
aatatataaa tatatataca catatatataa tatatataca tacatatata aacatatata   180
cataaatata tatgtataaa tatatataca cataaatata tgtatgaata tatatacata   240
aatatatatg tataaatata tatacataaa tatataaaga tatatacata aatatatata   300
aatatatata cataaatata tataaatata tataaataga tatataaata tatatataaa   360
tatataaata tatatataaa tatataaata tataaaaaata gatataataa tatatatata   420
aatatataaa tatatatata aatatatata aatatataaa tatatatata aatatatata   480
aatatataaa tatatataaa tatataaata tatatataaa tatatataaa tatataaata   540
tatataaata tatataaata tataaatata tatataaata tatataaata tataaatata   600
tatataaata tataaatata taaatatata tataaatata taaatatata taaatatata   660
taaatatata aatatatata aatatatata aatatataaa tatatataaa tatatataaa   720
tatatataaa tatataaata tatataaata tatataaata tatataaata tatataaata   780
tataaatata tatataaata taaatatata taaatatata aatatatata taaatatata   840
taaatat

```

847

<210> 188

<211> 784

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(784)

<223> MAR of chromosome 2 genomic contig; 6108986..6109769

<400> 188

```

atttattat atatttaata tataaaatat atatttaata tataaaatgt atatataac   60
atatattata tataatacaa tatatattat atataatata tattatatat aatattatat   120
attatattat aatataatat atattatata taatataata tatattatat attattatat   180
ataatataat atatattata tattattata tataatataa tatatattat atattattat   240
atataatata atatatatta tatattatta tataatataat aatatatatt atatataat   300
tttatatata taatatataa tatatatatt atatatatat ttatatata taatatataa   360
tatatatatt atatataat ttatatata taatatataa tatatatatt atatataat   420
tttatatga taatatataa tatatatatt atatatatat tatatatata taatatgtaa   480
tatatatatt atatataat tatatatata atatatatta tacataaaat atatattata   540
tataatatat ataatatata ttatatataa aatatatttt atgtataata tatattatat   600
ataatatata atgtatatt atatataaaa tatatatatta tatacaatgt atatttatat   660
ataaaatata tatttatata caatgtatat ttatatataat atgtgtttaa tatatgaaat   720
atatatttat atataatata tatttaattct ataaaatata tattaaatat atatttatat   780
ttaa

```

784

<210> 189

<211> 381

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(381)

<223> MAR of chromosome 2 genomic contig; 10389032..10389412

<400> 189

tatacacata tagagtatat agagtatata tagagtatat ctatagagta tatatgtata 60

tagagtatat aatacagcct accatatata tagtatacat atatatatac tctatatact 120

atatatatag tgtgtatata tatagtatag accctacat atatatatat aggagtatat 180

atatatacac actcctacta tatatagtat gtatatagag agtatataga gtatatatac 240

agtatatata cacagtatat atatgccata tatagtatct atatacttat atatagtatg 300

tatctatata ottatatata gtatgtatct atatactata tatagtatgt atctatatac 360

tatatagagt atatatgtat a 381

<210> 190

<211> 507

<212> DNA

<213> Homo sapiens

<220>

<221> misc_difference

<222> (1)..(507)

<223> MAR of chromosome 2 genomic contig; 11097807..11098313

<400> 190

SEL PCT 012.ST25

aattatatat aatttattat atataatfff atatttataa tatttttata tacatatfff 60

atatactfff ataattatat attacatata taatattata taatatatat aatatatata 120

atatatatta tatattatat aatatatatt atatatatta tatataatat atataatata 180

tataatatat ataatatata taatatataa tatatattat ataatatata ttatatataa 240

tatatattat atataatata tatttatatat aatatataat atatatataa tatataacat 300

ataataatat attatacata atttatatat aatttttata taattatata ttttatata 360

ttttatata atttatataa ttttatatt tttatataat tatatatatt tatatatfff 420

tatataatta tatatatataa ttttatataa atatatataa ttttatataa ttttatataa 480

ttataaaata tataattata tataatt 507

<210> 191

<211> 329

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(329)

<223> MAR of chromosome 2 genomic contig; 11234628..11234956

<400> 191

ttatagttaa atatatataa ataaaatata cagttttata cagtatatat aaaatatata 60

atatataata cataatacat tagttatata tactatatat actatatata ctacacgtat 120

agtatatata tgaaactata tatatactat acgtgtagta tatatatgaa actatatata 180

tactatacgt gtagtatata tatgaaacta tatactatac gtatagtata tatatgaaac 240

tatatatact atatatactt aactataatt gtatatagtt aaaaatataa atataaaata 300

SEL PCT 012.ST25

tacagttaaa tatattaata tataatagt 329

<210> 192

<211> 584

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(584)

<223> MAR of chromosome 2 genomic contig; 797844..798427

<400> 192

tattatttta tgttataaat agataaaaat atatactaata atatatgtac ttatatatac 60
atcaatatat aatgtattat ttatactaa cgtatattat atatactagt atataatcta 120
tattatttta tatgttataa atatataata aaatatataa atattttatg catatatataa 180
tatataatat atactaacat gctaatttat atatacttat atataattta tatagtatat 240
aatatataaa tgtatataat acataattta tatatttata tattaatagt ttatatatta 300
gtatatatac taattttata tactaataaa taaattatat aatatataaa ttatatatta 360
tagtacataa tatatattat atagttaaata aactatgttaa ctataatata taactatata 420
tgatatacag ttatatataa tataaatttt acatacagta tataaattat atactatata 480
tttatatata tatggatat aaattatata ctatacattt atatacatat ggtatataaa 540
ttgtatacta tataatgtgt attagtatat atactaatat atac 584

<210> 193

<211> 363

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(363)

<223> MAR of chromosome 2 genomic contig; 1093824..1094186

<400> 193

tatacacaca catatatata cacatatata tacacatata tatatacaca tatatatata 60

catatatata cacgtatatata tgtatacaca tatatatgta tatatatata catatatata 120

cacatatata cgtgtatatata cgtatatatcg tacatatata cgtgtatatata cgtatatgcg 180

tacatatata cgtgtatatata cgtatatgcg tacatatata cgtgtatatata cgtatatgcg 240

tacatatata cgtgtatatata cgtatatgcg tacatatata cgtgtatatata cgtatatgcg 300

tacatatata cgtgtatatata cgtatatgcg tacatatata cgtgtatatata cgtatatgcg 360

tac 363

<210> 194

<211> 545

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(545)

<223> MAR of chromosome 2 genomic contig; 3456187..3456731

SEL PCT 012.ST25

<400> 194

tattataata tatatttata tattataata tatatttat tatatattta tatattataa 60

tatatattat attatatatt tatatattat aatatatatt atattataat atatatattat 120

attataatat attatattat aatatatatt atattattat atattataat ataatatata 180

ttataatata tattatatta taatttatat attatatata ttataatata tattatatta 240

tatatatatt tatattataa tatatattat tatatattat atattataat ttatattata 300

ttacaatata tattataaat atatatatta tattataaat atatatattt atattacaat 360

atatattata aatatatatt ttatattaca atatatatta taaatatata tattatatta 420

