

# SEQUENCE LISTING

<110> GE HEALTHCARE BIO-SCIENCES AB

<120> METHODS FOR THE SELECTION OF NUCLEIC ACID SEQUENCES

<130> 327740-GB-1

<160> 16

<170> PatentIn version 3.5

<210> 1

<211> 47

<212> DNA

<213> Artificial sequence

<220>

<223> Partly modified native attP sequence

<400> 1

gtgccccaac tggggtaacc tttgagttct ctcagttggg ggcgtag

47

<210> 2

<211> 70

<212> DNA

<213> Artificial sequence

<220>

<223> Partly modified native attB sequence

<400> 2

ctcgaagccg cgggtgcgggt gccagggcgt gcccttgggc tccccgggcg cgtactccac

60

ctcaccatc

70

<210> 3

<211> 47

<212> DNA

<213> Artificial sequence

<220>

<223> Partly modified native attP sequence

<400> 3

gtgccccaac tggggtaacc taagagttct ctcagttggg ggcgtag

47

<210> 4  
<211> 70  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Partly modified native attB sequence

<400> 4  
ctcgaagccg cgggtgcgggt gccagggcgt gcccaagggc tccccgggcg cgtactccac 60  
ctcaccatc 70

<210> 5  
<211> 1941  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Partly modified native PhiC31 gene

<400> 5  
atgaccatga ttaccccatc tgcccagctg accctgacaa agggcaataa gagctgggtct 60  
agcctgggtga cagctgcttc tgtgctggag ttgcccacca tgatccaagg ggtcgctggg 120  
gaagtgactt atgccggggc gtacgaccgt cagtctcggg agcgcgagaa ctctagcgcg 180  
gCGTctccgg ccactcagcg tagcgctaac gaggccaaag ccgccgtctt ccagcgcgag 240  
atcgagcgcg ccggggggccg gtttcgtttc gtcgggtcact tcagcgaggc ccccggcaca 300  
tctgccttcg gtacagccga gcgccctgag ttcgaacgca ttctgaacga atgccgcgcc 360  
ggtcgggtga acatgattat cgtgtatgac gtgtctcgct tctctcgctt gaaggttatg 420  
gacgccatcc ctatcgtgtc agaattactg gccctgggcg tgacaatcgt ctctacgcag 480  
gaaggcgtgt tcagacaagg gaacgttatg gacctgatcc acctgatcat gcggctggac 540  
gcctctcaca aagaaagctc tctgaagtct gccaaagatcc tggacacaaa gaacctccag 600  
cgCGaacttg gcgggttacgt gggcggggaag gccccctacg gcttcgagct tGtcagcgag 660  
acaaaaggaga ttacacgcaa cggacgtatg gtcaatgtgg ttatcaacaa gctcgccccac 720  
tctaccacgc ctctcaccgg acctttcgag ttcgagccag acgtaattcg gtgggtgggtgg 780  
cgtgagatca agacacacaa acacctccct ttcaagcctg gcagtcaagc cgccatccac 840

|                                                                     |      |
|---------------------------------------------------------------------|------|
| cctggctcta ttaccggact ctgtaagcgc atggacgcgg acgccgtgcc taccagaggc   | 900  |
| gagacaatcg ggaagaagac cgcgtcgtct gcctgggacc ctgcgaccgt catgcgtatt   | 960  |
| ctcagagacc ctcgtatcgc cgggttcgct gcggagggtga tttacaagaa gaagccagac  | 1020 |
| ggcacaccta ccacaaagat cgagggatac cgcattccagc gcgaccctat tactctgcgg  | 1080 |
| cctgtggagc ttgattgcgg tcctattatc gagcctgcgg agtgggtatga gcttcaggcc  | 1140 |
| tggttggacg gacgtggctg cggcaagggt ctctctcggg gtcaagccat cctgtctgct   | 1200 |
| atggacaagc tgtactgcga gtgtggcgcc gttatgacga gcaagcgagg ggaagaatct   | 1260 |
| atcaaggaca gttaccgctg ccgtcgcaga aagggtgggtg acccttctgc gcccggtcag  | 1320 |
| cacgaaggca cttgcaacgt ctctatggcc gcgctggaca agttcgtcgc cgaacgcatt   | 1380 |
| ttcaacaaga tccgtcacgc cgaaggcgac gaagagacac ttgccctcct gtgggaagcc   | 1440 |
| gcccgtcgct tcggcaagct cacggaggcc cccgagaagt ctggcgaaaag agccaacctc  | 1500 |
| gtcgccgagc gcgccgacgc cctgaacgcc ctcgaagagc tgtacgaaga ccgcgctgcg   | 1560 |
| ggcgccctacg acggtcctgt cggacgaaaag cacttcagaa agcaacaggc ggccctgact | 1620 |
| ctgcgccagc aagggtgccga agagagactc gccgaactcg aagccgccga agccccaag   | 1680 |
| ctccctctcg accaatggtt cccagaagac gccgacgcgg accctaccgg ccccaagtct   | 1740 |
| tgggtggggtc gcgcctcggg agacgacaag cgcgtgttcg tgggtctgtt cgtagacaag  | 1800 |
| attgtcggtta caaagtctac gacaggccgt gggcagggga cacctatcga gaagcgcgcg  | 1860 |
| tctattactt gggccaagcc tcctaccgac gacgacgaag acgacgcca ggacggcaca    | 1920 |
| gaagacgtag ctgcttgata a                                             | 1941 |

