P450 SINGLE NUCLEOTIDE POLYMORPHISM BIOCHIP ANALYSIS

Inventors: Buena Chui, Chandler, AZ (US); Robert Elghanian, Skokie, IL (US); Vineet Gupta, Reading, MA (US); Krishnamurthy Jayaraman, Hoffman Estates, IL (US); Gretchen Kiser, Mesa, AZ (US); Changming Li, Schaumburg, IL (US); Chang-Gong Liu, Cherry Hill, NJ (US); Kenneth R. Luehrsen, Half Moon Bay, CA (US); Abhijit Mazumder, Buffalo Grove, IL (US); Ramesh Ramakrishnan, Vernon Hills, IL (US); Arkady Silbergleyt, Chandler, AZ (US); Todd Tuggle, Oceanside, CA (US); Carl Yamashiro, Chandler, AZ (US); Handy Yawanto, Walmut, CA (US); Ekaterina Pestova, Downers Grove, IL (US); David R. Fermin, Minneapolis, MN (US); David G. Wang, Deerfield, IL (US); Zhijie John Gu, San Diego, CA (US)

Correspondence Address:
DORSEY & WHITNEY LLP
Suite 3400
Four Embarcadero Center
San Francisco, CA 94111-4187 (US)

(21) Appl. No.: 10/114,908
(22) Filed: Apr. 1, 2002

Related U.S. Application Data
Provisional application No. 60/280,583, filed on Mar. 30, 2001.

Publication Classification
Int. Cl. 7 .......................... C12Q 1/68; G01N 33/53; C12M 1/34
U.S. Cl. ............................. 435/6; 435/7.5; 435/287.2

ABSTRACT

This invention relates to methods and compositions for determining single nucleotide polymorphisms (SNPs) in P450 genes. In preferred embodiments, self extension of interrogation probes is prevented by using novel non self-extension probes and/or methods, thereby improving the specificity and efficiency of P450 SNP detection in target samples with minimal false positive results. The invention thus describes a variety of methods to decrease self-extension of interrogation probes. In addition, the invention provides a unique collection of P450 SNP probes on one assay, primer sequences for specific amplification of each of the seven P450 genes and amplicon control probes to evaluate whether the intended p450 gene targets were amplified successfully. The invention also describes a variety of array platforms for performing the assays of the invention; for example: CodeLink™, cSensor™, multiplex arrays with cartridges etc., all described herein.
A. Target-dependent Extension

Target Binding, Polymerase Extension
Surface-attached Probe

Wash
Correct Signal

B. Target-independent Extension

Self-folding, Polymerase Extension
Surface-attached Probe

Wash
Incorrect Signal

FIG. 2
<table>
<thead>
<tr>
<th>SNP ID</th>
<th>GENE</th>
<th>POSITION</th>
<th>FROM TRANSLATION START</th>
<th>SEQ SOURCE</th>
<th>EXON/INTRON LOCATION</th>
<th>MUTATION</th>
<th>E FLANK</th>
<th>ENZYME EFFECT (IN VIVO)</th>
<th>P450 SNP CHARACTERIZATION LIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>1618</td>
<td>-1</td>
<td>EXON 1</td>
<td>E1</td>
<td>G &gt; A</td>
<td>GB</td>
<td>V.7 M</td>
<td>Normal</td>
</tr>
<tr>
<td>318</td>
<td>CYP2D6</td>
<td>1639</td>
<td>19</td>
<td>INTRON 1</td>
<td>S1</td>
<td>G &gt; A</td>
<td>GB</td>
<td>V.7 M</td>
<td>Glucuronidase</td>
</tr>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>1650</td>
<td>31</td>
<td>EXON 2</td>
<td>E2</td>
<td>G &gt; A</td>
<td>GB</td>
<td>R.38 H</td>
<td>Normal</td>
</tr>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>1767</td>
<td>77</td>
<td>EXON 3</td>
<td>E3</td>
<td>G &gt; A</td>
<td>GB</td>
<td>R.38 H</td>
<td>Normal</td>
</tr>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>1719</td>
<td>124</td>
<td>EXON 4</td>
<td>E4</td>
<td>G &gt; A</td>
<td>GB</td>
<td>R.38 H</td>
<td>Normal</td>
</tr>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>1743</td>
<td>174</td>
<td>EXON 5</td>
<td>E5</td>
<td>G &gt; A</td>
<td>GB</td>
<td>R.38 H</td>
<td>Normal</td>
</tr>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>1757</td>
<td>250</td>
<td>EXON 6</td>
<td>E6</td>
<td>G &gt; A</td>
<td>GB</td>
<td>R.38 H</td>
<td>Normal</td>
</tr>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>2502</td>
<td>293</td>
<td>EXON 7</td>
<td>E7</td>
<td>G &gt; A</td>
<td>GB</td>
<td>R.38 H</td>
<td>Normal</td>
</tr>
<tr>
<td>265</td>
<td>CYP2D6</td>
<td>2576</td>
<td>397</td>
<td>EXON 8</td>
<td>E8</td>
<td>G &gt; A</td>
<td>GB</td>
<td>R.38 H</td>
<td>Normal</td>
</tr>
<tr>
<td>SNP ID</td>
<td>GENE</td>
<td>POSITION</td>
<td>FROM TRANSLATION START</td>
<td>SEQ_SOURCE</td>
<td>EXON / INTRON LOCATION</td>
<td>MUTATION</td>
<td>CONSEQUENCE</td>
<td>ENZYME_EFFECT (IN VIVO)</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>----------</td>
<td>------------------------</td>
<td>------------</td>
<td>------------------------</td>
<td>----------</td>
<td>---------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>2D6.4089</td>
<td>CYP 2D6</td>
<td>4089</td>
<td>2470</td>
<td>GB [M33388]</td>
<td>EXON 5</td>
<td>T&gt;C</td>
<td>SILENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4099</td>
<td>CYP 2D6</td>
<td>4099</td>
<td>2480</td>
<td>GB [M33388]</td>
<td>EXON 5</td>
<td>C&gt;T</td>
<td>SILENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4102</td>
<td>CYP 2D6</td>
<td>4102</td>
<td>2483</td>
<td>GB [M33388]</td>
<td>EXON 5</td>
<td>G&gt;T</td>
<td>A 237 S</td>
<td>NORMAL</td>
<td></td>
</tr>
<tr>
<td>2D6.4168</td>
<td>CYP 2D6</td>
<td>4168</td>
<td>2549</td>
<td>GB [M33388]</td>
<td>EXON 5</td>
<td>A&gt;del</td>
<td>FRAMESHIFT</td>
<td>NO ACTIVITY</td>
<td></td>
</tr>
<tr>
<td>2D6.4194</td>
<td>CYP 2D6</td>
<td>4194</td>
<td>2575</td>
<td>GB [M33388]</td>
<td>EXON 5</td>
<td>C&gt;A</td>
<td>SILENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4206</td>
<td>CYP 2D6</td>
<td>4206</td>
<td>2587</td>
<td>GB [M33388]</td>
<td>EXON 5</td>
<td>GACT&gt;del</td>
<td>FRAMESHIFT</td>
<td>NO ACTIVITY</td>
<td></td>
</tr>
<tr>
<td>2D6.4232</td>
<td>CYP 2D6</td>
<td>4232</td>
<td>2613</td>
<td>GB [M33388]</td>
<td>EXON 5</td>
<td>AGA&gt;del</td>
<td>K 281 DEL</td>
<td>DECREASED</td>
<td></td>
</tr>
<tr>
<td>2D6.4469</td>
<td>CYP 2D6</td>
<td>4469</td>
<td>2850</td>
<td>GB [M33388]</td>
<td>EXON 6</td>
<td>C&gt;T</td>
<td>R 296 C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4472</td>
<td>CYP 2D6</td>
<td>4472</td>
<td>2853</td>
<td>GB [M33388]</td>
<td>EXON 6</td>
<td>A&gt;C</td>
<td>I 297 L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4554</td>
<td>CYP 2D6</td>
<td>4554</td>
<td>2935</td>
<td>GB [M33388]</td>
<td>EXON 6</td>
<td>A&gt;C</td>
<td>H 324 P</td>
<td>NONE</td>
<td></td>
</tr>
<tr>
<td>2D6.4557</td>
<td>CYP 2D6</td>
<td>4557</td>
<td>2938</td>
<td>GB [M33388]</td>
<td>EXON 6</td>
<td>C&gt;T</td>
<td>P 325 L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4558</td>
<td>CYP 2D6</td>
<td>4558</td>
<td>2939</td>
<td>GB [M33388]</td>
<td>EXON 6</td>
<td>G&gt;A</td>
<td>SILENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4802</td>
<td>CYP 2D6</td>
<td>4802</td>
<td>3183</td>
<td>GB [M33388]</td>
<td>EXON 7</td>
<td>G&gt;A</td>
<td>V 338 M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4817</td>
<td>CYP 2D6</td>
<td>4817</td>
<td>3198</td>
<td>GB [M33388]</td>
<td>EXON 7</td>
<td>C&gt;G</td>
<td>R 343 G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4896</td>
<td>CYP 2D6</td>
<td>4896</td>
<td>3277</td>
<td>GB [M33388]</td>
<td>EXON 7</td>
<td>T&gt;C</td>
<td>I 369 T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.4907</td>
<td>CYP 2D6</td>
<td>4907</td>
<td>3288</td>
<td>GB [M33388]</td>
<td>EXON 7</td>
<td>G&gt;A</td>
<td>G 373 S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.5447</td>
<td>CYP 2D6</td>
<td>5447</td>
<td>3828</td>
<td>GB [M33388]</td>
<td>EXON 8</td>
<td>G&gt;A</td>
<td>SILENT</td>
<td>NORMAL</td>
<td></td>
</tr>
<tr>
<td>2D6.5472</td>
<td>CYP 2D6</td>
<td>5472</td>
<td>3853</td>
<td>GB [M33388]</td>
<td>EXON 8</td>
<td>G&gt;A</td>
<td>E 410 K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.5496</td>
<td>CYP 2D6</td>
<td>5496</td>
<td>3877</td>
<td>GB [M33388]</td>
<td>EXON 8</td>
<td>G&gt;C</td>
<td>E 418 Q</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.5506</td>
<td>CYP 2D6</td>
<td>5506</td>
<td>3887</td>
<td>GB [M33388]</td>
<td>EXON 8</td>
<td>T&gt;C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.5661</td>
<td>CYP 2D6</td>
<td>5661</td>
<td>4042</td>
<td>GB [M33388]</td>
<td>EXON 9</td>
<td>G&gt;A</td>
<td>R 440 H</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6.5734</td>
<td>CYP 2D6</td>
<td>5734</td>
<td>4115</td>
<td>GB [M33388]</td>
<td>EXON 9</td>
<td>C&gt;T</td>
<td>SILENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1.1532</td>
<td>CYP 2E1</td>
<td>1532</td>
<td>-1293</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>G&gt;C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1.1627</td>
<td>CYP 2E1</td>
<td>1627</td>
<td>-1198</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>G&gt;C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1.1772</td>
<td>CYP 2E1</td>
<td>1772</td>
<td>-1053</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>C&gt;T</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1.1800</td>
<td>CYP 2E1</td>
<td>1800</td>
<td>-1025</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>T&gt;C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1.2019</td>
<td>CYP 2E1</td>
<td>2019</td>
<td>-806</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>T&gt;C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1.2473</td>
<td>CYP 2E1</td>
<td>2473</td>
<td>-352</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>A&gt;G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1.2492</td>
<td>CYP 2E1</td>
<td>2492</td>
<td>-333</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>T&gt;A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIG. 3B**
<table>
<thead>
<tr>
<th>SNP ID</th>
<th>GENE</th>
<th>POSITION</th>
<th>FROM TRANSLATION START</th>
<th>SEQ. SOURCE</th>
<th>EXON / INTRON LOCATION</th>
<th>MUTATION</th>
<th>CONSEQUENCE</th>
<th>ENZYME_EFFECT (IN VIVO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2E1_2754</td>
<td>CYP 2E1</td>
<td>2754</td>
<td>-71</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>G&gt;T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1_3104</td>
<td>CYP 2E1</td>
<td>3104</td>
<td>280</td>
<td>GB [J02843]</td>
<td>INTRON A</td>
<td>NONE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1_3085</td>
<td>CYP 2E1</td>
<td>3085</td>
<td>261</td>
<td>GB [J02843]</td>
<td>INTRON A</td>
<td>NONE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1_3956</td>
<td>CYP 2E1</td>
<td>3956</td>
<td>1132</td>
<td>GB [J02843]</td>
<td>EXON 2</td>
<td>G&gt;A</td>
<td>R 76 H</td>
<td>REDUCED</td>
</tr>
<tr>
<td>2E1_7592</td>
<td>CYP 2E1</td>
<td>7592</td>
<td>4768</td>
<td>GB [J02843]</td>
<td>EXON 4</td>
<td>G&gt;A</td>
<td>V 179 I</td>
<td>NORMAL</td>
</tr>
<tr>
<td>2E1_10456</td>
<td>CYP 2E1</td>
<td>10456</td>
<td>7632</td>
<td>GB [J02843]</td>
<td>INTRON F</td>
<td>T&gt;A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1_12720</td>
<td>CYP 2E1</td>
<td>12720</td>
<td>9896</td>
<td>GB [J02843]</td>
<td>INTRON G</td>
<td>C&gt;G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2E1_12847</td>
<td>CYP 2E1</td>
<td>12847</td>
<td>10023</td>
<td>GB [J02843]</td>
<td>EXON 8</td>
<td>G&gt;A</td>
<td>V 389 I</td>
<td>NORMAL</td>
</tr>
<tr>
<td>2E1_12945</td>
<td>CYP 2E1</td>
<td>12945</td>
<td>10121</td>
<td>GB [J02843]</td>
<td>EXON 8</td>
<td>C&gt;T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3A4_816</td>
<td>CYP 3A4</td>
<td>816</td>
<td>-392</td>
<td>GB [D11131]</td>
<td>5' FLANK</td>
<td>A&gt;G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3A4_918</td>
<td>CYP 3A4</td>
<td>918</td>
<td>-290</td>
<td>GB [D11131]</td>
<td>5' FLANK</td>
<td>A&gt;G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A1_1213</td>
<td>CYP 1A1</td>
<td>1213</td>
<td>-3229</td>
<td>GB [X04300]</td>
<td>5' FLANK</td>
<td>G&gt;A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A1_1223</td>
<td>CYP 1A1</td>
<td>1223</td>
<td>-3219</td>
<td>GB [X04300]</td>
<td>5' FLANK</td>
<td>C&gt;T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A1_5568</td>
<td>CYP 1A1</td>
<td>5568</td>
<td>2453</td>
<td>GB [X04300]</td>
<td>EXON 7</td>
<td>C&gt;A</td>
<td>T 461 N</td>
<td></td>
</tr>
<tr>
<td>1A1_6570</td>
<td>CYP 1A1</td>
<td>6570</td>
<td>2455</td>
<td>GB [X04300]</td>
<td>EXON 7</td>
<td>A&gt;G</td>
<td>I 462 V</td>
<td></td>
</tr>
<tr>
<td>1A1_7320</td>
<td>CYP 1A1</td>
<td>7320</td>
<td>3205</td>
<td>GB [X04300]</td>
<td>3' FLANK</td>
<td>T&gt;C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A1_7916</td>
<td>CYP 1A1</td>
<td>7916</td>
<td>3800</td>
<td>GB [X04300]</td>
<td>3' FLANK/FLOATER</td>
<td>T&gt;C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1B1_3793</td>
<td>CYP 1B1</td>
<td>3793</td>
<td>-13</td>
<td>GB [U56438]</td>
<td>INTRON 1 (not translated)</td>
<td>T&gt;C</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>1B1_3947</td>
<td>CYP 1B1</td>
<td>3947</td>
<td>142</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>C&gt;G</td>
<td>R 48 G</td>
<td></td>
</tr>
<tr>
<td>1B1_3976</td>
<td>CYP 1B1</td>
<td>3976</td>
<td>171</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>G&gt;C</td>
<td>W 57 C</td>
<td></td>
</tr>
<tr>
<td>1B1_3987</td>
<td>CYP 1B1</td>
<td>3987</td>
<td>182</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>G&gt;A</td>
<td>G 61 E</td>
<td></td>
</tr>
<tr>
<td>1B1_4035</td>
<td>CYP 1B1</td>
<td>4035</td>
<td>230</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>T&gt;C</td>
<td>L 77 P</td>
<td>GLAUCOMA</td>
</tr>
<tr>
<td>1B1_4160</td>
<td>CYP 1B1</td>
<td>4160</td>
<td>355</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>G&gt;T</td>
<td>A 119 S</td>
<td></td>
</tr>
<tr>
<td>1B1_4306</td>
<td>CYP 1B1</td>
<td>4306</td>
<td>501</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>insT</td>
<td>FRAMESHIFT</td>
<td></td>
</tr>
<tr>
<td>1B1_4646</td>
<td>CYP 1B1</td>
<td>4646</td>
<td>841</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>G&gt;T</td>
<td>E 281 X</td>
<td></td>
</tr>
<tr>
<td>1B1_4668</td>
<td>CYP 1B1</td>
<td>4668</td>
<td>863</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>insC</td>
<td>FRAMESHIFT</td>
<td></td>
</tr>
<tr>
<td>1B1_7930</td>
<td>CYP 1B1</td>
<td>7930</td>
<td>4125</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>G&gt;T</td>
<td>G 365 W</td>
<td></td>
</tr>
<tr>
<td>1B1_7940</td>
<td>CYP 1B1</td>
<td>7940</td>
<td>4135</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>G&gt;A</td>
<td>R 388 H</td>
<td>GLAUCOMA</td>
</tr>
<tr>
<td>1B1_7957</td>
<td>CYP 1B1</td>
<td>7957</td>
<td>4152</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>G&gt;A</td>
<td>D 374 N</td>
<td>GLAUCOMA</td>
</tr>
</tbody>
</table>