caatatatat tataaatata tattatatta caatatatat tatattataa tatatattta 480

tatatgatat attatattta atatatatta taacataata tataatatat aatatattaa 540

tataa

545

<210> 195

<211> 356

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(356)

<223> MAR of chromosome 2 genomic contig; 5001567..5001922

<400> 195

tataaaatat atgttatata tataatatat attatataat atataatata tataatatat 60

aaaatatata aaatatataa tatataatat aatatataat atatatata tataaaatat 120

atataatata aaatatatat aatatataat atatatata tataatacat ataatatata 180

SEL PCT 012.ST25

atatataata tataatatat ataatatata atatataata tataatatat ataatatata 240

atatataata tatataatat ataatatata atatataata tataaatata taaatatata 300

tacacacata cacacacata tatgcatata tatacatata catgtgtaca tagata 356

<210> 196

<211> 321

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(321)

<223> MAR of chromosome 2 genomic contig; 5457330..5457650

<400> 196

tatacaatat attataaatt atatataatt tatatataat atatattata tataaattat 60

atataattta tataatatat aaattatata taatataaat tatatataat ttatataata 120

tataaattat atattatata aattaaatat aatttatatt atatataaat tatatttaatt 180

ttatataata tataaattat atttaattta tatataatat aaatttatatt tttatatatt 240

atgtataatt tatatattta tacatatata cattataata tattgtatag tatatataat 300

atatagtata tataaagcat a 321

<210> 197

<211> 361

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(361)

<223> MAR of chromosome 2 genomic contig; 8124469..8124829

<400> 197

tatatataat atatattata tatattatat aaattatata taatatgtaa tataaatttt 60

gtaatatataa ttatatatat aaattatata taatatatat taatatatat aatataaatt 120

aatatatata atatataatt atatataatt tatatgatat atataaatat atattatata 180

taaattatat atatacataa ttatatatca tataaattat atataatata cattatgtac 240

ataatatatg atatataata tataatatat attatatata attatatata tataattata 300

taatatatat aaattataat atataatata tataaattat aatatataat atataataat 360

t 361

<210> 198

<211> 418

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(418)

<223> MAR of chromosome 2 genomic contig; 11151485..11151902

<400> 198

atgtaactat atatatagta tatatagtat atatatacta tatagtgtgt atatatagta 60

SEL PCT 012.ST25

tatatatact atatagtggt tatatatagt atatatatag tgtatatatc gtatatatac 120
 tatatactat atagtggtata tatagtatat gtagtatata tagtatatat agtatagtat 180
 atatagtata tatagtggtat atatactgta tatatagtggt acatagtata ctatatagta 240
 tacatatagt acactgtata gtatatatag tatagtatat atagtataca tagtatacta 300
 tatatagtat agtatacata gtatactata tagtatatag agtatatata cagtatacta 360
 tatagtatat agagtatata tacagtatac tatatcgtgt gtatagagta tatataca 418

<210> 199

<211> 394

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(394)

<223> MAR of chromosome 2 genomic contig; 13591477..13591870

<400> 199

ttatatatat ttatatata ttatatatat ttatatata ttatatatat attatatata 60
 tattatatat aattatatat aatatatatt atatataatta tatataatta tatataatat 120
 atattatata tattatatat ataatatata tataatatat atattttata tatgtattat 180
 atatatttta tatatattat atatattata tatatatattt atatatatta tattttatat 240
 atataatata acatatataa tatataatta tatattatat atatattata ttatatataa 300
 tatatattat atataatata atatataatt atatataatta tatattttat atattttat 360
 aaaaattatt ttattatttt ttatatataa atat 394

<210> 200

SEL PCT 012.ST25

<211> 1194

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1194)

<223> MAR of chromosome 2 genomic contig; 14996824..14998017

<400> 200

```

taatatttat atatacatat aaaatttata tataatatat aatatttata tatacatata   60
aaatttatat atatataata tatttatata tacatatataa atttatatat aatatataat   120
atttatatat acatatataaa ttatatata atatataata ttatatata catataaaat   180
ttatatataa taaatattta tatatacata taaaatttat atataattta tatataacat   240
ataatattta tatataaaat ttatatataa catatattta tatataattt atatataaca   300
tataatattt atatataata tataatttat tatacaattt atatataata tataataactt   360
atatatacat acataattta tatgatatat atttatata taatttatat gatatatataat   420
atatctaata tatatttat atatttatata tatttatatat aatttatata atatataatta   480
tatatataat ttatataata tatatattat atatataatt tatataatat atatattata   540
tatataattt atataatata tatttatatat ataatttata taatatatat tatatatata   600
atttatataa tatatattat atataattta tatataacat atttatata catatataat   660
ttatatataa tatatattta catatacata tataattttt atataatata aaatatttct   720
atatacatat ataattttta tataatataa aatatttcta tatacatata taatttttat   780
ataatatata ttctatata catgtctaatt ttatatataa tatatatttc tatatacata   840
tataattttt atataatata taatattttt atatacataa ttttatata atatataatt   900

```

SEL PCT 012.ST25

acatatacat atataatfff tatataatat atatttatat atacatatat aatttttaca 960
 taatatatat tatatatata tatataatff atatacaaca tataatatat acatatataa 1020
 tttatatata acatataata tttatgtata catatatataat gtatacacia tatataatat 1080
 ttatatatac atataataatt tatatgtaat atatacatat ataatttata tgtaatatat 1140
 atacatgtat aatttatatg tagtatatat acatgtataa tttatatgta gtat 1194

<210> 201

<211> 487

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(487)

<223> MAR of chromosome 2 genomic contig; 14998429..14998915

<400> 201

tagtatatac ttacacatac atgtataatt atatgtaata tataatatff acatatataa 60
 ttatagataa tatatatffa catatacata tataattata tataatatat aatgtttaca 120
 tatacataca taattatata taatatatat ttaaatacac atatacaatt atataataa 180
 tatatttaca tatgcatata taattataga taatatatat ttacatatatac atataaatt 240
 atataataa tataatgtff acatatacat atataattat atataatata tatttaaata 300
 tacatatata attatatata atatatatff acatatgcat atataattat agataatata 360
 tattttacata tacatatata attatatata atataataa ttacatatata catatatata 420
 gtatatataa tatataatat ttacatatatac atataaatt tatatatata atatatata 480
 tatatta 487

SEL PCT 012.ST25

<210> 202

<211> 421

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(421)

<223> MAR of chromosome 2 genomic contig; 16562490..16562910

<400> 202

tatatgaata tatatatgaa tatatacgta tatatgaata tatacatgta tgtatatatg 60

aatatatgta tatatatgaa tatatatgta tatatgaata tatgtatata tatgaatata 120

tatgtatata tgtatatata tgaatatata tgtatatatg tatatatatg aatatatatg 180

tatatatgta tatatgtata tatatgaata tatatgtata tatgaatata tatgaatata 240