<210> 6

<211> 34

<212> DNA

<213> Artificial sequence

<220>

<223> Partly modified native lox gene

<400> 6

ataacttcgt ataggatact ttatacgaag ttat

34

<210> 7  
<211> 1031  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Partly modified native Cre gene

<400> 7  
atgtcaaacc ttctcaccgt ccacaaaaac ctccccgcac tccccgttga cgccacctcc 60  
gacgagggtca gaaaaaacct catggacatg ttccgggacc gccaggcctt ttccgaacac 120  
acttgaaaaa tgcttctcag cgtttgccgt agttgggccg cttggtgtaa actcaacaac 180  
cgcaagtgggt tccccgccga acccgaggac gtccgcgatt accttctgta ttgcaagcg 240  
cgaggactgg ccgtgaaaac catccagcaa catctgggtc agcttaacat gttgcaccgg 300  
aggagcggcc tgccacggcc tagcgactcc aacgcggtgt ccctcgtgat gaggagaatc 360  
cgcaaggaga atgtggacgc cggagaaaaga gcaaagcagg ccctggcctt cgagaggact 420  
gacttcgacc aagtccggtc gctgatggag aactcggacc gatgtcagga catcaggaac 480  
ctcgcattct cggcattgcc tacaacaccc tgctgagaat tgcagagatc gcccgcattc 540  
gcgtcaagga catttcgaga accgacggag ggcggtatgct gattcacatc ggcaggacta 600  
agaccctcgt gtcaaccgcc ggagtggaaa aggccctcag cctgggagtg acaaagctcg 660  
tggagcgctg gatctccgtg tcgggggtgg ccgacgatcc gaacaattac ctgttctgcc 720  
gggtccgcaa aaatgggggtg gccgccccgt ctgctacaag ccagttgtcc actcgcgccc 780  
tggaagggaat cttcgaggcc acgcaccgcc tgatctatgg ggcaaaggac gattccggcc 840  
agaggatatct cgcgtggtcc ggtcactccg cgcgcgtggg cgcgggcccg gacatggccc 900  
gggctggagt gtccatccct gaaatcatgc aggccggtgg atggaccaac gtgaacatcg 960  
tgatgaacta cattcggaac ctggacagcg aaactgggtc tatggtccgc ctgctggagg 1020  
acggagattg a 1031

<210> 8  
<211> 1428

<212> DNA

<213> Artificial sequence

<220>

<223> Enhanced Green Fluorescent Protein fused to an FC from IgG1  
(FC-eGFP gene)