**FIG. 3C**
<table>
<thead>
<tr>
<th>SNP ID</th>
<th>GENE</th>
<th>POSITION</th>
<th>FROM TRANSLATION START</th>
<th>SEQ_SOURCE</th>
<th>EXON / INTRON LOCATION</th>
<th>MUTATION</th>
<th>CONSEQUENCE</th>
<th>ENZYME_EFFECT (IN VIVO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1B1_7973</td>
<td>CYP 1B1</td>
<td>7973</td>
<td>4168</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>C&gt;T</td>
<td>P 379 L</td>
<td></td>
</tr>
<tr>
<td>1B1_7996</td>
<td>CYP 1B1</td>
<td>7996</td>
<td>4191</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>G&gt;A</td>
<td>E 387 K</td>
<td></td>
</tr>
<tr>
<td>1B1_8006</td>
<td>CYP 1B1</td>
<td>8006</td>
<td>4201</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>G&gt;A</td>
<td>R 390 H</td>
<td></td>
</tr>
<tr>
<td>1B1_8131</td>
<td>CYP 1B1</td>
<td>8131</td>
<td>4326</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>C&gt;G</td>
<td>L 432 V</td>
<td></td>
</tr>
<tr>
<td>1B1_8147</td>
<td>CYP 1B1</td>
<td>8147</td>
<td>4342</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>C&gt;T</td>
<td>P 437 L</td>
<td></td>
</tr>
<tr>
<td>1B1_8184</td>
<td>CYP 1B1</td>
<td>8184</td>
<td>4379</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>T&gt;C</td>
<td>449 SILENT</td>
<td></td>
</tr>
<tr>
<td>1B1_8195</td>
<td>CYP 1B1</td>
<td>8195</td>
<td>4390</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>A&gt;G</td>
<td>N 453 V</td>
<td></td>
</tr>
<tr>
<td>1B1_8242</td>
<td>CYP 1B1</td>
<td>8242</td>
<td>4437</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>C&gt;T</td>
<td>R 469 W</td>
<td></td>
</tr>
<tr>
<td>1B1_8587</td>
<td>CYP 1B1</td>
<td>8587</td>
<td>4782</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>C&gt;G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1B1_8807</td>
<td>CYP 1B1</td>
<td>8807</td>
<td>5002</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>T&gt;A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1B1_9164</td>
<td>CYP 1B1</td>
<td>9164</td>
<td>5359</td>
<td>GB [U56438]</td>
<td>3' FLANK</td>
<td>T&gt;G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1A2_2640</td>
<td>CYP 1A2</td>
<td>2640</td>
<td>-164</td>
<td>GB [M31664]</td>
<td>INTRON A (5' FLANK)</td>
<td>C&gt;A</td>
<td>HIGHER INDUCIBILITY</td>
<td></td>
</tr>
<tr>
<td>1A2_2866</td>
<td>CYP 1A2</td>
<td>2866</td>
<td>63</td>
<td>GB [M31664]</td>
<td>EXON 1</td>
<td>C&gt;G</td>
<td>F 21 L</td>
<td></td>
</tr>
<tr>
<td>2C19_99</td>
<td>CYP 2C19</td>
<td>99</td>
<td>99</td>
<td>GB [NM_000769]</td>
<td>EXON 1</td>
<td>C&gt;T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2C19_276</td>
<td>CYP 2C19</td>
<td>276</td>
<td>276</td>
<td>GB [NM_000769]</td>
<td>EXON 2</td>
<td>G&gt;C</td>
<td>E 92 D</td>
<td></td>
</tr>
<tr>
<td>2C19_395</td>
<td>CYP 2C19</td>
<td>395</td>
<td>395</td>
<td>GB [NM_000769]</td>
<td>EXON 3</td>
<td>G&gt;A</td>
<td>R 132 Q</td>
<td></td>
</tr>
<tr>
<td>2C19_430</td>
<td>CYP 2C19</td>
<td>430</td>
<td>430</td>
<td>GB [NM_000769]</td>
<td>EXON 3</td>
<td>C&gt;T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2C19_636</td>
<td>CYP 2C19</td>
<td>636</td>
<td>636</td>
<td>GB [NM_000769]</td>
<td>EXON 4</td>
<td>G&gt;A</td>
<td>STOP CODON</td>
<td></td>
</tr>
<tr>
<td>2C19_681</td>
<td>CYP 2C19</td>
<td>681</td>
<td>681</td>
<td>GB [NM_000769]</td>
<td>EXON 5</td>
<td>G&gt;A</td>
<td>SPLICE DEFECT</td>
<td></td>
</tr>
</tbody>
</table>

**FIG. 3D**
<table>
<thead>
<tr>
<th>ALLELE / HAPLOTYPE FAM</th>
<th>IUPAC CODE</th>
<th>ALLELE_1</th>
<th>ALLELE_2</th>
<th>SNP_SOURCE INFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>*22</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*15</td>
<td></td>
<td></td>
<td></td>
<td>Sachse C.: Pharmacogenetics 1996</td>
</tr>
<tr>
<td>*23</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*4</td>
<td>M</td>
<td>C</td>
<td>A</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*17</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*6</td>
<td>T</td>
<td>DEL</td>
<td></td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*2</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*4</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*1</td>
<td>Y</td>
<td>T</td>
<td>C</td>
<td>Sachse C.: Pharmacogenetics 1996</td>
</tr>
<tr>
<td>*20</td>
<td>INS</td>
<td>G</td>
<td></td>
<td>Marez-Allorge D.: Pharmacogenetics 1999</td>
</tr>
<tr>
<td>*20</td>
<td>Y</td>
<td>T</td>
<td>C</td>
<td>Marez-Allorge D.: Pharmacogenetics 1999</td>
</tr>
</tbody>
</table>