tatgtatata tatgaatata tatgaatata tgtgtatata tatgaatata tatgtatata 300

tatgaatata tgtatatata tatgaatata tatgtatata tgtatatatg aatatatatg 360

tgtatatgaa tatatatatg aatatatatg tgtatatgaa tatatatgaa tatatatgtg 420

t 421

<210> 203

<211> 479

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(479)

<223> MAR of chromosome 2 genomic contig; 21592301..21592779

<400> 203

tatatgtata cgtatataat atattatata ttatatacgt gtacgtatat atgtaatata 60

taatgtatat gtacacgtat ataatatata atatattata tacgtatacg tatacattat 120

atattacata tatacgtata tacgtatata aatatatgt atatattata tatacgtata 180

taatatatat tatataatat ataatatata cgtatacata taatatatta tatatacata 240

ttatatatta tatatttaaa ttatatatta tatcatatat aatatatatg atataatata 300

taatatacat atattacata atatatatta tatacatata catatataat atataatata 360

ttatatacat atacatatat aatatataat atattatata catatacata tataatatat 420

aatatattat atacatatat atatataata tataatatat tatatataca tattatata 479

<210> 204

<211> 870

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(870)

<223> MAR of chromosome 2 genomic contig; 22557584..22558453

<400> 204

tataatatat aatatata atagtatat ttatacaca atataaataa tatacataac 60

SEL PCT 012.ST25

atatatgtat attttatata tgtatatatt atatatatt tatatatatt atatatatgt 120
 atattttata tataatatat atattgtata taataatata taatatatta tattatatat 180
 aatatatata atatatatat aaatatatat tatatataat atgtataata tataatattt 240
 tatatataat atgtataata tatattttat atataataat atgtacaata tatattttat 300
 atataataat atgtacaata tatattttat atataataat atgtacaata tatattttat 360
 gtataaatg tataatatat attttatgta taatatatat ttatgtata atatatattt 420
 tacgtatatt ttatatataa tatataatat ttatatata atataaaca ttatatatat 480
 aatatataat attatatata ttatatattt tatatataat atataaaat atatatattt 540
 tatatataat atattttata tataatatat ataaatatat atattatata taatatattt 600
 tatatataat atattttata tataatatat aatatattt atattatata tataatatat 660
 tatatattat atataatata ttatatataa tatataatat ataatatatt atataataa 720
 tataatatat aatatattat atataatata taatatataa tatataatat attatatata 780
 atataataa tgtaatatat aatatattt atataatata taatataata tataatattt 840
 tatatataat atataatata taatatataa 870

<210> 205

<211> 1086

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1086)

<223> MAR of chromosome 2 genomic contig; 30591960..30593045

<400> 205

SEL PCT 012.ST25

gtatatataa tatatattat attatgttat atattatgta gactatgtat taaatatatg 60
 tatatattat atataaatat ataatatata ttataaattt ataattataa atatatttat 120
 aatatatttt tctaaatatt tatatattat atattatatac taatgatata taataaatat 180
 atttctaata tttttatat ttataaatat ttatatata ttatatattt tatatatatact 240
 atatattata tattatatat ttatatata ctatatatta tatagtatat attttatata 300
 tactatatat tatatattat atattttata tatactatat attatatatt atatatttta 360
 tatatactat atactatttta ttatatattt tatatatatact atatactatt tattatatat 420
 ttatatata ctatafacta ttattatat attttatata tactatatat tatatattat 480
 atattttata tataatatat atttattata tttttatat attatatata ttatatatta 540
 tatatttata tattatataa tatatattat atatagaata tataatatat attatatata 600
 atataatata atatataffa tataaaatat atataatata taaaatatat aatatatgat 660
 atatataata tatattctat atttatacat atatatttaa tattatatta atatataatt 720
 atatattatc atagttaata atagatatata tatgtaatat ataaattata attatatatt 780
 aatattatat attatttaat atgtatattt acacatatat taattattaa atatatatat 840
 ttaatatatt aaatattatg tattaaatat atataatata ttataaata ttttatatat 900
 aatatataca tatattaaca tatatgtata tatgtatata ttatatataa cattatatat 960
 attatgttac atatactata ttttatatgt tacatatatact atatattata tgttacatat 1020
 aatatatata acatatatta taatatgttaa catattatat ataacatata atatatagta 1080
 tatata 1086

<210> 206

<211> 406

<212> DNA

<213> Homo sapiens

<220>

SEL PCT 012.ST25

<221> misc_binding

<222> (1)..(406)

<223> MAR of chromosome 2 genomic contig; 36233909..36234314

<400> 206

attataaata tatattatag atattagata ttatagatat aatatatata atatatatta 60

tagatattat agatatagat ataatagata ttatagatat tatagatata atatatatta 120

tagatattat agatataata tatattatag atattataga tataatatat attatagata 180

ttatagatat aatatatatt atagatataa tatatattat agatattata gatatagata 240

ttatagatat tatatatatt atagatataa tatatattat agatattata gatatagata 300

ttatagatat aatatatatt atagatatta tagatataat atatatata gatattatag 360

atataatata tattatagat ataagatata ttatagatat tacaga 406

<210> 207

<211> 797

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(797)

<223> MAR of chromosome 2 genomic contig; 36271745..36272541

<400> 207

atataaacat atacgtatat acacatatat acaaatacat atatacatat attatatata 60

tgtatatata ttatattata catatattat atatatatta tattatacat atatacatat 120

SEL PCT 012.ST25

acacataaac atattacata catatacaaa ttatacacat atacatatat acatatatgt 180
 atatacatatc attatatata aatatatgta tataaaatgt acattatata tacatatata 240
 ttatgtataa ataatatata aaataaacat aatatatatt tatagatatg atatatataa 300
 tatatatgta tacatatata catatatgta tatataatgt acattatata tacataaaca 360
 tcatatataa atgttatata tataatataa atatatataa tatataatat atactttata 420
 tactatatat aatatatata atagtatata acatatataa tatataactat atataatata 480
 tactatatat actgtatata atatatataa taatatatac tatatatatac aaatatataa 540
 tacataatat aatatatac atatatataa tataatatat aatatagtat atatactata 600
 tataataatt acatattata tattatatac tatatatatt ataattatta tatataatta 660
 tatattacat acttgtata taatgtaaat atacattaga atatatatg tatatatatg 720
 tacatatata atgtatatat gtatacatata tataaactat atataaacat tatattatat 780
 aaacattata tataaac 797

<210> 208

<211> 423

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(423)

<223> MAR of chromosome 2 genomic contig; 36498521..36498943

<400> 208

tattatatta tatatttaatt attatatatt taatatatta tatatttaatt attatatatt 60

taatatatta tatatttaatt attatatatt taatatatta tatatttaatt attatatata 120

SEL PCT 012.ST25

taatatatta tatatttaaat attatatata taatattata tatataatat tatatattta 180
 atattatata tataatatta tatatataat atattatata tttagtatta tgtatttaaat 240
 atattatata tttagtatta tgtatttaaat atattattta tttagtatta tatatttaaat 300
 atattattta tttagtatta tatatttaaat atattatata ttaatatat tatatattta 360
 ttatatattg tatatttaaat atattatata ttattatat attatatata attatatatt 420