<400> 8

|                                    |                                     |                             |
|------------------------------------|-------------------------------------|-----------------------------|
| gacaaaactc acacatgccc accgtgccc    | gcacctgaac tcctgggggg accgtcagtc    | 60                          |
| ttcctcttcc ccccaaaacc caaggacacc   | ctcatgatct cccggacccc tgaggtcaca    | 120                         |
| tgcgtggtgg tggacgtgag ccacgaagac   | cctgaggtca agttcaactg gtacgtggac    | 180                         |
| ggcgtggagg tgcataatgc caagacaaaag  | ccacgggagg agcagtacaa cagcacgtac    | 240                         |
| cgtgtggtca gcgtcctcac cgtcctgcac   | caggactggc tgaatggcaa ggagtacaag    | 300                         |
| tgcaaggtct ccaacaaagc cctcccagcc   | cccatcgaga aaaccatctc caaagccaaa    | 360                         |
| gggcagcccc gagaaccaca ggtctacacc   | ctgcccccat cccgggagga gatgaccaag    | 420                         |
| aaccagggtca gcctgacctg cctgggtcaaa | ggcttctatc ccagcgacat cgccgtggag    | 480                         |
| tgggagagca atgggcagcc ggagaacaac   | tacaagacca cgcctcccgt gctggactcc    | 540                         |
| gacggctcct tcttcctcta cagcaagctc   | accgtggaca agagcaggtg gcagcagggg    | 600                         |
| aacgtcttct catgctccgt gatgcacgag   | gctctgcaca accactacac gcagaagagc    | 660                         |
| ctctccctgt ctccgggtaa aggttcctcc   | agttccggca gctccagttc cggtatgagt    | 720                         |
| aaaggagagg aactcttcac cggagtcgtc   | ccgatactcg tcgagctaga cggagacgtc    | 780                         |
| aacggccaca aattctccgt ctccggcgag   | ggggaggggg acgccaccta cggaaaactc    | 840                         |
| acccttaagt ttatttgcac taccggaaaa   | ctccccgtcc cttggccaac cctagtcacc    | 900                         |
| acgctgacat acggagtcca atgtttctcg   | cggatatcccg accacatgaa gcagcatgac   | 960                         |
| tttttcaa                           | at cgcgatgcc tgagggtctac gtgcaggaac | gcaccatctt cttcaaggac 1020  |
| gacgggaatt                         | acaagactag agccgaggtc aagtttgaag    | gagacaccct cgtgaatcgc 1080  |
| atcgagctta                         | agggcattga cttcaaggag gacggcaaca    | tcctgggtca caagctggag 1140  |
| tacaactaca                         | actcgcataa cgtctacatc atggccgaca    | agcaaaaagaa cggtatcaag 1200 |
| gtcaacttca                         | agattaggca caacattgag gatgggtccg    | tccaactggc cgaccactac 1260  |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| cagcagaaca ccccatcgg cgacggacct gtgctcctgc ctgataacca ctatctcagc  | 1320 |
| actcagagcg cactgtccaa ggaccctaac gaaaaacggg accacatggt cttgctggag | 1380 |
| ttcgtgacag ccgctgggtat taccctgggc atggatgaac tgtataag             | 1428 |

<210> 9  
 <211> 1431  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Blue Fluorescent Protein variant fused to an FC from IgG1  
 (FC-TagBFP2 gene)

|                                                                     |      |
|---------------------------------------------------------------------|------|
| <400> 9                                                             |      |
| gacaaaactc acacatgccc accgtgcccc gcacctgaac tcctgggggg accgtcagtc   | 60   |
| ttcctcttcc ccccaaaacc caaggacacc ctcatgatct cccggacccc tgaggtcaca   | 120  |
| tgcgtgggtg tggacgtgag ccacgaagac cctgagggtca agttcaactg gtacgtggac  | 180  |
| ggcgtggagg tgcataatgc caagacaaaag ccacgggagg agcagtacaa cagcacgtac  | 240  |
| cgtgtgggtca gcgtcctcac cgtcctgcac caggactggc tgaatggcaa ggagtacaag  | 300  |
| tgcaagggtct ccaacaaagc cctcccagcc cccatcgaga aaaccatctc caaagccaaa  | 360  |
| gggcagcccc gagaaccaca ggtctacacc ctgcccccat cccgggagga gatgaccaag   | 420  |
| aaccagggtca gcctgacctg cctgggtcaaa ggcttctatc ccagcgacat cgccgtggag | 480  |
| tgggagagca atgggcagcc ggagaacaac tacaagacca cgcctcccgt gctggactcc   | 540  |
| gacggctcct tcttctctca cagcaagctc accgtggaca agagcagggtg gcagcagggg  | 600  |
| aacgtcttct catgctccgt gatgcacgag gctctgcaca accactacac gcagaagagc   | 660  |
| ctctccctgt ctccgggtaa aggttcctcc agttccggca gctccagttc cggtatgggtg  | 720  |
| tcgaaggggag aggagctgat taaggagaac atgcacatga agctgtatat ggaaggggacg | 780  |
| gtggacaacc accacttcaa gtgcaccagc gaaggagaag gaaagcctta cgaaggcact   | 840  |
| caaactatgc ggatcaaagt ggtggaaggc ggtcctcttc cgttcgctt cgacatcttg    | 900  |
| gccacctcct tcctctacgg ctccaagacc tttatcaacc acaccaggga aatcccggac   | 960  |
| ttctttaagc agagcttccc tgagggttc acctgggaaa gagtgacaac ctacgaggac    | 1020 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ggtggcgtcc tgaccgcgac ccaggacacc tccctgcaag acggctgcct gatctacaac  | 1080 |
| gtcaagattc gcggcgtgaa cttcacctcc aatgggtccag tgatgcagaa gaaaactctg | 1140 |
| ggatgggagg ccttcactga aactctgtac cccgccgatg gaggactgga ggggaggaac  | 1200 |
| gatatggctt tgaagctcgt ggggggatcg cacctgattg cgaatgcca gaccacctac   | 1260 |
| agatccaaga aacccgcca gaacctcaag atgcccggag tctactacgt ggactataga   | 1320 |
| ctggaacgga tcaaggaagc caacaacgag acttacgtgg aacagcacga ggtcgctgtg  | 1380 |
| gcacgctact gtgatctgcc gtcaaagctc gggcataagc tcaactgata a           | 1431 |