**FIG. 3E**
<table>
<thead>
<tr>
<th>ALLELE / HAPLOTYP</th>
<th>IUPAC CODE</th>
<th>ALLELE_1</th>
<th>ALLELE_2</th>
<th>SNP_SOURCE INFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>*2</td>
<td>Y</td>
<td>T</td>
<td>C</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*3</td>
<td>K</td>
<td>G</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*33</td>
<td>M</td>
<td>C</td>
<td>A</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*1, *2</td>
<td>M</td>
<td>C</td>
<td>DEL</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*35</td>
<td>GACT</td>
<td>DEL</td>
<td></td>
<td>Leathart JB.: Pharmacogenetics 1998</td>
</tr>
<tr>
<td>*9</td>
<td>AGA</td>
<td>DEL</td>
<td></td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*7</td>
<td>M</td>
<td>A</td>
<td>C</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*4</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*28</td>
<td>Y</td>
<td>T</td>
<td>C</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*1</td>
<td>R</td>
<td>G</td>
<td>A</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*2</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Yokota, H.: Pharmacogenetics 1993</td>
</tr>
<tr>
<td>*2</td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>6, 8, 10, 11, 12, 14, 17, 19, 20, 28, 29, 30, 31, 32, 33</td>
<td>S</td>
<td>G</td>
<td>C</td>
<td>Marez, D.: Pharmacogenetics 1997</td>
</tr>
<tr>
<td>*7C</td>
<td>R</td>
<td>A</td>
<td>G</td>
<td>Fairbrother, KS.: Pharmacogenetics 1998</td>
</tr>
<tr>
<td>*7ALL</td>
<td>W</td>
<td>T</td>
<td>A</td>
<td>Fairbrother, KS.: Pharmacogenetics 1998</td>
</tr>
</tbody>
</table>

FIG. 3F
<table>
<thead>
<tr>
<th>ALLELE / HAPLOTYPE FAM</th>
<th>IUPAC CODE</th>
<th>ALLELE_1</th>
<th>ALLELE_2</th>
<th>SNP_SOURCE INFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>*2</td>
<td>R</td>
<td>G</td>
<td>A</td>
<td>Fairbrother, KS.: Pharmacogenetics 1998</td>
</tr>
<tr>
<td>*3</td>
<td>R</td>
<td>Y</td>
<td>C</td>
<td>HGBASE, [SNP000000637]</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>R</td>
<td>A</td>
<td>OMIM, [124010.0001]</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Y</td>
<td>C</td>
<td>Smart and Daly: Pharmacogenetics 1999</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>Y</td>
<td>C</td>
<td>Smart and Daly: Pharmacogenetics 1999</td>
</tr>
<tr>
<td>*4</td>
<td>M</td>
<td>C</td>
<td>A</td>
<td>Cacorbi, I.: Cancer Res 1996</td>
</tr>
<tr>
<td>*3</td>
<td>Y</td>
<td>Y</td>
<td>T</td>
<td>HGBASE, [SNP0000002686]</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>Y</td>
<td>C</td>
<td>GeneBank SNP database, [D12525]</td>
</tr>
<tr>
<td>*12</td>
<td>Y</td>
<td>Y</td>
<td>T</td>
<td>Bejani: Hum Mol Genetics 2000</td>
</tr>
<tr>
<td>*15</td>
<td>R</td>
<td>R</td>
<td>G</td>
<td>Bejani: Hum Mol Genetics 2000</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>R</td>
<td>G</td>
<td>Bejani: Am J. Hum Genetics 1998</td>
</tr>
</tbody>
</table>

**FIG. 3G**
<table>
<thead>
<tr>
<th>ALLELE / HAPLOTYPE FAM</th>
<th>IUPAC CODE</th>
<th>ALLELE_1</th>
<th>ALLELE_2</th>
<th>SNP_SOURCE INFO</th>
</tr>
</thead>
<tbody>
<tr>
<td>*3</td>
<td>Y</td>
<td>T</td>
<td>C</td>
<td>Stfoilov, I: Am. J. Human Genetics 1998</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>C</td>
<td>G</td>
<td>HGBASE, [SNP000000547]</td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>T</td>
<td>A</td>
<td>HGBASE, [SNP000001727], NCBI [rs2872]</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>T</td>
<td>G</td>
<td>NCBI, [rs10916]</td>
</tr>
<tr>
<td>*2</td>
<td>S</td>
<td>C</td>
<td>G</td>
<td>Huang, JD: Drug Metab Dispos 1999</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>Richardson, TH: Arch Biochem Biophys 1995</td>
</tr>
<tr>
<td>*2B</td>
<td>S</td>
<td>G</td>
<td>C</td>
<td>Ibeanu, GC: J. Pharmacol Exp Ther 1998</td>
</tr>
<tr>
<td>*6</td>
<td>R</td>
<td>G</td>
<td>A</td>
<td>Ibeanu, GC: J. Pharmacol Exp Ther 1998</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>C</td>
<td>T</td>
<td>HGBASE, [SNP00000186]</td>
</tr>
<tr>
<td>*3</td>
<td>R</td>
<td>G</td>
<td>A</td>
<td>De Morais, SM: Mol Pharmacol 1994</td>
</tr>
<tr>
<td>*2A</td>
<td>R</td>
<td>G</td>
<td>A</td>
<td>De Morais, SM: J Biol Chem 1994</td>
</tr>
</tbody>
</table>

**FIG. 3H**
### CYP2D6 and its Pseudogenes: Similarity by Region

<table>
<thead>
<tr>
<th>Length, bp</th>
<th>CYP2D6</th>
<th>CYP2D7</th>
<th>CYP2D8</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPSTREAM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXON 1</td>
<td>774</td>
<td>777</td>
<td>265</td>
</tr>
<tr>
<td>EXON 2</td>
<td>189</td>
<td>186</td>
<td>183</td>
</tr>
<tr>
<td>INTRON 1</td>
<td>268</td>
<td>269</td>
<td>265</td>
</tr>
<tr>
<td>EXON 2</td>
<td>703</td>
<td>701</td>
<td>1620*</td>
</tr>
<tr>
<td>INTRON 2</td>
<td>172</td>
<td>172</td>
<td>172</td>
</tr>
<tr>
<td>EXON 3</td>
<td>550</td>
<td>528</td>
<td>546</td>
</tr>
<tr>
<td>INTRON 3</td>
<td>153</td>
<td>153</td>
<td>153</td>
</tr>
<tr>
<td>EXON 4</td>
<td>88</td>
<td>88</td>
<td>88</td>
</tr>
<tr>
<td>EXON 5</td>
<td>161</td>
<td>161</td>
<td>161</td>
</tr>
<tr>
<td>INTRON 4</td>
<td>433</td>
<td>425</td>
<td>449</td>
</tr>
<tr>
<td>EXON 6</td>
<td>177</td>
<td>177</td>
<td>177</td>
</tr>
<tr>
<td>INTRON 5</td>
<td>190</td>
<td>192</td>
<td>186</td>
</tr>
<tr>
<td>EXON 7</td>
<td>142</td>
<td>142</td>
<td>142</td>
</tr>
<tr>
<td>INTRON 6</td>
<td>207</td>
<td>194</td>
<td>204</td>
</tr>
<tr>
<td>EXON 8</td>
<td>188</td>
<td>188</td>
<td>185</td>
</tr>
<tr>
<td>INTRON 7</td>
<td>454</td>
<td>454</td>
<td>449</td>
</tr>
<tr>
<td>EXON 8</td>
<td>142</td>
<td>142</td>
<td>142</td>
</tr>
<tr>
<td>INTRON 8</td>
<td>98</td>
<td>98</td>
<td>96</td>
</tr>
<tr>
<td>EXON 9</td>
<td>252</td>
<td>252</td>
<td></td>
</tr>
<tr>
<td>3'-FLANKING</td>
<td>180</td>
<td>160</td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>538</td>
<td>528</td>
<td>181</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Similarity, %</th>
<th>D6 / D7</th>
<th>D7 / D8</th>
<th>D6 / D8</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPSTREAM</td>
<td>97</td>
<td>92</td>
<td>89</td>
</tr>
<tr>
<td>EXON 1</td>
<td>97</td>
<td>94</td>
<td>93</td>
</tr>
<tr>
<td>EXON 2</td>
<td>98</td>
<td>90</td>
<td>89</td>
</tr>
<tr>
<td>INTRON 1</td>
<td>95</td>
<td>94</td>
<td>91</td>
</tr>
<tr>
<td>EXON 2</td>
<td>74</td>
<td>78</td>
<td>77</td>
</tr>
<tr>
<td>INTRON 2</td>
<td>98</td>
<td>93</td>
<td>92</td>
</tr>
<tr>
<td>EXON 3</td>
<td>98</td>
<td>91</td>
<td>93</td>
</tr>
<tr>
<td>INTRON 3</td>
<td>98</td>
<td>89</td>
<td>91</td>
</tr>
<tr>
<td>EXON 4</td>
<td>94</td>
<td>85</td>
<td>86</td>
</tr>
<tr>
<td>INTRON 4</td>
<td>99</td>
<td>93</td>
<td>92</td>
</tr>
<tr>
<td>EXON 5</td>
<td>97</td>
<td>84</td>
<td>83</td>
</tr>
<tr>
<td>INTRON 5</td>
<td>94</td>
<td>92</td>
<td>96</td>
</tr>
<tr>
<td>EXON 6</td>
<td>82</td>
<td>87</td>
<td>90</td>
</tr>
<tr>
<td>INTRON 6</td>
<td>98</td>
<td>94</td>
<td>95</td>
</tr>
<tr>
<td>EXON 7</td>
<td>98</td>
<td>91</td>
<td>91</td>
</tr>
<tr>
<td>INTRON 7</td>
<td>99</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>EXON 8</td>
<td>100</td>
<td>97</td>
<td>97</td>
</tr>
<tr>
<td>3'-FLANKING</td>
<td>94</td>
<td>95</td>
<td>92</td>
</tr>
</tbody>
</table>