 taa 423

<210> 209

<211> 304

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(304)

<223> MAR of chromosome 2 genomic contig; 37179891..37180194

<400> 209

gtgtatatat atcatatata ttatatcata tataatgtgta tatatatcat atattatattc 60
 atatatatgt gtatatatat catatatata tcatatatgt gtatatatca tatatattat 120
 atatcatata tgtgtatata tatcatatat tatatatcat atatatgtgt atatatcata 180
 tatattatat atatctcata tgtgtatata tatcatatat aatatatatg tgtatatatc 240
 atatatcata tataacatat atagtgtgt atatatcata tataacatat atcatatatg 300
 tgta 304

<210> 210

<211> 693

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(693)

<223> MAR of chromosome 2 genomic contig; 38440448..38441140

<400> 210

tatatattct ttatatatt atatataata tatattcttt tatatattat atatagtata 60

tattctttta tatattatat atagtatata ttcttttata tattatatat agtatatatt 120

cttttatata ttatatatag tatatatatt ttatatatt atatatagta tatattcttt 180

tatatattat atatagtata tattctttta tatattatat atataatata tattctttta 240

tatatcatat ataatatata ttcttttata tattatatat aatatatatt cttttatata 300

ttatatatca tgtatatata atatacaaaa tatatataga ttatatatat agattattac 360

ataatagaat atattatata ttatatataa tatatacata atatataata ttatatatga 420

tataatatat atcatatata tcatataata tatattatat atcatatatt atatataata 480

atatatagat tatatataat tatatatata atatatataa ttatatatat tatctatata 540

tagataatat atataattat atataatata ttatatagat tatatataat tatattatat 600

acaaaatcta tatataatat atattatatt atatataata tacataacta tataaaaaat 660

ataatatata atatataata tatataatat ata 693

<210> 211

<211> 471

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(471)

<223> MAR of chromosome 2 genomic contig; 38887582..38888052

<400> 211

aacatatata ctatatatat tatatactat attatatatt atatatataa acatatatac 60

tatatataat atataaacat attatattat acatgatata gataaacata tatattatat 120

ataatataga taaaatatgt tatatataat ataattgata gacatatatt atatatatat 180

atattctaca tatattatat atatttcta cacatattat attatatata catatattct 240

acatatatta tatatacata tattctacat atattatata tacatatatt ctacatatac 300

atatatacat atattatata tacatatatt atagatatat aatatataaa catatataat 360

attattatat ataatatata taataatatt atataatata taataatatt atattttata 420

tataaataat atatatattt tatatatata atattatata tatataatat a 471

<210> 212

<211> 1221

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1221)

<223> MAR of chromosome 2 genomic contig; 43885944..43887164

SEL PCT 012.ST25

<400> 212

catataaaca tatattatat gtaacatata aacatattat atgtaacata taatatataa 60
 tatataaaca tataatttat atattatatg ttacatataa tatataatat ataaacatat 120
 attatatatt atagttaaca tataatatat aatatataaa catatatatt atataataa 180
 tataaacata ttttatatat aatatataaa catattttat atataatata taaacatata 240
 ttttatatat aatatataaa catattttat atataatata taaacatata ttttatataa 300
 tatataaaca tataatatat ataatatata aaagtatata atataaatat atataatata 360
 aacatatata atataaatat atataaaata taaacatatg taatatataa acatatatta 420
 tatataatat ataaacatat attatacgta caatatataa acatatattg tacgtacaat 480
 atataaacat atattatacg tacaatatat aaacatatat tatacgtaca atataaaac 540
 atatatataa cgtacaatat ataaacatat attatacgta caatatataa acatatatta 600
 tacgtacaat atataaacat atattatacg tacaatatat aaacatatat tatacgtaca 660
 atataaaac atatatataa cgtacaatat ataaacatat attatacgta caataaacat 720
 atattatacg tacaatatat aaacatatat tatacgtaca atataaaac atatatata 780
 cgtacaatat ataaacatat attgtacgta caatatataa acatatatta tatgtataat 840
 atataaacat ataatatata atatatatta tatatatgtt tattatatat gtttatatat 900
 tatataaac atatatatt atattatata tgtttatata ttatatatta tataatatat 960
 atgtttatat attatatatt atataatata tatgtttata tattatatat tatataatat 1020
 atagtttat atattatata ttatatataa tatatgttta tatattatat attatataat 1080
 atatatgttt atatatataa tattatataa tatatatgtt tatatatatt atattatata 1140
 atatatatgt ttatatatta tatattatat aatatatatg tttatatatt atataaataa 1200
 taaacttaca tttttatta a 1221

<210> 213

<211> 543

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(543)

<223> MAR of chromosome 2 genomic contig; 45818200..45818742

<400> 213

tatgtatata tacatatata ttatacatg tatatatgta tatatacata tatattata	60
catgtatata tatacatata tttatataca tgtatatata tacatatata ttatacatg	120
tatatatata catatatatt tatacatgta tgtatatata catatatatt tatacatgta	180
tgtatatata catatatatt tatacatgta tgtatatata catatatatt tatacatgta	240
tgtatatata catatatatt tatacatgta tgtatatata catatatatt tatacatgta	300
tgtatatata catatatatt tatacatgta tgtatatata catatatatt tatacatgta	360
tgtatatata catatatatt tatacatgta tgtatatata catatatatt tatacatgta	420
tgtatatata catgtatatt tatacatgta tgtatatata catgtatatt tatacatgta	480
tgtatatata catgtatatt tatacatgta tgtatatata catgtatatt tatacatgta	540
tac	543

<210> 214

<211> 463

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(463)

<223> MAR of chromosome 2 genomic contig; 47055478..47055940

<400> 214

atacatacat atatacatat atacacatat atacatatataa tacacacata ttatcatata 60
 tacacacata tatacatata tacatatata cacatatata catgcataca catatatata 120
 tatatacaca catatacaca catatatata tatatacaca tatatacaca tatacacata 180
 tatacacaca tatacatata tacacatata tacatatata catatatata cacatatata 240
 catatatata tatacacata tatacacata tacatatata cacatatata cacatatata 300
 catatatata catatatata tatatacaca tatatacaca catatacaca tatatacata 360
 tatacatatg tatacacata tatacatatg tatacacata tatacacata tacatatata 420
 catacacata tatacgtata tatgtgtata tatacacata tac 463

<210> 215

<211> 2482

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(2482)

<223> MAR of chromosome 2 genomic contig; 47492696..47495177

<400> 215

aatatatata aaatatatta tattctatgt aatatataga atatatataaa tatattctat 60
 atattatata gaatatatat ttataatat atattattta tatattttta tatatttata 120

SEL PCT 012.ST25

ttatttatat atttatatat aatttatata atttatatat ataatttata tataatttat 180
 ataaattata tatataattt atatataatt tatatataat ttatataaat tatatatata 240
 atttatatat aatttatatg atttttatat ataatttata tataatttat ataattttta 300
 tatataattt atatataatt tatataattt ttatatataa ttatataat atatatatat 360
 aatttatata taatttatat aatttatata tataatttat atataattta tataatttat 420