<210> 10  
 <211> 738  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Red Fluorescent Protein Variant (TagRFP-T gene)

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| <400> 10                                                            |     |
| atgggtgtcaa agggagagga actgattaag gagaatatgc acatgaaact ctacatggag  | 60  |
| gggaccgtga acaaccacca cttcaagtgc acctccgagg gcgaagggaa gccgtacgag   | 120 |
| ggaactcaga ccatgcggat taaggctgtc gaagggggtc ctctgccatt cgccttcgac   | 180 |
| atcctcgcca catcctttat gtacggatcg cggaccttca tcaaccacac tcagggtatc   | 240 |
| cccgacttct tcaagcaatc gttcccggaa ggctttactt gggagcgcgt gaccacctac   | 300 |
| gaggatggag ggggtgctgac ggccactcag gacaccagcc tgcaagacgg ctgtcttatc  | 360 |
| tacaacgtga agattcgcg cgtgaacttc cctagcaacg gtccgggtcat gcagaaaaag   | 420 |
| accctggggtt gggagggctaa caccgaaatg ctctatcctg cggacggagg attggaaggc | 480 |
| cggactgaca tggccctgaa acttgtgggc ggcggacatc tgatctgcaa tttcaagacc   | 540 |
| acttaccgct ccaagaagcc cgccaagaac ctgaagatgc ctggagtgtg ctacgtggac   | 600 |
| cacagactcg aaaggatcaa ggaggcggat aaggaaacct acgtggaaca gcatgaagtg   | 660 |
| gcagtggcca gatactgcga tctgccgtcc aagctcggcc acaagctgaa cggaatggac   | 720 |
| gagctgtata agtgataa                                                 | 738 |

<210> 11  
<211> 714  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Enhanced Green Fluorescent Protein (eGFP gene)

<400> 11  
atgagtaaag gagaggaact cttcaccgga gtcgtcccga tactcgtcga gctagacgga 60  
gacgtcaacg gccacaaatt ctccgtctcc ggcgaggggg aggggggacgc cacctacgga 120  
aaactcacc ttaagtttat ttgcactacc ggaaaactcc ccgtcccttg gccaaccccta 180  
gtcaccacgc tgacatacgg agtccaatgt ttctcgcggt atcccgaacca catgaagcag 240  
catgactttt tcaaatccgc gatgcctgag ggctacgtgc aggaacgcac catcttcttc 300  
aaggacgacg ggaattacaa gactagagcc gaggtcaagt ttgaaggaga caccctcgtg 360  
aatcgcacgc agcttaaggg cattgacttc aaggaggacg gcaacatcct gggtcacaag 420  
ctggagtaca actacaactc gcataacgtc tacatcatgg ccgacaagca aaagaacggt 480  
atcaagggtca acttcaagat taggcacaac attgaggatg ggtccgtcca actggccgac 540  
cactaccagc agaacacccc catcggcgac ggacctgtgc tcctgcctga taaccactat 600  
ctcagcactc agagcgact gtccaaggac cctaacgaaa aacgggacca catgggtcttg 660  
ctggagttcg tgacagccgc tgggtattacc ctgggcatgg atgaactgta taag 714

<210> 12  
<211> 717  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Blue Fluorescent Protein variant (TagBFP2 gene)

<400> 12  
atgggtgtcga agggagagga gctgattaag gagaacatgc acatgaagct gtatatggaa 60  
gggacggtgg acaaccacca cttcaagtgc accagcgaag gagaaggaaa gccttacgaa 120  
ggcactcaaa ctatgcggat caaagtgggtg gaaggcggtc ctcttccgtt cgccttcgac 180