* *3* Alu Repeats Insertion

**FIG. 4**
## P450 PRIMER LIST 02/22/01

<table>
<thead>
<tr>
<th>GENE SEQ ID</th>
<th>START POSITION</th>
<th>DIRECTION</th>
<th>SEQ_SOURCE</th>
<th>EXON / INTRON LOCATION</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP 2D6</td>
<td>1278</td>
<td>FORWARD</td>
<td>GB [M33388]</td>
<td>5' FLANK</td>
<td>CCCAGAAGGCTTTGCAGGCTTCA</td>
</tr>
<tr>
<td>CYP 2D6</td>
<td>5799</td>
<td>REVERSE</td>
<td>GB [M33388]</td>
<td>EXON 9</td>
<td>CTCACCAGGAAAGCAAGACACCAT</td>
</tr>
<tr>
<td>CYP 2E1</td>
<td>1487</td>
<td>FORWARD</td>
<td>GB [J02843]</td>
<td>5' FLANK</td>
<td>TGAAGCCTCTGACCAGGCAA</td>
</tr>
<tr>
<td>CYP 2E1</td>
<td>4022</td>
<td>REVERSE</td>
<td>GB [J02843]</td>
<td>EXON 2</td>
<td>GAGAACTCTGTCTTGTAGTCAGCA</td>
</tr>
<tr>
<td>CYP 2E1</td>
<td>7445</td>
<td>FORWARD</td>
<td>GB [J02843]</td>
<td>INTRON 3</td>
<td>CACCTTCTACAGGCTTGTTGAA</td>
</tr>
<tr>
<td>CYP 2E1</td>
<td>10689</td>
<td>REVERSE</td>
<td>GB [J02843]</td>
<td>INTRON 6</td>
<td>CACCCCTGCTCTCTGAGGTGAA</td>
</tr>
<tr>
<td>CYP 2E1</td>
<td>12667</td>
<td>FORWARD</td>
<td>GB [J02843]</td>
<td>INTRON 8</td>
<td>CCTGACCCCTGACTGCTTTCTATCTAA</td>
</tr>
<tr>
<td>CYP 2E1</td>
<td>13030</td>
<td>REVERSE</td>
<td>GB [J02843]</td>
<td>INTRON 8</td>
<td>AGGAGTGTGGGCTGCTCCTCAA</td>
</tr>
<tr>
<td>CYP 3A4</td>
<td>747</td>
<td>FORWARD</td>
<td>GB [D11131]</td>
<td>5' FLANK</td>
<td>CTGTAGGTGTGCTTGTGGATGAA</td>
</tr>
<tr>
<td>CYP 3A4</td>
<td>941</td>
<td>REVERSE</td>
<td>GB [D11131]</td>
<td>5' FLANK</td>
<td>CACACCACTCACTGACCTCTTGTGA</td>
</tr>
<tr>
<td>CYP 1A1</td>
<td>7080</td>
<td>FORWARD</td>
<td>GB [X04300]</td>
<td>EXON 7</td>
<td>CTCTGTTACAGGAAGCTATGGGCTCAA</td>
</tr>
<tr>
<td>CYP 1A1</td>
<td>8206</td>
<td>REVERSE</td>
<td>GB [X04300]</td>
<td>3' FLANK</td>
<td>CATGCAAGCTCAATGCGAGCTAGAATAGAA</td>
</tr>
<tr>
<td>CYP 1B1</td>
<td>3698</td>
<td>FORWARD</td>
<td>GB [U56438]</td>
<td>INTRON 1</td>
<td>CTTCACCACCAACGGCACCTCA</td>
</tr>
<tr>
<td>CYP 1B1</td>
<td>4704</td>
<td>REVERSE</td>
<td>GB [U56438]</td>
<td>EXON 2</td>
<td>GCAGAGAGAGGATAAAGCGTCATCA</td>
</tr>
<tr>
<td>CYP 1B1</td>
<td>7583</td>
<td>FORWARD</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>ATCCCTGATGTGCAAGACTGAGTGCA</td>
</tr>
<tr>
<td>CYP 1B1</td>
<td>9235</td>
<td>REVERSE</td>
<td>GB [U56438]</td>
<td>EXON 3</td>
<td>CAGCCAGGAGGCTGACATCAGCTGCA</td>
</tr>
<tr>
<td>CYP 1A2</td>
<td>2595</td>
<td>FORWARD</td>
<td>GB [M31664]</td>
<td>5' FLANK</td>
<td>CCAGCTCTCAGATTCTGTGATGCTCAA</td>
</tr>
<tr>
<td>CYP 1A2</td>
<td>2872</td>
<td>REVERSE</td>
<td>GB [M31664]</td>
<td>EXON 1</td>
<td>GGGTGACGACATGCGCGAGCAA</td>
</tr>
<tr>
<td>CYP 2C19</td>
<td>197</td>
<td>FORWARD</td>
<td>GB [NM 000769]</td>
<td>EXON 2</td>
<td>CTCTGTATTTTGCGCTGGAAACGCGATG</td>
</tr>
<tr>
<td>CYP 2C19</td>
<td>787</td>
<td>REVERSE</td>
<td>GB [NM 000769]</td>
<td>EXON 5</td>
<td>TCCCGAGGGCTTGTATGTCCATC</td>
</tr>
<tr>
<td>P450 PRIMER LIST 02/22/01</td>
<td>BLAST RESULTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PRIMER LENGTH</strong></td>
<td>23</td>
<td>25</td>
<td>22</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td><strong>Tm</strong></td>
<td>71</td>
<td>67</td>
<td>71</td>
<td>65</td>
<td>69</td>
</tr>
<tr>
<td><strong>ACCEPTABLE</strong></td>
<td>ACCEPTABLE</td>
<td>ACCEPTABLE</td>
<td>ACCEPTABLE</td>
<td>ACCEPTABLE</td>
<td>ACCEPTABLE</td>
</tr>
</tbody>
</table>
### P450 COMPILED PROBE LIST/ BETA

<table>
<thead>
<tr>
<th>PROBE ID</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP1A1 V 2.70.6568.A.S</td>
<td>GCAAGCGGAAGTGATCGGTGAGAA</td>
</tr>
<tr>
<td>CYP1A1 V 2.70.6568.C.S</td>
<td>GCAAGCGGAAGTGATCGGTGAGAC</td>
</tr>
<tr>
<td>CYP1A1 60.6570.A.A</td>
<td>TCCACGCGGCGAAT</td>
</tr>
<tr>
<td>CYP1A1 60.6570.G.A</td>
<td>TCCACGCGGCGAAC</td>
</tr>
<tr>
<td>CYP1A1 V 2.60.7320.C.A</td>
<td>ATAGGCTTCAAGCGCAG</td>
</tr>
<tr>
<td>CYP1A1 V 2.60.7320.T.A</td>
<td>AATAAGGCTTCAAGCGCAG</td>
</tr>
<tr>
<td>CYP1A2 60+1.2640.A.A</td>
<td>CATCTACATGGCTCTGCTG</td>
</tr>
<tr>
<td>CYP1A2 60+1.2640.C.A</td>
<td>ATCTACATGGCTCTGCTG</td>
</tr>
<tr>
<td>CYP1A2 V2.2866.C.S</td>
<td>TGGCCTCTGCCATCTTCT</td>
</tr>
<tr>
<td>CYP1A2 V2.2866.G.S</td>
<td>TGCCCCGCTGCAATCTT</td>
</tr>
<tr>
<td>CYP1A2 V3.2866.C.S</td>
<td>TGGCCTCTGCCATCTTCT</td>
</tr>
<tr>
<td>CYP1A2 V3.2866.G.S</td>
<td>TGCCCCGCTGCAATCTT</td>
</tr>
<tr>
<td>CYP1B1 60.3793.C.A</td>
<td>CCATGCTGAGGACAGAG</td>
</tr>
<tr>
<td>CYP1B1 60.3793.T.A</td>
<td>CCATGCTGAGGACAGAA</td>
</tr>
<tr>
<td>CYP1B1 60.3947.C.S</td>
<td>GACGAGGCAGCTCC</td>
</tr>
<tr>
<td>CYP1B1 60.3947.G.S</td>
<td>GACGAGGCAGCTCG</td>
</tr>
<tr>
<td>CYP1B1 60.3976.C.S</td>
<td>GCCCGTTGCGTTCG</td>
</tr>
<tr>
<td>CYP1B1 60.3976.G.S</td>
<td>GCCCGTTGCGTTC</td>
</tr>
<tr>
<td>CYP1B1 60.3987.A.A</td>
<td>GCCCGCGGCATT</td>
</tr>
<tr>
<td>CYP1B1 60.3987.G.A</td>
<td>GCCCGCGGCATT</td>
</tr>
<tr>
<td>1B1 4035.C.S</td>
<td>GTTCGCTCGCG</td>
</tr>
<tr>
<td>1B1 4035.T.S</td>
<td>CTCGTTGCTGCTG</td>
</tr>
<tr>
<td>CYP1B1 60+2.14160.G.A</td>
<td>GAAAGAGGAGAAGGAC</td>
</tr>
<tr>
<td>CYP1B1 60+2.14160.T.A</td>
<td>GAAAGAGGAGAAGGAC</td>
</tr>
<tr>
<td>1B1 V2.4306.A.A</td>
<td>TCAGCACTGACCCCT</td>
</tr>
<tr>
<td>1B1 V2.4306.T.A</td>
<td>CAGCACTGACCCAG</td>
</tr>
<tr>
<td>CYP1B1 60+1.4646.G.S</td>
<td>AGTTCTTGAAGCAGTTCGA</td>
</tr>
<tr>
<td>CYP1B1 60+1.4646.T.S</td>
<td>CAAGTTCTTGAAGCAGTTC</td>
</tr>
<tr>
<td>1B1 4668.C.A</td>
<td>TCACGCGGCGGG</td>
</tr>
<tr>
<td>1B1 4668.G.A</td>
<td>TCACGCGGCGG</td>
</tr>
<tr>
<td>CYP1B1 60.7930.G.S</td>
<td>GAAATTGAGTACAGTGCTG</td>
</tr>
<tr>
<td>CYP1B1 60.7930.T.S</td>
<td>GAAATTGAGTACAGTGCTG</td>
</tr>
<tr>
<td>1B1 V2.7940.A.A</td>
<td>TGGTIAACCAATACAGAG</td>
</tr>
<tr>
<td>1B1 V2.7940.G.A</td>
<td>GGTIAACCAATACAGAG</td>
</tr>
<tr>
<td>CYP1B1 60+1.7957.A.S</td>
<td>CGTCTGCTCTTTGATGTGAA</td>
</tr>
<tr>
<td>CYP1B1 60+1.7957.G.S</td>
<td>CGTCTGCTCTTTGATGT</td>
</tr>
<tr>
<td>CYP1B1 60.7973.C.A</td>
<td>GGAAGCCAGGACACAG</td>
</tr>
<tr>
<td>CYP1B1 60.7973.T.A</td>
<td>GGAAGCCAGGACACAG</td>
</tr>
<tr>
<td>CYP1B1 60.7996.A.S</td>
<td>TATGTCCCTGCCGCTTATA</td>
</tr>
<tr>
<td>CYP1B1 60.7996.G.S</td>
<td>GTTCCTGCTCTTCTTAT</td>
</tr>
<tr>
<td>CYP1B1 60+1.8131.C.S</td>
<td>GTCTGTGATCATGACCGCCAGT</td>
</tr>
<tr>
<td>CYP1B1 60+1.8131.G.S</td>
<td>GTCTGTGATCATGACCGCCAGT</td>
</tr>
<tr>
<td>CYP1B1 60+1.8184.C.A</td>
<td>GTCTGTGATCATGACCGCCAGT</td>
</tr>
<tr>
<td>CYP1B1 60+1.8184.T.A</td>
<td>GTCTGTGATCATGACCGCCAGT</td>
</tr>
<tr>
<td>CYP1B1 60.8195.A.A</td>
<td>TGCTGTGATCATGACCGCCAGT</td>
</tr>
<tr>
<td>CYP1B1 60.8195.G.A</td>
<td>GCTGTGATCATGACCGCCAGT</td>
</tr>
<tr>
<td>CYP1B1 60.8242.C.S</td>
<td>TTTCAGTGGGCAAAAGG</td>
</tr>
<tr>
<td>CYP1B1 60.8242.T.S</td>
<td>TTTCAGTGGGCAAAAGG</td>
</tr>
</tbody>
</table>