 atatataatt tatatataat ttatataatt tatatatata atttatatat aatttatata 480
 atttatatat ataatttata cataatttat ataatttata tatataattt atataattta 540
 tatatataat ttatatatat aatttatata atttatatat atgatttata taatttatat 600
 atataattta tataatttat atatataaat tatatatata atttttatat aatttatata 660
 ttataattt atatatttat ataatttata tatttataat ttatatattt atataattta 720
 tatatttata atttatatat ttatataatt tatatataat tattcatata ttatataat 780
 ttacatatata ttatttatat attcatatat aatttatata ttatatata atttatatat 840
 aatttttac atatttatat atttatatat aatttatata tatttatata taatttataa 900
 ataaaatata taatatataa tatataatat tataatagat aaaatatata ctatatatta 960
 tatattttac atttatattta atatttatatg tataattttta tatcatatat aatatatatg 1020
 atatatatataa ttttatatca tatataatat ataggtata tataattttta tatcatatat 1080
 aatatatatg gtatatataa ttttatatca tatataatat atgatatata attttatatc 1140
 atataatata tatttatatat aatttttatat ctacatatata tatatttatat atacaatttt 1200
 atatctatct ataatatata ttatatataac aatttttatat ctatataata tatatttatat 1260
 atacttttat atttatatata aaatgtatat tatatatact ttatatattat atataaaatg 1320
 tatatttatat ataattttat ttatatata aaatgtatat tatatataat ttatttttat 1380
 atataaaatg tatatttatat ataattttat ttatatata aaatgtatat tatatataat 1440
 ttatttttat atataaaatg tatatttatat ataattttat ttatatataa aaatgtatat 1500
 tatatataat ttatatattat atataaatg tatatttatat ataattttat atttatatata 1560
 ataggtatat tatatataat ttatatattat atataaatg tatatttatat ataattttat 1620

SEL PCT 012.ST25

attatatata atatgtatat tatatatataat ttgtattat atataaatg tatattatat 1680
 ataattttat attatatata atatgtatat tatatatataat ttatattat atataaatg 1740
 tatattatat ataattttat attatatata atatgtatat tatatatataat ttatattat 1800
 atataaatg tatattatat ataattttat attatatata aaatgtatat tatatatataat 1860
 ttatattat atataaatg tatattatat ataattttat attatatata atatgtatat 1920
 tatatatataat ttatattat atataaaatg tatattatat ataattttat attatatata 1980
 aaatgtatat tatatatatt atatataaaa tgtatattat atatattata tataaaatgt 2040
 atattatata tattatatat aaaatgtata ttatatatat tatatatataaa atgtatatta 2100
 tgtatattat atataatgta tattatgtat attatatata atgtatatta tatataatat 2160
 atattatata taatgtatat tatataatat atattatata ttataatata taatatacat 2220
 tatatattac atattatata taatatatta tatattatat attacatatt atatataata 2280
 tattatatat tatattaaat atatatttta tatattatat attatatatt atataaaaata 2340
 tatatattat atattatata aaatatatat atattatatt atatattata ttaaatatat 2400
 attttatata taatatatat aatatataat atataaaaata tatattatat attatatata 2460
 aattatatat attatatata aa 2482

<210> 216

<211> 539

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(539)

<223> MAR of chromosome 2 genomic contig; 47561069..47561607

SEL PCT 012.ST25

<400> 216

aacagtaata taccactaat atataataat atataacagt aatatatcat taatatataa 60

tatatcatta gtatataata ttaatatata ttaatatata atatatcata tacaatatta 120

atatatatta atatataata atatattatt aatgtataat agtaatatata tatattatca 180

atatatatta ctaatatata ataatatatc gtaatatat aatagatcat taatatataa 240

tgtaatatata ttatgaatag ataatatatc agtatataat attaatatat taatatatta 300

tatattattt aataatatat aatatattaa taaataatta tatattaata tagcaatata 360

ttaatatatg actgtattat attattaata tataacaata tattatatat tatataataa 420

tttattatat aatatataat aatatattat atattatata acatattaat aatacataat 480

aacattaata atatataata atgtaatat attattatat tatatattaa tatataata 539

<210> 217

<211> 336

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(336)

<223> MAR of chromosome 2 genomic contig; 52853648..52853983

<400> 217

tatatacata aaatatatat attttatata tatacataat atatatatgt atattttatg 60

tatatatcta taatatatat aatataataa aatatacata tatattttat atatataata 120

tatacatata aaatatacat acataaaata tacatgtata tttatgtat atataatata 180

tatataaaat atacatgtat attttatata tataatatatc atgtataatt aatatacatg 240

SEL PCT 012.ST25

tatgttatat atattacatg tatattatat ataatatata tataaatftt aaatttagtg 300

tatattacat gtatattata tataatatat gtatat 336

<210> 218

<211> 406

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(406)

<223> MAR of chromosome 2 genomic contig; 54866263..54866668

<400> 218

tacgtatata aaaatgtata ttacatata taaaataaat attttatata cgtatataaa 60

atatatattt attttatata cgtatataaa atatttattt tatatatgta tataaaatat 120

ttattttata tacatgtata ttaaatatat atttatatat gtatatataaa atatataatta 180

tatacatgta tataaaatat atatttatata tgtatatataa aatatatatg tatataaaat 240

atatataatta tatagatata taaaatatat atttatataga tatataaaat atatataatta 300

tatagatata taaaatatat atatttatata gatatatataa atatatatat tatatagata 360

tataaaatat atatattata tagatatata aaatatatat attata 406

<210> 219

<211> 1452

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(1452)

<223> MAR of chromosome 2 genomic contig; 55113305..55114756

<400> 219

ataatatata atatatttg tatattatat tattatatat tatatattat taaatatata 60
 tattatatta tatattatat aatatatatt atatataata atatagaata tataattata 120
 tattatatta tatattatat aatatatatt atatataata atatagaata tataattata 180
 tattatacta tatattatat aatatatatt atatataata atatagaata tataattata 240
 tattatataa tatgtgaata atgtaatata taattatatt atttaccat tatataatat 300
 ataattatat tatataatat ataattatat tatttgata ttatatataa catatacatt 360
 atattatata taatataatt atatataatt aattataaat taattatata taattatata 420
 atataatata taatatacat aatatataat atataatata taatatacat aatataatat 480
 atattatata taatataata tatataatat aatataatat aatgtataat ataattatat 540
 attatatata atatataatg ttatatataatt atattatatt atataattaa ttatatgtaa 600
 ttaatataat ataattatta tatataattt ttatatataat ataatatata attatataat 660
 ataatatat tatattatat tatataatat atatattatta tataatataa tataattata 720
 ttatatataat atataatata atataattat atattatatt atataataaa tataattata 780
 taatataata tgattatata atatattatg tatattatat attatatatt gtattatgta 840
 tattatatat tatatattat gtatattata tattatgtat attatatatt atgtatatta 900
 tatattatat attatattat gtataatata ttatgtatgt tatatataat ataaattata 960
 ttatatatta tgtatattat atataaatta tattatatat tatgtatatt atatataata 1020
 taaagtatat attatgtata ttatatataa tataaagtat atattatgta tattatatat 1080
 aatataaagt atatattatg tatattatat ataataataa gtatatatta tgtatattat 1140