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atcttggcca cctccttcct ctacggctcc aagaccttta tcaaccacac ccagggaatc | 240 |
| ccggacttct ttaagcagag cttccctgag ggcttcacct gggaaagagt gacaacctac | 300 |
| gaggacggtg gcgtcctgac cgcgacccag gacacctccc tgcaagacgg ctgcctgac  | 360 |
| tacaacgtca agattcgcgg cgtgaacttc acctccaatg gtccagtgat gcagaagaaa | 420 |
| actctgggat gggaggcctt cactgaaact ctgtaccccg ccgatggagg actggagggg | 480 |
| aggaacgata tggctttgaa gctcgtgggg ggatcgcacc tgattgcgaa tgccaagacc | 540 |
| acctacagat ccaagaaacc cgccaagaac ctcaagatgc ccggagtcta ctacgtggac | 600 |
| tatagactgg aacggatcaa ggaagccaac aacgagactt acgtggaaca gcacgaggtc | 660 |
| gctgtggcac gctactgtga tctgccgtca aagctcgggc ataagctcaa ctgataa    | 717 |

<210> 13  
 <211> 792  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Neomycin resistance gene (NeoR gene)

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| <400> 13                                                           |     |
| atgattgaac aagatggatt gcacgcagg tctccggccg cttgggtgga gaggctattc   | 60  |
| ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca  | 120 |
| gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg  | 180 |
| caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg  | 240 |
| ctcgacgttg tcaactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag | 300 |
| gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg  | 360 |
| cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc  | 420 |
| atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa  | 480 |
| gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgcac | 540 |
| ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat  | 600 |
| ggccgctttt ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac  | 660 |

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atagcgttgg ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc | 720 |
| ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt | 780 |
| gacgagttct tc                                                     | 792 |

<210> 14  
 <211> 1122  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Partially modified native Glutamine Synthetase gene (GS gene)

|                                                                    |      |
|--------------------------------------------------------------------|------|
| <400> 14                                                           |      |
| atggccacct cagcaagttc ccacttgaac aaaaacatca agcaaatgta cttgtgcctg  | 60   |
| ccccagggtg agaaagtcca agccatgtat atctgggttg atggtactgg agaaggactg  | 120  |
| cgctgcaaaa ccgcaccct ggactgtgag cccaagtgtg tagaagagtt acctgagtgg   | 180  |
| aattttgatg gctctagtag ctttcagtct gagggctcca acagtgacat gtatctcagc  | 240  |
| cctgttgcca tgtttcggga ccccttccgc agagatccca acaagctggg gttctgtgaa  | 300  |
| gttttcaagt acaaccggaa gcctgcagag accaatttaa ggcaactcgtg taaacggata | 360  |
| atggacatgg tgagcaacca gcacccctgg tttggaatgg aacaggagta tactctgatg  | 420  |
| ggaacagatg ggcacccttt tggttggcct tccaatggct ttcctgggcc ccaaggccg   | 480  |
| tattactgtg gtgtgggcgc agacaaagcc tatggcaggg atatcgtgga ggctcactac  | 540  |
| cgcgcctgct tgtatgctgg ggtcaagatt acaggaacaa atgctgaggt catgcctgcc  | 600  |
| cagtgggagt tccaaatagg accctgtgaa ggaatccgca tgggagatca tctctgggtg  | 660  |
| gcccgtttca tcttgcacgc agtatgtgaa gactttgggg taatagcaac ctttgacccc  | 720  |
| aagcccattc ctgggaactg gaatggtgca ggctgccata ccaacttttag caccaaggcc | 780  |
| atgcgggagg agaatggtct gaagcacatc gaggaggcca tcgagaaaact aagcaagcgg | 840  |
| caccgctacc acattcgagc ctacgatccc aagggggggcc tggacaatgc ccgtcgtctg | 900  |
| actgggttcc acgaaacgtc caacatcaac gacttttctg ctggtgtcgc caatcgcagt  | 960  |
| gccagcatcc gcattccccg gactgtcggc caggagaaga aaggttactt tgaagaccgc  | 1020 |

cgccccctctg ccaactgtga cccctttgca gtgacagaag ccatcgtccg cacatgcctt 1080

ctcaatgaga ctggcgacga gcccttccaa tacaaaaact aa 1122

<210> 15

<211> 39

<212> DNA

<213> Artificial sequence

<220>

<223> 462 5' part of 5'UTR

<400> 15

accatgggtt gaaccatggg ttgaaccatg ggttgaacc 39

<210> 16

<211> 15

<212> DNA

<213> Artificial sequence

<220>

<223> 464 5' part of 5'UTR

<400> 16

caaatgggtt gaacc 15