**FIG. 6A**
<table>
<thead>
<tr>
<th>PROBE ID</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP1B1 60.8587.G.S</td>
<td>TCAATTAGCGTTAAGGTGAAGC</td>
</tr>
<tr>
<td>CYP1B1 60.8587.G.S</td>
<td>TCAATTAGCGTTAAGGTGAAGC</td>
</tr>
<tr>
<td>CYP1B1 60.8807.A.S</td>
<td>CCAAAACTTACACCAAACA</td>
</tr>
<tr>
<td>CYP2B1 60.8807.T.S</td>
<td>ACCCAAAACTTACACCAAACA</td>
</tr>
<tr>
<td>CYP2B1 60-1.9164.G.S</td>
<td>GAGTATAGTGGGGTCTCATGATT</td>
</tr>
<tr>
<td>CYP2B1 60-1.9164.T.S</td>
<td>GAGTATAGTGGGGTCTCATGATT</td>
</tr>
<tr>
<td>CYP2C19 EXONS 70.276.C.A</td>
<td>GAACTGGCCCTCTTCCAGAAAACCTCG</td>
</tr>
<tr>
<td>CYP2C19 EXONS 70.276.G.A</td>
<td>GAGAAATGGCCTCTTCCAGAAAACCTCC</td>
</tr>
<tr>
<td>CYP2C19 EXONS 70.395.A.A</td>
<td>CTCTCTTCCACATCCAAAATCTCT</td>
</tr>
<tr>
<td>CYP2C19 EXONS 70.395.G.A</td>
<td>CTCTCTTCCACATCCAAAATCTCT</td>
</tr>
<tr>
<td>CYP2C19 EXONS 70.430.C.A</td>
<td>GCCGGCTTCTTCTTGAAACGC</td>
</tr>
<tr>
<td>CYP2C19 EXONS 70.430.T.A</td>
<td>AGCCGGCTTCTTCTTGAAACGC</td>
</tr>
<tr>
<td>CYP2C19 EXONS 60.636.A.S</td>
<td>GATTGTAAGCCACCCCTGGA</td>
</tr>
<tr>
<td>CYP2C19 EXONS 60.636.G.S</td>
<td>TTGTAAACACCCCTTGGG</td>
</tr>
<tr>
<td>CYP2C19 EXONS 60.681.A.S</td>
<td>CCAACTACATGTAGAATTTCCTCCA</td>
</tr>
<tr>
<td>CYP2C19 EXONS 60.681.G.S</td>
<td>CCAACTACATGTAGAATTTCCTCCA</td>
</tr>
<tr>
<td>CYP2D6 70.1638.A.S</td>
<td>AGCCAGTGTGGGGCTAAGGACTGA</td>
</tr>
<tr>
<td>CYP2D6 70.1638.G.S</td>
<td>AGCCAGTGTGGGGCTAAGGACTGA</td>
</tr>
<tr>
<td>CYP2D6 70.1650.A.A</td>
<td>AGGCAGCAGGAGAGTGCCACATCAT</td>
</tr>
<tr>
<td>CYP2D6 70.1650.G.A</td>
<td>AGGCAGCAGGAGAGTGCCACATCAT</td>
</tr>
<tr>
<td>CYP2D6 70.1696.G.S</td>
<td>GGCGTGGACCGTCCG</td>
</tr>
<tr>
<td>CYP2D6 70.1696.G.S</td>
<td>GGCGTGGACCGTCCG</td>
</tr>
<tr>
<td>CYP2D6 70.1696.C.A</td>
<td>TGGTGAGGGGTGACCCCGCC</td>
</tr>
<tr>
<td>CYP2D6 70.1701.C.A</td>
<td>TGGTGAGGGGTGACCCCGCC</td>
</tr>
<tr>
<td>CYP2D6 70.1701.T.A</td>
<td>GTGGTACGGTGCAGGCCAGCA</td>
</tr>
<tr>
<td>CYP2D6 60.1719.C.A</td>
<td>GGGGCGCTTGGTG</td>
</tr>
<tr>
<td>CYP2D6 60.1719.T.A</td>
<td>AGGGGGCTCTTGTA</td>
</tr>
<tr>
<td>CYP2D6 70-1743.A.S</td>
<td>CCCCCGTCCACTTCCCA</td>
</tr>
<tr>
<td>CYP2D6 70-1743.G.S</td>
<td>CCCCCGTCCACTTCCCA</td>
</tr>
<tr>
<td>2D6H1757.G.S</td>
<td>CCGCTGGACACTGGCCG</td>
</tr>
<tr>
<td>2D6H1757.T.S</td>
<td>CCGCTGGACACTGGCCG</td>
</tr>
<tr>
<td>2D6H21757.G.S</td>
<td>CTGCCACTTGCCGACCTTCT</td>
</tr>
<tr>
<td>2D6H21757.T.S</td>
<td>CTGCCACTTGCCGACCTTCT</td>
</tr>
<tr>
<td>CYP2D6 60-1.2502.C.A</td>
<td>CGGCACGCGCAAGT</td>
</tr>
<tr>
<td>CYP2D6 60-1.2502.G.A</td>
<td>CGGCACGCGCAAGT</td>
</tr>
<tr>
<td>CYP2D6 60-1.2502.C.S</td>
<td>TGACCTTCCCTCTGCAGT</td>
</tr>
<tr>
<td>CYP2D6 60-1.2502.G.S</td>
<td>TGACCTTCCCTCTGCAGT</td>
</tr>
<tr>
<td>CYP2D6 60.2576.C.S</td>
<td>GCTCAATGGGGCTGG</td>
</tr>
<tr>
<td>CYP2D6 60.2576.T.S</td>
<td>GTCTCATAGGGGCTGGGT</td>
</tr>
<tr>
<td>CYP2D6 60.2593.A.A</td>
<td>CGCCCGGCTGCACTCAT</td>
</tr>
<tr>
<td>CYP2D6 60.2593.C.A</td>
<td>CGCCCGGCTGCACTCAT</td>
</tr>
<tr>
<td>CYP2D6 70+.2603.A.S</td>
<td>GAGGCGTGGTGAGCCACCCG</td>
</tr>
<tr>
<td>CYP2D6 70+.2603.G.S</td>
<td>GAGGCGTGGTGAGCCACCCG</td>
</tr>
<tr>
<td>CYP2D6 60+2-1.2616.C.A</td>
<td>GCGGTGCGGGCCGT</td>
</tr>
<tr>
<td>CYP2D6 60+2-1.2616.G.A</td>
<td>GCGGTGCGGGCCGT</td>
</tr>
<tr>
<td>CYP2D6 60+2-1.2616.C.S</td>
<td>GCGGTGCGGGCCGT</td>
</tr>
<tr>
<td>CYP2D6 60+2-1.2616.T.S</td>
<td>GCGGTGCGGGCCGT</td>
</tr>
<tr>
<td>CYP2D6 60+2.2642.C.A</td>
<td>CCIAAAACCCAGGATCTGGGT</td>
</tr>
<tr>
<td>CYP2D6 60+2.2642.T.A</td>
<td>CCIAAAACCCAGGATCTGGGT</td>
</tr>
</tbody>
</table>

**FIG. 6B**
<table>
<thead>
<tr>
<th>PROBE ID</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2D6 70+1.2658.C.A</td>
<td>TGGGAACGGCGGCCCAGA</td>
</tr>
<tr>
<td>CYP2D6 70+1.2658.T.A</td>
<td>TGGGAACGGCGGCCCCAA</td>
</tr>
<tr>
<td>CYP2D6 60+1.3278.A.S</td>
<td>CAGAGGCGCTCTTCTCAT</td>
</tr>
<tr>
<td>CYP2D6 60+1.3278.G.S</td>
<td>CAGAGGCGCTCTTCTCGGT</td>
</tr>
<tr>
<td>CYP2D6 60.3280.C.S</td>
<td>CAGAGGCGCTCTTCTCCITC</td>
</tr>
<tr>
<td>CYP2D6 60.3280.G.S</td>
<td>CAGAGGCGCTCTTCTCCITG</td>
</tr>
<tr>
<td>CYP2D6 70+1.3323.C.S</td>
<td>TGGGCAAGAAGTGCCTGGAGGA</td>
</tr>
<tr>
<td>CYP2D6 70+1.3323.G.S</td>
<td>TGGGCAAGAAGTGCCTGGAGGA</td>
</tr>
<tr>
<td>2D6H.3326.G.A</td>
<td>GCIGCCTCTCCGTGGCTACCCC</td>
</tr>
<tr>
<td>2D6H.3326.T.A</td>
<td>GCIGCCTCTCCGTGGCTACCCA</td>
</tr>
<tr>
<td>2D6H_V2.3326.G.A</td>
<td>GCIGCCTCTCCGTGGCTACCCCT</td>
</tr>
<tr>
<td>2D6H_V2.3326.T.A</td>
<td>GCIGCCTCTCCGTGGCTACCCCAC</td>
</tr>
<tr>
<td>CYP2D6 70.3343.C.A</td>
<td>CGGCACAAAAGGCAGGCAGG</td>
</tr>
<tr>
<td>CYP2D6 70.3343.T.A</td>
<td>CGGCACAAAAGGCAGGCAGA</td>
</tr>
<tr>
<td>CYP2D6 70.3368.A.S</td>
<td>TGTCGCCGCTGTGGACCA</td>
</tr>
<tr>
<td>CYP2D6 70.3368.G.S</td>
<td>GTGCAGCGCTGTCGCAGG</td>
</tr>
<tr>
<td>CYP2D6 70.3377.G.S</td>
<td>CGCCTTCGCGCAACACTCCG</td>
</tr>
<tr>
<td>CYP2D6 70.3377.T.S</td>
<td>CGCCTTCGCGCAACACTCCT</td>
</tr>
<tr>
<td>CYP2D6 60.3465.A.S</td>
<td>CATCTCCACCACCCCAC</td>
</tr>
<tr>
<td>CYP2D6 60.3465.G.S</td>
<td>CATCTCCACCACCCCAG</td>
</tr>
<tr>
<td>CYP2D6 60.3477.C.A</td>
<td>AGAIGCCGTTGGGCCG</td>
</tr>
<tr>
<td>CYP2D6 60.3477.T.A</td>
<td>AGAIGCCGTTGGGCCGA</td>
</tr>
<tr>
<td>CYP2D6 60.3488.C.A</td>
<td>CGGCTTGTCCAAAGGGA</td>
</tr>
<tr>
<td>CYP2D6 60.3488.T.A</td>
<td>AGGCTTGTCCAAAGGAGA</td>
</tr>
<tr>
<td>CYP2D6 60.3562.A.A</td>
<td>CCAGCAGCCTGAGGAAGT</td>
</tr>
<tr>
<td>CYP2D6 60.3562.G.A</td>
<td>CCAGCAGCCTGAGGAAGC</td>
</tr>
<tr>
<td>2D6H.3592.A.S</td>
<td>TCAAGCTGCTGAGCTTAGCTCAGGA</td>
</tr>
<tr>
<td>2D6H.3592.G.S</td>
<td>TCAAGCTGCTGAGCTTAGCTCAGGGA</td>
</tr>
<tr>
<td>CYP2D6 60.3595.A.S</td>
<td>TGACCTAGCTCAGAGGA</td>
</tr>
<tr>
<td>CYP2D6 60.3595.G.S</td>
<td>GGACCTAGCTCAGAGGG</td>
</tr>
<tr>
<td>CYP2D6 70.3597.C.S</td>
<td>GCTGCTGAGCTTACGTAGCAGGAGGIACTG</td>
</tr>
<tr>
<td>CYP2D6 70.3597.T.S</td>
<td>GCTGCTGAGCTTACGTAGCAGGAGGIAT</td>
</tr>
<tr>
<td>CYP2D6 60.3598.C.A</td>
<td>CCCGAACCTTCCTCTCAGG</td>
</tr>
<tr>
<td>CYP2D6 60.3598.T.A</td>
<td>CCCGAACCTTCCTCTCAGC</td>
</tr>
<tr>
<td>CYP2D6 70.4099.C.S</td>
<td>TCCTCTCTGCAATCCCAAGGC</td>
</tr>
<tr>
<td>CYP2D6 70.4099.T.S</td>
<td>TCCTCTCTGCAATCCCAAGGT</td>
</tr>
<tr>
<td>2D6H_V2.4168.A.S</td>
<td>CTGGATGAGCTCTGTAAGTACGGACACGG</td>
</tr>
<tr>
<td>2D6H_V2.4168.G.S</td>
<td>CTGGATGAGCTCTGTAAGTACGGACCGG</td>
</tr>
<tr>
<td>2D6H_V3.4168.A.S</td>
<td>GCTGAGTGAGCTCTAAGTACGGACACA</td>
</tr>
<tr>
<td>2D6H_V3.4168.G.S</td>
<td>GCTGAGTGAGCTCTAAGTACGGACCGG</td>
</tr>
<tr>
<td>CYP2D6 70.4194.A.S</td>
<td>GGGACCCAGCCCGACCA</td>
</tr>
<tr>
<td>CYP2D6 70.4194.C.S</td>
<td>GGGACCCAGCCCGGCC</td>
</tr>
<tr>
<td>2D6H.4206.G.S</td>
<td>GCCCAGCICCAGGGACCTGAGG</td>
</tr>
<tr>
<td>2D6H.4206.T.S</td>
<td>GCCCAGCICCAGGGACCTGACT</td>
</tr>
<tr>
<td>2D6H.4232.A.A</td>
<td>TGGGAGCGACTCTCACCTCCT</td>
</tr>
<tr>
<td>2D6H.4232.G.A</td>
<td>TGGGAGCGACTCTCACCTCCT</td>
</tr>
<tr>
<td>CYP2D6 60.4469.C.S</td>
<td>GCTTTCAATGATGAGAACCTGCG</td>
</tr>
<tr>
<td>CYP2D6 60.4469.T.S</td>
<td>GCTTTCAATGATGAGAACCTGCT</td>
</tr>
</tbody>
</table>