SEL PCT 012.ST25

atataatata aagtatatat tatgtatatt atatataata taaagtatat attatgtata 1200
 ttatatataa tataaagtat atattatgta tattatatat aatataaagt atatattatg 1260
 tatattatat ataataataa gtatatatta tgtatattat atataatata aagtatatat 1320
 tatgtatatt atatataata taaagtatat attatgtata ttatatataa tataaagtat 1380
 atattatgta tattatatat aatataaagt atatattata tggataaat tatatattgt 1440
 tatatattat at 1452

<210> 220

<211> 502

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(502)

<223> MAR of chromosome 2 genomic contig; 56350637..56351138

<400> 220

atatattata gaaatataaa tatatagata tatctatata ttatagaaat ataaatatat 60
 agatatatct atatattata gaaatataaa tatatagata tatctatata ttatagaaat 120
 ataaatatat agatatacct atatattata gaaatataaa tatatagata tacctatata 180
 ttatagaaat ataaatatat agatatacct atatattata gaaatataaa tatatagata 240
 tatctatata ttatagaaat ataaatatat agatatatct atatattata gaaatataaa 300
 tatatagata tatctatata ttatagaaat ataaatatat agatatatct atatattata 360
 gaaatataaa tatatagata tatctatata ttatagaaat ataaatatat agatatatac 420
 aacatatatg ttacatatta tatattatat atctatatat ctatataaca ttatatatct 480

SEL PCT 012.ST25
502

atatatctat ataacatata ta

<210> 221

<211> 794

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(794)

<223> MAR of chromosome 2 genomic contig; 57051633..57052426

<400> 221

aactatatat actatattat atagttatac tatatatact atattatata gttatataac 60
tattatataa ctgtattata tagttatata actattatat aactgtatta tatagttata 120
taactattat ataactgtat tatatagtta tataactata ttatataact gtgttatata 180
gttatatatt atataactat attatataac tgtattatat agttatatat tatataacta 240
tattatataa ctgtattata tagttatata ttatataact atattatata actgtattat 300
atagttatat attatataac tgtattatat agttataaaa ctatattata taactgtatt 360
atatagttat aaaactacta tataactgta ttatataatt ataaaattat actatataac 420
tgtattatat agttataaaa ctatactata taactgtatt atatagttat aaaactatac 480
tatataactg tattatatag ttataaagct atactatata actgtattat atagttatat 540
aactatacta tataactgta ttatagtt ataaaactat actatataac tgtattatat 600
agttataaaa ttatattata taactgtatt atatagttat ataactatat tatataactg 660
tattatatag ttatataact atattatata agtgtattat atagttatat aactatatta 720
tataactgta ttatacagtt atataactat attatataac tgtattatat acttatataa 780

SEL PCT 012.ST25
794

ctatattata taac

<210> 222

<211> 300

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(300)

<223> MAR of chromosome 2 genomic contig; 57069272..57069571

<400> 222

acacatacat atatgtatat atgcacacac atatatatgt atatatacac atacatatat 60

gtatatatac atatatgtat atacgcacat acatatatgt atatatacac gtacatatat 120

gtctctatat atacacatac acatatgtat atacatatat gtgtatatat acacaatcat 180

atatgtatat acatatatac acatatacac aaacatatat gtatatatac atatgtatat 240

acatatatac acatatacac aaacatatat gtatatatac atatgtatat acatacacia 300

<210> 223

<211> 370

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(370)

SEL PCT 012.ST25

<223> MAR of chromosome 2 genomic contig; 57235143..57235512

<400> 223

```
tattttata tataactata tatattttat atataaatta tatatatgat catatatata    60
atcatatata taatcatata tgattatata tgatcatata tatatttata tatataatta   120
tatataactta tatataaatta tatatatatt tatatatata attatgtata cttatatata   180
tttatatata taattatata tacaatttat atataaatt atataaatt tatatataat   240
tatatatata aattatatat aagtatatat aattatatat atgtttatat ataattatat   300
atataaatga tatgtataat atataactat atataattat atataaatat atatatagat   360
tttatatata                                     370
```

<210> 224

<211> 306

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(306)

<223> MAR of chromosome 2 genomic contig; 57693125..57693430

<400> 224

```
tacgtatata cacgtataaa tataaatata tacatgtata tacgtatata catgtataaa    60
tataaatata tatatgtata tacgtatata catgtataaa tatatatatg tatatacgta   120
tatacatgta taaatatata tatatgtata tacgtatata catgtataaa tatatataca   180
tgtatatacg tatgttgtgt atacatacaa atctgtacat atatacatat atgttgtgtg   240
```

SEL PCT 012.ST25

tatatataca tctatacatg tgtatgcgta tatatgtata tgtatatata gtatatataa 300

tacatg 306

<210> 225

<211> 500

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(500)

<223> MAR of chromosome 2 genomic contig; 59810331..59810830

<400> 225

tttattatat gtaatatata ttgtattatt atatatatta tatataatat atattgtatt 60

attatatata ttatatataa tatatttgtt attattatat atattatata taatatatat 120

tgtattatta tatatattat atataatata tattgtatat tatatatatt atatattata 180

ttattatata ttatatatat tatattatta tatattatat attatatata ttattattata 240

tattatatat tatattatat atattatatt atatattata tattatatta tatatattat 300

attatatatt atatattata ttatatatat tatattatat atattatata ttatatatta 360

tatatattat atattatata ttatatatat tatatattat atataatata tattatatta 420

ttatataata ttatatatta tatatattat atattatata taatatatat tatattatta 480

tataatatta tatattatat 500

<210> 226

<211> 565

<212> DNA

SEL PCT 012.ST25

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(565)

<223> MAR of chromosome 2 genomic contig; 59974589..59975153

<400> 226

atatatgtat aatatgtata tatgtatata ttatgtatat gttatatatg taatatatgt 60
 atgtatatat tatatatcat atataatata taatgtgtat atatgtatat atgtatgtat 120
 acatgtatat actatgtata tattgtatat attatatatg tatatataca tatacatata 180
 taatatatac atatattata tacaatatat acatgtatat tatatacgat atatacatat 240
 atattatata caatatatac atagtatata aatgtataca tacatacata tatacatatt 300
 atatatgtat atatgtatac ataaatgtat atataatata tatacatata taaatgtata 360
 catacgtaca tatacgtata tgtatatgca tatatgtata tatgtgcata catatatatg 420
 tatatacata tatgtacata tgtacatata cgtatatatg tacatatgta catatacgta 480
 tatatgtaca tatgtacata tacgtatata tgtacatatg tacatatatcg tatatatgta 540
 catatgtaca tatatacata tatat 565

<210> 227

<211> 427

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(427)

<223> MAR of chromosome 2 genomic contig; 60605573..60605999

<400> 227

tatataatgt atataatgga tatagatata gatatagata tatatattat ataatatata 60

ttatatatta tatataatat atggtatata tattatatat ttatatatat atatata 120

tataaattat atatataata tatataatat atatattata tatatattat ataatatata 180

ttaatatatta tctattatat atttatata atatataatt tatataatat ataatatata 240

atatataatt tacataatat ataatatata atacgtatta tatataatat ataatacgta 300

tttatataa tatataatac gtattatata taatacgta tatatattat ataatatata 360

atacgtaata tataatatac gtaattatat ttattataa taagtattat atattatata 420

atatata 427

<210> 228

<211> 1199

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1199)

<223> MAR of chromosome 2 genomic contig; 61229949..61231147

<400> 228

gtatacatat ataaagtga tatataatgt atatacatat atacatatat aaagtatata 60

tataatatat acatatataa agtatatata taatatatac atatatataag tatatataat 120

SEL PCT 012.ST25

atatacatat ataaagtata tataatatat acatatataa agtatatata tcatatatac 180
 atatataaag tatatatata atatatacat atatacatat ataaagtata tataacatat 240
 atacatatat aaagtatata taacatatat acatatataa agtatatata taatatatac 300
 atatatacat atataaagta tatataacat atatacatat atacagtata tataacatat 360
 atacatatat acagtatata taacatatat acatatatac agtatatata acatatatac 420
 atatatacag tatatataac atatatacat atatacatga agtatatata acatatatac 480
 atatatacat gaagtatata taacatatat acatatatac atgaagtata tataacatat 540
 atacatatat acatgaagta tatataacat atatacatat atacatatat aaagtatata 600
 taacatatac atatatacat atataaagta taacatatac atatatacat atataaagta 660
 tatataatat ataacatata catatataaa gtatatataa tatataacat atacatatat 720
 aaagtatata taatatataa catatacata tataaagtat atataatata tacatatata 780
 catatataaa gtatatataa tatatatata catatataaa gtatatataa tatatatata 840
 tatatacata tataaagtat atataatata tatacatata taaagtatat ataatatata 900
 tacatatata catatataaa gtatatataa tatatatata tatatacata tataaagtat 960
 atataatata tatacatata tacatatata aagtatatat aatatatata catatatata 1020
 tatataaagt atatataata tatatacata tatacatata taaagtatat ataatatata 1080
 tacatatata catatataaa gtatatataa tatgtatata tatatacata tataaagtat 1140
 atataatatg tatacatata tacatatata aagtatatat ataatatgta tacatatat 1199

<210> 229

<211> 454

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(454)

<223> MAR of chromosome 2 genomic contig; 62181058..62181511

<400> 229

tatatatcat atattatata tgatatatat tatgtatata atacatatta tatataataa	60
atatttatta tatatgatat atattatgta tataatacat attatatata ataaatatat	120
attatattat atataataaa tatatattat attatatata atatatattt atatataaat	180
atattatata taaatatata ttatatataa aatatttata tattatatat aaatatatat	240
tatatataaa ttttatata ttatatataa atatttatat attatatata aatatttata	300
tattatatat aaaatatatt atatatatta tatatattat atatttatata taatatattt	360
aatatataat atataaacat atatttatata taatatataa acatatataa atatatttat	420
atataataga taaaaatata tataatatat ataa	454

<210> 230

<211> 658

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(658)

<223> MAR of chromosome 2 genomic contig; 62190919..62191576

<400> 230

tatatacaca actatatata taactatata tataacaacta tatatacaac tatatatata	60
actatatata taactatata taactatata tataactata tataactata tatataacta	120

SEL PCT 012.ST25

tatatataac tatataaact tatatatata actatatata actatatata actatatata 180
 taactatata tataaactata tatataaacta tatatataac tatatataact atatataata 240
 ctatatatat ataactatat atataactat atatataata ctatatataa ctatatatat 300
 ataactatat atataaactat atatataata ctatatatat aactatatat atataactat 360
 atatataact atatatatat aactatatat aactatatat atataactat atatataact 420
 atatatatat aactatatat ataactatat atatataact atatataata ctatatatat 480
 ataactatat atataaactat atatataact atatataata ctatatatat ataactatat 540
 atataactat atatataata ctatatatat aactatatat ataactatat atataactat 600
 atatataata ctatatatat aactatatat atataactat atatataact atatatat 658

<210> 231

<211> 1486

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(1486)

<223> MAR of chromosome 2 genomic contig; 62384127..62385612

<400> 231

attatatcta atctattata tattatatct aatacatatt atatctaact tattgtatat 60
 tatatctaact atataatata ttatatataa tatattatat attatatatt atatacaata 120
 tattatatat tatataatata ataatatatt atatataata tattatatct aatatattac 180
 atattatact taatctatta tagatataat atgtaatata ttatatatta tatctaataag 240
 atattagata taatatataa tatattatta atataatata ttagatataa tatataatat 300

SEL PCT 012.ST25

aataatatat aatatatatt attggaata tataatatat aattaataat atatattata	360
tataattatt atgaataata tatcatatat aatatctagt atattatata ttaataacat	420
ataaatatta tattaataat aaataacata ttaatattat attaataata tataatatac	480
taatattata ttaataatat ataatafact aatattatat taataatata taatatacta	540
atattatatt aataatatat aatatactaa tattatatta ataatatata atatactaata	600
attatattaa taatatataa tatactaata tattaagaat atataatata ctaatatatt	660
aagaatatat aatatactaa tattatatta ataatatata ttatattaa taatatatta	720
attattatta attaattatt aataattata taatattgat tatattaata ttatcaattt	780
aataatattg attatatatt atatattata tattatatat tatatattat atattatata	840
ttatatatta ataatatata ttagatataa tataatatat taataatata taagatataa	900
tataatatat taataatata tattagatat aatataatat attaataata tatattagat	960
ataatatat atattaataa tatatattag atataatata atatattaat aatatatatt	1020
agatgtaata taatatatta ataatatata ttagatgtaa tataatatat taataatata	1080
tattagatgt aatataatat attaataata tatattagat gtaatatat atattaataa	1140
tatatattag atgtaatatata atatattaat aatatatatt agatgtaata taatatatta	1200
ataatatata ttagatgtaa tataatatat taatatatat tagatgtaat ataatatatt	1260
aataatatat attagatata atataatata ttaataatat attagatata atataatata	1320
ttaataatat ataagatata atataatata ttaataatat ataagatata atataatata	1380
ttaataatat ataagatata atataatata ttaataatat atattagata tataatatat	1440
taataatata tattagatat ctaatatcta ttagatatct aataga	1486

<210> 232

<211> 333

<212> DNA

<213> Homo sapiens

SEL PCT 012.ST25

<220>

<221> misc_binding

<222> (1)..(333)

<223> MAR of chromosome 2 genomic contig; 62538649..62538981

<400> 232

ttatatatat tatatatata tttttatat atatattata tatatatatt atatatatat 60

tatatatata ttttatatat atattatata tatattttat atatatatatta tatatatatt 120

ttatatatat tatatatata ttttatatat attatatata tttttatat atatattata 180

tatatatatt atatatatat tatatatata ttttatatat atattatata tatattttat 240

atatatatatta tatatatatt ttatatatat attatatata tttttatat atatattata 300

tatatatatt atatatatat tatatatata ttt 333

<210> 233

<211> 480