FIG_6C
<table>
<thead>
<tr>
<th>PROBE ID</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2D6 70.4472.A.A</td>
<td>GCAGAGAACAGGTGCAGCCACACACTAT</td>
</tr>
<tr>
<td>CYP2D6 70.4472.C.A</td>
<td>GCAGAGAACAGGTGCAGCCACACACTAG</td>
</tr>
<tr>
<td>CYP2D6  V3.4554.C.A</td>
<td>TGGGCTACGCCTGCACATCTGGG</td>
</tr>
<tr>
<td>CYP2D6  V3.4554.A.A</td>
<td>GCTACGCTGACTGCATGACATCTGGG</td>
</tr>
<tr>
<td>CYP2D6  V2.4554.C.A</td>
<td>GCTACGCTGACTGCATGACATCTGGG</td>
</tr>
<tr>
<td>CYP2D6  V2.4554.A.A</td>
<td>GCTACGCTGACTGCATGACATCTGGG</td>
</tr>
<tr>
<td>CYP2D6  70.4557.C.S</td>
<td>GGCCCTCTGGTCAACACACATCCTACTCTC</td>
</tr>
<tr>
<td>CYP2D6  70.4557.T.S</td>
<td>GGCCCTCTGGTCAACACACATCCTACTCTC</td>
</tr>
<tr>
<td>CYP2D6  70.4558.A.A</td>
<td>TGCCCTACGCTGCACACTCT</td>
</tr>
<tr>
<td>CYP2D6  70.4558.G.A</td>
<td>GGCTGCAGCTGCACTCCCCCA</td>
</tr>
<tr>
<td>CYP2D6  70.4802.A.S</td>
<td>GTGCACACAGAGTGGGTCAAGCGAAGG</td>
</tr>
<tr>
<td>CYP2D6  70.4802.G.S</td>
<td>TGCCACACAGAGTGGGTCAAGCGAAGG</td>
</tr>
<tr>
<td>CYP2D6  70.1+4517.C.S</td>
<td>TCCAGACTGATAGGCGAGGTGCGG</td>
</tr>
<tr>
<td>CYP2D6  70.1+4517.G.S</td>
<td>ATGCAGCTGATAGGCGAGGTGCGG</td>
</tr>
<tr>
<td>CYP2D6  70.4896.C.S</td>
<td>TGCAAGCCTTGTTGGGACAC</td>
</tr>
<tr>
<td>CYP2D6  70.4896.T.S</td>
<td>GTGACAGCTTGTTGGGACAC</td>
</tr>
<tr>
<td>CYP2D6  60.4907.A.S</td>
<td>GGCAACAGGCTCCCCCCTGA</td>
</tr>
<tr>
<td>CYP2D6  60.4907.G.S</td>
<td>GGCAACAGGCTCCCCCCTGA</td>
</tr>
<tr>
<td>CYP2D6  70.5447.A.A</td>
<td>AGACGCGGCTATCTCCTTCAGCACT</td>
</tr>
<tr>
<td>CYP2D6  70.5447.G.A</td>
<td>AGACGCGGCTATCTCCTTCAGCACT</td>
</tr>
<tr>
<td>CYP2D6  70.5472.A.S</td>
<td>CTGAGGACACACGGGCAGTGCTGGA</td>
</tr>
<tr>
<td>CYP2D6  70.5472.G.S</td>
<td>CTGAGGACACACGGGCAGTGCTGGA</td>
</tr>
<tr>
<td>CYP2D6  70.5496.C.S</td>
<td>CTTCCGGCTCCACCC</td>
</tr>
<tr>
<td>CYP2D6  70.5496.G.S</td>
<td>CTTCCGGCTCCACCC</td>
</tr>
<tr>
<td>CYP2D6  70.1+5506.C.S</td>
<td>CGTTCGACCCCAACACTCTCCCG</td>
</tr>
<tr>
<td>CYP2D6  70.1+5506.T.S</td>
<td>CGTTCGACCCCAACACTCTCCCG</td>
</tr>
<tr>
<td>CYP2D6  70.1+5561.A.S</td>
<td>CCCCTCCCTCCACGGGCAC</td>
</tr>
<tr>
<td>CYP2D6  70.1+5561.G.S</td>
<td>CCCCTCCCTCCACGGGCAC</td>
</tr>
<tr>
<td>CYP2D6  70.5734.C.A</td>
<td>CAGTGGGGCACAGGAGAAGCTG</td>
</tr>
<tr>
<td>CYP2D6  70.5734.T.A</td>
<td>CAGTGGGGCACAGGAGAAGCTG</td>
</tr>
<tr>
<td>CYP2E1  60.1+1.1532.C.A</td>
<td>CTGACCTAAACACTTCAG</td>
</tr>
<tr>
<td>CYP2E1  60.1+1.1532.G.A</td>
<td>CTGACCTAAACACTTCAG</td>
</tr>
<tr>
<td>CYP2E1  60.1627.A.A</td>
<td>CATTCTATAACTTGCATTAATTTAACAAAAATGAGAG</td>
</tr>
<tr>
<td>CYP2E1  60.1627.G.A</td>
<td>CATTCTATAACTTGCATTAATTTAACAAAAATGAGAG</td>
</tr>
<tr>
<td>CYP2E1  V2.1772.C.A</td>
<td>TCTTAAATTCTAGGTCAATTGGAATTTTGT</td>
</tr>
<tr>
<td>CYP2E1  V2.1772.T.A</td>
<td>TCTTAAATTCTAGGTCAATTGGAATTTTGT</td>
</tr>
<tr>
<td>CYP2E1  V3.1800.T.S</td>
<td>TTGGAACCTAGTAAAGGTAAGACTCTTA</td>
</tr>
<tr>
<td>CYP2E1  V3.1800.C.S</td>
<td>ATTGCAACCTTAGGTAAGAAGACTCTTCC</td>
</tr>
<tr>
<td>CYP2E1  60.1+2019.C.A</td>
<td>GATTGTTTTTACATTAGGTGAATTGTTGG</td>
</tr>
<tr>
<td>CYP2E1  60.1+2019.T.A</td>
<td>GATTGTTTTTACATTAGGTGAATTGTTGG</td>
</tr>
<tr>
<td>2E1 2492.A.A</td>
<td>GTGGGGGTGAGTACCGT</td>
</tr>
<tr>
<td>2E1 2492.T.A</td>
<td>GTGGGGGTGAGTACCGT</td>
</tr>
<tr>
<td>2E1 2492.A.S</td>
<td>TGCCAAAGGGCACGA</td>
</tr>
<tr>
<td>2E1 2492.T.S</td>
<td>TGCCAAAGGGCACGA</td>
</tr>
<tr>
<td>2E1 2473.A.A</td>
<td>GCCCTTGTCGACTGTG</td>
</tr>
<tr>
<td>2E1 2473.G.A</td>
<td>GCCCTTGTCGACTGTG</td>
</tr>
<tr>
<td>2E1 2473.A.S</td>
<td>GGAGTTCCCCGTTGTCTA</td>
</tr>
<tr>
<td>2E1 2473.G.S</td>
<td>GGAGTTCCCCGTTGTCTA</td>
</tr>
</tbody>
</table>

**FIG. 6D**
<table>
<thead>
<tr>
<th>PROBE ID</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYP2E1</td>
<td>GGGTCACCCCTCCTTCTCAG</td>
</tr>
<tr>
<td>CYP2E1</td>
<td>TGGGCTCGCAGGCGCA</td>
</tr>
<tr>
<td>CYP2E1.V2.3956.A.A</td>
<td>CGGTGACATCACACCATGT</td>
</tr>
<tr>
<td>CYP2E1.V2.3956.A.GA</td>
<td>GTGCATACACACCATGCG</td>
</tr>
<tr>
<td>CYP2E1.60.10456.A.S</td>
<td>CACACCCAGCTGATTTAAAAATTA</td>
</tr>
<tr>
<td>CYP2E1.60.10456.T.S</td>
<td>CACACCCAGCTGATTTAAAAATTT</td>
</tr>
<tr>
<td>CYP2E1.60.12720.C.S</td>
<td>TACACTAAGCAACTCTTTACTCTC</td>
</tr>
<tr>
<td>CYP2E1.60.12720.C.A</td>
<td>TACACTAAGCAACTCTTTACTCTG</td>
</tr>
<tr>
<td>CYP2E1.60.12847.A.S</td>
<td>TTTCCTCCGACACAGTCG</td>
</tr>
<tr>
<td>CYP2E1.60.12847.G.S</td>
<td>TTTCCTCCGACACAGTCG</td>
</tr>
<tr>
<td>CYP2E1.60.12945.C.A</td>
<td>GGCTTGAATAATGTCATGTACTTG</td>
</tr>
<tr>
<td>CYP2E1.60.12945.T.A</td>
<td>AATGGCTTAATAGTCATGTACTTA</td>
</tr>
<tr>
<td>CYP3A4.60.816.A.S</td>
<td>GCCATAGAGCACAGGGGCAA</td>
</tr>
<tr>
<td>CYP3A4.60.816.G.S</td>
<td>GCCATAGAGCACAGGGGCA</td>
</tr>
<tr>
<td>CYP3A4.60.918.A.S</td>
<td>CCACTACATTGATAGTTGTTTATTA</td>
</tr>
<tr>
<td>CYP3A4.60.918.G.S</td>
<td>CAGTAACATTGATAGTTGTTTG</td>
</tr>
</tbody>
</table>

**Amplicon Control Probes**

- 1A1.23F22R.A.X.A: GCAGGATCCCTTAGGCTTG
- 1A1.23F22R.B.X.S: AGCCAGGAGGCTCCTGA
- 1A2.5F3R.A.X.S: TATCCAGCTGGGAGCCAA
- 1A2.5F3R.B.X.S: CCAGCAGCCATGCTCT
- 1B1.2F4R.A.X.S: CACGACGGACCCCGAGTT
- 1B1.2F4R.B.X.S: CGGTGCCAACCGT
- 1B1.8F11R.A.X.A: TTGGGTGGGCTCGTGGAA
- 1B1.8F11R.B.X.S: TGGGCTAATGACCGAGCTT
- 2C19.3F6R.A.X.A: GCACAGCCGAGGATGAA
- 2C19.3F6R.B.X.A: CATGCAAGCACCACCATG
- 2D6.1F1R.A.X.S: AGCCCATTTGTGAGGAGGCGAG
- 2D6.1F1R.B.X.S: GACTCCATTGTGAGGAGGCGAG
- 2E1.1F6R.A.X.A: AGGTGTTATTGGAACAAACACCAA
- 2E1.1F6R.B.X.A: ATTCAGGTAATTCACACAGAGC
- 2E1.8F19R.A.X.S: GACTGTGGGCGACTCCTG
- 2E1.8F19R.B.X.S: GCACAGTGCGAGCGCTTT
- 2E1.11F13R.A.X.S: CCAGATGAAAGGGCCACATT
- 2E1.11F13R.B.X.S: AAGCGCCACATTTGTAACATG
- 3A4.1F1R.A.X.S: GCTTTTGGGATGAATTCTC
- 3A4.1F1R.B.X.S: CTGATAAGAACCAGAAACCCCTT