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(480)

<223> MAR of chromosome 2 genomic contig; 63240325..63240804

<400> 233

tatatataaa atatatatatt tttaaataata aaatatatat atattttaat attaatatat 60

atatatttta atatataata tatatatatt atattttata tataaaatat atatattata 120

SEL PCT 012.ST25

tattttatat ataaaatata tatattatat attttatata ttaaaatata tattttatat 180

attttaatta ttaaaatata tatattatat attttaaata taaaatatat atattatata 240

ttttaatata taaaatatat atattatata ttttaatata taaaatatat atattttata 300

tttatatata taaatata tattatatat ttttaatatat aaaatata tattatatat 360

tttaatatat aaaatata tattatatat ttttaatatat aaaatata tattatatat 420

tttaatatat ataaaatata tatattatat attttatata tattaaatat atattttata 480

<210> 234

<211> 302

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(302)

<223> MAR of chromosome 2 genomic contig; 63935480..63935781

<400> 234

atatatataa atatatatat aattatatat agatatatat aattatatat agatatatat 60

attctatatt ctatatatat ataatatata atatataaat tatatataga atatatatta 120

tatataatat attatatata ttatatataa tatatatatt atatatatta tatataatat 180

atatattata tatattatat ataatttata tatattatat atagaatata tattatatat 240

agaatataga atatatataa tatatataga atacagaata tatatagaat atagaatata 300

ta 302

<210> 235

<211> 407

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(407)

<223> MAR of chromosome 2 genomic contig; 63935888..63936294

<400> 235

tataatatat taatataata tatagacagt atataatata atatacagac agtatataat 60

atacagacag tatataatat ataattatt atataatatt atatataata ttatataata 120

tattatatta tatattatt ataattatt atattatata taatatatgt aattattat 180

attattatt acataatata ttatatataa tatattatat ataattatt atatattata 240

tataatatat ataataataa tattataata tataatatat aatagtacag tatattatt 300

atatataatt ctatatataa tatatagaat tctatctatt tataatatat atagaattct 360

atatataata tataatatac agaattctat atatattata tatagaa 407

<210> 236

<211> 302

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(302)

<223> MAR of chromosome 2 genomic contig; 66958350..66958651

SEL PCT 012.ST25

<400> 236

tattatatat attgtatata tatgtatatt atatatattg tatatataat gtatattata 60

tatattatat atatatgtat attatatata ttgtatatat atgtatatta tatatattgt 120

atatagtat atgtatatat gtatgtgtat atatatcac atatacacat atatgtgtat 180

gtatatatat gtgtgtatat acgtatatat acatatatac aatttttgta tatatacata 240

tatacacata tatatgtgta tgtgtatata tatacacata tatgtgtgtg tatatacaca 300

ta 302

<210> 237

<211> 651

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(651)

<223> MAR of chromosome 2 genomic contig; 68307125..68307775

<400> 237

gatattatat attgtatata ttatatatgt atataatata ctattatata ttatatatgt 60

atataatttt attaatatat atattatatt atattatata ttatattata ttatattata 120

tatataatat taatattata tattattata tattatatta tattaatatt atatatatat 180

aatatatata atatatataa tagtattata tataatatat ataatagtat tatatattat 240

atatatatata tactattata tatattatat ataatagtat tatatatatt atatatataa 300

tactattata tataatatat actattatat aatatatata atactattat atatattata 360

SEL PCT 012.ST25

tataatacta ttatatataa tatatataat actattatat ataatatata taatactatt 420
 atatataata tatataatac tatttatatat aatatatata atactattat atataatata 480
 tataatacta ttatatataa tatatatatt atatataatt atattaatat ataatagtat 540
 catatataat aatagtatat ataatatata atatatatat tatatatatt ataatagtat 600
 atataacata taatatagta tatatatatt atattatata taaaatattt a 651

<210> 238

<211> 367

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(367)

<223> MAR of chromosome 2 genomic contig; 68308243..68308609

<400> 238

atatatatat atgagtcaac catacacata tatatatata atgtttatat atataatgta 60
 tatatataat gtttatatat aatgtatata tataatgttt atatatatata tgtatatata 120
 taatgtttat atatataatg tatatatata atgtttatat atataatgtg tatatataat 180
 gtttatatat ataatgtgta tatataatgt ttatatataa tgtgtatata taatgtttat 240
 atatataatg tgtatatata atgtttatat atataatgtg tatatataat gtttatatat 300
 ataatgtgta tatatataat gtttatatat ataatgtgta tatatataat gtttatatat 360
 ataatgt 367

<210> 239

<211> 499

SEL PCT 012.ST25

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(499)

<223> MAR of chromosome 2 genomic contig; 410241..410739

<400> 239

ataatatgta tatatattat attatatatt atattacata ttatatatta tattacatat 60

tatatattta tatattacat attatatatt atattttata ttatatatta tatcatatat 120

atgttatgca ttatataata cataatatat tatatatgat ataatatata ttatatatta 180

ttatatataa tataattaat atattatgta ttatataata tatattatgt tataatatat 240

aatatatatt atataattat ataatatatt atgtattata taatatatat tatgttataa 300

tatattatat tatatatatt atatatatat tatatatata atgtatatta tatataatac 360

ataatatatt atatattata tattatttta tataatatat tatataatgt gatatattat 420

ataatatatt atataacata gtatattata taatatatta tataatgtaa tatattatat 480

attatataat atattgtat 499

<210> 240

<211> 402

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

SEL PCT 012.ST25

<222> (1)..(402)

<223> MAR of chromosome 2 genomic contig; 31531..31932

<400> 240

cacattatat atataaacat tatatatata cacattatat atataaacat tatatatata 60

cacattatat atataaacat tatatatata cattatatat ataaacatta tatatacaca 120

ttatatatat aaacattata tatacaaatt atatatataa acattatata tacaattat 180

atatataaac attatatata tacattatat atataaacat tatatatata cattatatat 240

ataaacatta tatatatata ttatatatat aaacattata tatatacatt atatatataa 300

acattatata tatacattat atatatataac gttatatata tacattatat atataaacat 360

tatatgtata cattatatat ataaacatta tatatatatg tg 402

<210> 241

<211> 421

<212> DNA

<213> Homo sapiens

<220>

<221> misc_binding

<222> (1)..(421)

<223> MAR of chromosome 2 genomic contig; 32415..32835

<400> 241

ataaatattt tatatatata atataatata tatactatat tatatgttat atatactatt 60

ataatatata taatatatat attatatatt atatatata ttattatata tgatactatt 120

atatattaat ataattatat ataatatata tattatataa tatactatta tatattatat 180

SEL PCT 012.ST25

ataatagtat attatataat atatatatta tatataatag tattatatat actattatat 240

attatatata ttatatatat ataaaatata atataatata tataatatat aatattaata 300

ttatatatat aatataatat aatatataat ataataat atatatatta ataaaattat 360

attaatatat aatatataat agtatattat atacatatat aatatatata atatataata 420

t 421