**FIG. 6E**
<table>
<thead>
<tr>
<th>1A1</th>
<th>CYP</th>
<th>SNP ID</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>NONE</td>
<td>CYP1A1</td>
<td>1213</td>
<td>CYP1A1_V.2_60.1213.A.S</td>
</tr>
<tr>
<td>NONE</td>
<td>CYP1A1</td>
<td>1223</td>
<td>CYP1A1_V.2_60.1223.C.A</td>
</tr>
<tr>
<td>23F22R</td>
<td>CYP1A1</td>
<td>6568</td>
<td>CYP1A1_V.2_70.6568.A.S</td>
</tr>
<tr>
<td>23F22R</td>
<td>CYP1A1</td>
<td>6570</td>
<td>CYP1A1_60.6570.A.A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CYP1A1_V.2_70.6570.A.A</td>
</tr>
<tr>
<td>23F22R</td>
<td>CYP1A1</td>
<td>7320</td>
<td>CYP1A1_V.2_60.7320.C.A</td>
</tr>
<tr>
<td>NONE</td>
<td>CYP1A1</td>
<td>7547</td>
<td>CYP1A1_60+1.7547.A.S</td>
</tr>
<tr>
<td>23F22R</td>
<td>CYP1A1</td>
<td>CNTRL</td>
<td>1A1.23F22R_A.X.A</td>
</tr>
<tr>
<td>23F22R</td>
<td>CYP1A1</td>
<td>CNTRL</td>
<td>1A1.23F22R_B.X.S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1A2</th>
<th>CYP</th>
<th>SNP ID</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>5F3R</td>
<td>CYP1A2</td>
<td>2640</td>
<td>CYP1A2_60+1.2640.A.A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CYP1A2_60+1.2640.A.A</td>
</tr>
<tr>
<td>5F3R</td>
<td>CYP1A2</td>
<td>2866</td>
<td>CYP1A2_V2.2866.C.S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CYP1A2_V2.2866.C.S</td>
</tr>
<tr>
<td>5F3R</td>
<td>CYP1A2</td>
<td>CNTRL</td>
<td>1A2.5F3R_A.X.S</td>
</tr>
<tr>
<td>5F3R</td>
<td>CYP1A2</td>
<td>CNTRL</td>
<td>1A2.5F3R_B.X.S</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1B1</th>
<th>CYP</th>
<th>SNP ID</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>3793</td>
<td>CYP1B1_60.3793.C.A</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>3947</td>
<td>CYP1B1_60.3947.C.S</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>3976</td>
<td>CYP1B1_60.3976.C.S</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>3987</td>
<td>CYP1B1_60.3987.A.A</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>4035</td>
<td>1B1_4035.C.S</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>4160</td>
<td>CYP1B1_60+2.1460.G.A</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>4306</td>
<td>1B1_V2.4306.A.A</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>4646</td>
<td>CYP1B1_60+1.4646.G.S</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>4668</td>
<td>1B1_4668.C.A</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>CNTRL</td>
<td>1B1.2F4R_A.X.S</td>
</tr>
<tr>
<td>2F4R</td>
<td>CYP1B1</td>
<td>CNTRL</td>
<td>1B1.2F4R_B.X.S</td>
</tr>
</tbody>
</table>

![FIG_7](image-url)
<table>
<thead>
<tr>
<th>SNP ID</th>
<th>PROBE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>2E1_1772</td>
<td>CYP2E1_60.1772.C.A</td>
</tr>
<tr>
<td>2E1_1772</td>
<td>CYP2E1_60.1772.T.A</td>
</tr>
<tr>
<td>2E1_2019</td>
<td>CYP2E1_60.2019.C.S</td>
</tr>
<tr>
<td>2E1_2019</td>
<td>CYP2E1_60.2019.T.S</td>
</tr>
<tr>
<td>2E1_2754</td>
<td>CYP2E1_60.2754.G.S</td>
</tr>
<tr>
<td>2E1_2754</td>
<td>CYP2E1_60.2754.T.S</td>
</tr>
<tr>
<td>2E1_3085</td>
<td>CYP2E1_60.3085.A.A</td>
</tr>
<tr>
<td>2E1_3085</td>
<td>CYP2E1_60.3085.T.A</td>
</tr>
<tr>
<td>2E1_3104</td>
<td>CYP2E1_60.3104.A.A</td>
</tr>
<tr>
<td>2E1_3104</td>
<td>CYP2E1_60.3104.G.A</td>
</tr>
<tr>
<td>2E1_7592</td>
<td>CYP2E1_60.7592.A.S</td>
</tr>
<tr>
<td>2E1_7592</td>
<td>CYP2E1_60.7592.G.S</td>
</tr>
<tr>
<td>2E1_10456</td>
<td>CYP2E1_60.10456.A.S</td>
</tr>
<tr>
<td>2E1_10456</td>
<td>CYP2E1_60.10456.T.S</td>
</tr>
<tr>
<td>2E1_12720</td>
<td>CYP2E1_60.12720.C.S</td>
</tr>
<tr>
<td>2E1_12720</td>
<td>CYP2E1_60.12720.G.S</td>
</tr>
<tr>
<td>2E1_12847</td>
<td>CYP2E1_60.12847.A.S</td>
</tr>
<tr>
<td>2E1_12847</td>
<td>CYP2E1_60.12847.G.S</td>
</tr>
<tr>
<td>2E1_12945</td>
<td>CYP2E1_60.12945.C.A</td>
</tr>
<tr>
<td>2E1_12945</td>
<td>CYP2E1_60.12945.T.A</td>
</tr>
<tr>
<td>2D6_4802</td>
<td>CYP2D6_70.4802.A.S</td>
</tr>
<tr>
<td>2D6_4802</td>
<td>CYP2D6_70.4802.G.S</td>
</tr>
<tr>
<td>2D6_4896</td>
<td>CYP2D6_70.4896.C.S</td>
</tr>
<tr>
<td>2D6_4896</td>
<td>CYP2D6_70.4896.T.S</td>
</tr>
<tr>
<td>2D6_4907</td>
<td>CYP2D6_60.4907.A.S</td>
</tr>
<tr>
<td>2D6_4907</td>
<td>CYP2D6_60.4907.G.S</td>
</tr>
<tr>
<td>2D6_5447</td>
<td>CYP2D6_70.5447.A.A</td>
</tr>
<tr>
<td>2D6_5447</td>
<td>CYP2D6_70.5447.G.A</td>
</tr>
<tr>
<td>2D6_5472</td>
<td>CYP2D6_70.5472.A.S</td>
</tr>
<tr>
<td>2D6_5472</td>
<td>CYP2D6_70.5472.G.S</td>
</tr>
<tr>
<td>2D6_3595</td>
<td>CYP2D6_60.3595.A.A</td>
</tr>
<tr>
<td>2D6_3595</td>
<td>CYP2D6_60.3595.G.A</td>
</tr>
<tr>
<td>2D6_3597</td>
<td>CYP2D6_60.3597.G.S</td>
</tr>
<tr>
<td>2D6_3597</td>
<td>CYP2D6_60.3597.C.S</td>
</tr>
<tr>
<td>2D6_4089</td>
<td>CYP2D6_70.4089.C.S</td>
</tr>
<tr>
<td>2D6_4089</td>
<td>CYP2D6_70.4089.T.S</td>
</tr>
<tr>
<td>2D6_4099</td>
<td>CYP2D6_70.4099.C.S</td>
</tr>
<tr>
<td>2D6_4099</td>
<td>CYP2D6_70.4099.T.S</td>
</tr>
<tr>
<td>2D6_4102</td>
<td>CYP2D6_70.4102.G.A</td>
</tr>
<tr>
<td>2D6_4102</td>
<td>CYP2D6_70.4102.T.A</td>
</tr>
<tr>
<td>2D6_2642</td>
<td>CYP2D6_60.2642.C.S</td>
</tr>
<tr>
<td>2D6_2642</td>
<td>CYP2D6_60.2642.T.S</td>
</tr>
</tbody>
</table>

**FIG. 8A**
<table>
<thead>
<tr>
<th>SNP ID</th>
<th>PROBE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D6 2642</td>
<td>CYP2D6_70.2642.C.S</td>
</tr>
<tr>
<td>2D6 2642</td>
<td>CYP2D6_70.2642.T.S</td>
</tr>
<tr>
<td>2D6 2658</td>
<td>CYP2D6_70.2658.C.A</td>
</tr>
<tr>
<td>2D6 2658</td>
<td>CYP2D6_70.2658.T.A</td>
</tr>
<tr>
<td>2D6 3278</td>
<td>CYP2D6_60.3278.A.S</td>
</tr>
<tr>
<td>2D6 3278</td>
<td>CYP2D6_60.3278.G.S</td>
</tr>
<tr>
<td>2D6 3280</td>
<td>CYP2D6_60.3280.C.S</td>
</tr>
<tr>
<td>2D6 3280</td>
<td>CYP2D6_60.3280.G.S</td>
</tr>
<tr>
<td>2D6 3343</td>
<td>CYP2D6_70.3343.C.A</td>
</tr>
<tr>
<td>2D6 3343</td>
<td>CYP2D6_70.3343.T.A</td>
</tr>
<tr>
<td>2D6 1618</td>
<td>CYP2D6_70.1618.A.S</td>
</tr>
<tr>
<td>2D6 1618</td>
<td>CYP2D6_70.1618.G.S</td>
</tr>
<tr>
<td>2D6 1638</td>
<td>CYP2D6_70.1638.A.S</td>
</tr>
<tr>
<td>2D6 1638</td>
<td>CYP2D6_70.1638.G.S</td>
</tr>
<tr>
<td>2D6 1650</td>
<td>CYP2D6_70.1650.A.A</td>
</tr>
<tr>
<td>2D6 1650</td>
<td>CYP2D6_70.1650.G.A</td>
</tr>
<tr>
<td>2D6 1696</td>
<td>CYP2D6_70.1696.A.S</td>
</tr>
<tr>
<td>2D6 1696</td>
<td>CYP2D6_70.1696.G.S</td>
</tr>
<tr>
<td>2D6 1701</td>
<td>CYP2D6_70.1701.T.A</td>
</tr>
<tr>
<td>2D6 1701</td>
<td>CYP2D6_70.1701.C.A</td>
</tr>
<tr>
<td>2D6 1719</td>
<td>CYP2D6_60.1719.C.A</td>
</tr>
<tr>
<td>2D6 1719</td>
<td>CYP2D6_60.1719.T.A</td>
</tr>
<tr>
<td>2D6 1743</td>
<td>CYP2D6_70.1743.A.S</td>
</tr>
<tr>
<td>2D6 1743</td>
<td>CYP2D6_70.1743.G.S</td>
</tr>
<tr>
<td>2D6 2502</td>
<td>CYP2D6_60.2502.C.A</td>
</tr>
<tr>
<td>2D6 2502</td>
<td>CYP2D6_60.2502.G.A</td>
</tr>
<tr>
<td>2D6 2576</td>
<td>CYP2D6_60.2576.T.S</td>
</tr>
<tr>
<td>2D6 2576</td>
<td>CYP2D6_60.2576.C.S</td>
</tr>
<tr>
<td>2D6 2593</td>
<td>CYP2D6_60.2593.A.A</td>
</tr>
<tr>
<td>2D6 2593</td>
<td>CYP2D6_60.2593.C.A</td>
</tr>
<tr>
<td>2D6 2603</td>
<td>CYP2D6_70.2603.A.S</td>
</tr>
<tr>
<td>2D6 2603</td>
<td>CYP2D6_70.2603.G.S</td>
</tr>
<tr>
<td>2D6 2616</td>
<td>CYP2D6_60.2616.C.A</td>
</tr>
<tr>
<td>2D6 2616</td>
<td>CYP2D6_60.2616.G.A</td>
</tr>
<tr>
<td>2D6 3368</td>
<td>CYP2D6_70.3368.A.A</td>
</tr>
<tr>
<td>2D6 3368</td>
<td>CYP2D6_70.3368.G.S</td>
</tr>
<tr>
<td>2D6 3377</td>
<td>CYP2D6_70.3377.G.S</td>
</tr>
<tr>
<td>2D6 3377</td>
<td>CYP2D6_70.3377.T.S</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_60.3465.A.A</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_60.3465.G.A</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_60.3465.A.S</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_60.3465.G.S</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_70.3465.A.A</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_70.3465.G.A</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_70.3465.A.S</td>
</tr>
<tr>
<td>2D6 3465</td>
<td>CYP2D6_70.3465.G.S</td>
</tr>
<tr>
<td>2D6 3477</td>
<td>CYP2D6_60.3477.C.A</td>
</tr>
<tr>
<td>2D6 3477</td>
<td>CYP2D6_60.3477.T.A</td>
</tr>
</tbody>
</table>

**FIG. 8B**
<table>
<thead>
<tr>
<th>SNP ID</th>
<th>PROBE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D6 3488</td>
<td>CYP2D6_60.3488.C.A</td>
</tr>
<tr>
<td>2D6 3488</td>
<td>CYP2D6_60.3488.T.A</td>
</tr>
<tr>
<td>2D6 3562</td>
<td>CYP2D6_60.3562.A.A</td>
</tr>
<tr>
<td>2D6 3562</td>
<td>CYP2D6_60.3562.G.A</td>
</tr>
<tr>
<td>2D6 4194</td>
<td>CYP2D6_70.4194.A.S</td>
</tr>
<tr>
<td>2D6 4194</td>
<td>CYP2D6_70.4194.C.S</td>
</tr>
<tr>
<td>2D6 4469</td>
<td>CYP2D6_60.4469.C.S</td>
</tr>
<tr>
<td>2D6 4469</td>
<td>CYP2D6_60.4469.T.S</td>
</tr>
<tr>
<td>2D6 4472</td>
<td>CYP2D6_70.4472.A.A</td>
</tr>
<tr>
<td>2D6 4472</td>
<td>CYP2D6_70.4472.C.A</td>
</tr>
<tr>
<td>2D6 4472</td>
<td>CYP2D6_70.4472.A.S</td>
</tr>
<tr>
<td>2D6 4472</td>
<td>CYP2D6_70.4472.C.S</td>
</tr>
<tr>
<td>2D6 4554</td>
<td>CYP2D6_70.4554.A.A</td>
</tr>
<tr>
<td>2D6 4554</td>
<td>CYP2D6_70.4554.C.A</td>
</tr>
<tr>
<td>2D6 4557</td>
<td>CYP2D6_70.4557.C.S</td>
</tr>
<tr>
<td>2D6 4557</td>
<td>CYP2D6_70.4557.T.S</td>
</tr>
<tr>
<td>2D6 4558</td>
<td>CYP2D6_70.4558.A.A</td>
</tr>
<tr>
<td>2D6 4558</td>
<td>CYP2D6_70.4558.G.A</td>
</tr>
<tr>
<td>2D6 5496</td>
<td>CYP2D6_70.5496.C.S</td>
</tr>
<tr>
<td>2D6 5496</td>
<td>CYP2D6_70.5496.G.S</td>
</tr>
<tr>
<td>2D6 5506</td>
<td>CYP2D6_60.5506.C.A</td>
</tr>
<tr>
<td>2D6 5506</td>
<td>CYP2D6_60.5506.T.A</td>
</tr>
<tr>
<td>2D6 5661</td>
<td>CYP2D6_70.5661.A.S</td>
</tr>
<tr>
<td>2D6 5661</td>
<td>CYP2D6_70.5661.G.S</td>
</tr>
<tr>
<td>2D6 5734</td>
<td>CYP2D6_70.5734.C.A</td>
</tr>
<tr>
<td>2D6 5734</td>
<td>CYP2D6_70.5734.T.A</td>
</tr>
<tr>
<td>2D6 5799</td>
<td>CYP2D6_60.5799.C.S</td>
</tr>
<tr>
<td>2D6 5799</td>
<td>CYP2D6_60.5799.G.S</td>
</tr>
<tr>
<td>2E1 1627</td>
<td>CYP2E1_60.1627.C.A</td>
</tr>
<tr>
<td>2E1 1627</td>
<td>CYP2E1_60.1627.G.A</td>
</tr>
<tr>
<td>3A4 816</td>
<td>CYP3A4_60.816.G.S</td>
</tr>
<tr>
<td>3A4 816</td>
<td>CYP3A4_60.816.A.S</td>
</tr>
<tr>
<td>3A4 918</td>
<td>CYP3A4_60.918.A.S</td>
</tr>
<tr>
<td>3A4 918</td>
<td>CYP3A4_60.918.G.S</td>
</tr>
<tr>
<td>2C19 430</td>
<td>CYP2C19EXONS_70.430.C.A</td>
</tr>
<tr>
<td>2C19 430</td>
<td>CYP2C19EXONS_70.430.T.A</td>
</tr>
<tr>
<td>2C19 636</td>
<td>CYP2C19EXONS_60.636.A.S</td>
</tr>
<tr>
<td>2C19 636</td>
<td>CYP2C19EXONS_60.636.G.S</td>
</tr>
<tr>
<td>2C19 681</td>
<td>CYP2C19EXONS_60.681.A.S</td>
</tr>
<tr>
<td>2C19 681</td>
<td>CYP2C19EXONS_60.681.G.S</td>
</tr>
<tr>
<td>1B1 4160</td>
<td>CYP1B1_60.4160.G.A</td>
</tr>
<tr>
<td>1B1 4160</td>
<td>CYP1B1_60.4160.T.A</td>
</tr>
<tr>
<td>1B1 7973</td>
<td>CYP1B1_60.7973.C.A</td>
</tr>
<tr>
<td>1B1 7973</td>
<td>CYP1B1_60.7973.T.A</td>
</tr>
<tr>
<td>1B1 7996</td>
<td>CYP1B1_60.7996.A.S</td>
</tr>
<tr>
<td>1B1 7996</td>
<td>CYP1B1_60.7996.G.S</td>
</tr>
<tr>
<td>1B1 8006</td>
<td>CYP1B1_60.8006.A.S</td>
</tr>
<tr>
<td>1B1 8006</td>
<td>CYP1B1_60.8006.G.S</td>
</tr>
</tbody>
</table>

FIG. 8C
<table>
<thead>
<tr>
<th>SNP_ID</th>
<th>PROBE_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1B1_8195</td>
<td>CYP1B1_60.8195.A.A</td>
</tr>
<tr>
<td>1B1_8195</td>
<td>CYP1B1_60.8195.G.A</td>
</tr>
<tr>
<td>1B1_8242</td>
<td>CYP1B1_60.8242.C.S</td>
</tr>
<tr>
<td>1B1_8242</td>
<td>CYP1B1_60.8242.T.S</td>
</tr>
<tr>
<td>1B1_8587</td>
<td>CYP1B1_60.8587.C.S</td>
</tr>
<tr>
<td>1B1_8587</td>
<td>CYP1B1_60.8587.G.S</td>
</tr>
<tr>
<td>1A1_1223</td>
<td>CYP1A1_60.1223.C.A</td>
</tr>
<tr>
<td>1A1_1223</td>
<td>CYP1A1_60.1223.T.A</td>
</tr>
<tr>
<td>2D6_3326</td>
<td>2D6_66.3326.G.A</td>
</tr>
<tr>
<td>2D6_3326</td>
<td>2D6_66.3326.T.A</td>
</tr>
<tr>
<td>2D6_3326</td>
<td>2D6_66.3326.G.S</td>
</tr>
<tr>
<td>2D6_3326</td>
<td>2D6_66.3326.T.S</td>
</tr>
<tr>
<td>2D6_4168</td>
<td>2D6_66.4168.A.A</td>
</tr>
<tr>
<td>2D6_4168</td>
<td>2D6_66.4168.C.A</td>
</tr>
<tr>
<td>2D6_4168</td>
<td>2D6_66.4168.A.S</td>
</tr>
<tr>
<td>2D6_4168</td>
<td>2D6_66.4168.C.S</td>
</tr>
<tr>
<td>1A1_6568</td>
<td>CYP1A1_60.6568.A.A</td>
</tr>
<tr>
<td>1A1_6568</td>
<td>CYP1A1_60.6568.C.A</td>
</tr>
<tr>
<td>1A1_6568</td>
<td>CYP1A1_60.6568.A.S</td>
</tr>
<tr>
<td>1A1_6568</td>
<td>CYP1A1_60.6568.C.S</td>
</tr>
<tr>
<td>1A1_6570</td>
<td>CYP1A1_60.6570.A.A</td>
</tr>
<tr>
<td>1A1_6570</td>
<td>CYP1A1_60.6570.G.A</td>
</tr>
<tr>
<td>1A2_2640</td>
<td>CYP1A2_60.2640.A.A</td>
</tr>
<tr>
<td>1A2_2640</td>
<td>CYP1A2_60.2640.C.A</td>
</tr>
<tr>
<td>1A2_2866</td>
<td>CYP1A2_60.2866.C.S</td>
</tr>
<tr>
<td>1A2_2866</td>
<td>CYP1A2_60.2866.G.S</td>
</tr>
<tr>
<td>1B1_3793</td>
<td>CYP1B1_60.3793.C.A</td>
</tr>
<tr>
<td>1B1_3793</td>
<td>CYP1B1_60.3793.T.A</td>
</tr>
<tr>
<td>1B1_3793</td>
<td>CYP1B1_60.3793.C.S</td>
</tr>
<tr>
<td>1B1_3947</td>
<td>CYP1B1_60.3947.C.S</td>
</tr>
<tr>
<td>1B1_3947</td>
<td>CYP1B1_60.3947.G.S</td>
</tr>
<tr>
<td>1B1_3976</td>
<td>CYP1B1_60.3976.C.S</td>
</tr>
<tr>
<td>1B1_3976</td>
<td>CYP1B1_60.3976.G.S</td>
</tr>
<tr>
<td>1B1_3987</td>
<td>CYP1B1_60.3987.A.A</td>
</tr>
<tr>
<td>1B1_3987</td>
<td>CYP1B1_60.3987.G.A</td>
</tr>
<tr>
<td>1B1_8807</td>
<td>CYP1B1_60.8807.A.S</td>
</tr>
<tr>
<td>1B1_8807</td>
<td>CYP1B1_60.8807.T.S</td>
</tr>
<tr>
<td>2C19_276</td>
<td>CYP2C19EXONS_70.276.C.A</td>
</tr>
<tr>
<td>2C19_276</td>
<td>CYP2C19EXONS_70.276.G.A</td>
</tr>
<tr>
<td>2C19_395</td>
<td>CYP2C19EXONS_70.395.A.A</td>
</tr>
<tr>
<td>2C19_395</td>
<td>CYP2C19EXONS_70.395.G.A</td>
</tr>
</tbody>
</table>

**CONTROL PROBES**

PBR322WSNPS.4056.C.S
PBR322WSNPS.4056.T.S
WIAF-1648.107.A.A
WIAF-1648.107.G.A

**FIG. 8D**
<table>
<thead>
<tr>
<th>SNP_ID</th>
<th>PROBE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WIAF-198.38.C.A</td>
</tr>
<tr>
<td></td>
<td>WIAF-198.38.T.A</td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.111.C.A</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.111.G.A</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.111.C.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.111.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.23.A.A</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.23.G.A</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.2370.A.S</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.2370.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.2592.C.S</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.2592.T.S</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.3471.C.A</td>
<td></td>
</tr>
<tr>
<td>2D7APSEUDOGENECONTR_60.3471.T.A</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.600.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.600.T.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.1760.C.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.1760.T.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.2108.A.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.2108.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.3539.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.3539.T.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.3647.A.A</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.3647.G.A</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.3766.A.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.3766.C.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.4506.C.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.4506.G.S</td>
<td></td>
</tr>
<tr>
<td>2D8_60.105.A.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.105.G.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.3080.A.S</td>
<td></td>
</tr>
<tr>
<td>2D8_60.3080.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.1360.C.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.1360.T.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.3030.A.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.3030.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.3148.A.S</td>
<td></td>
</tr>
<tr>
<td>2D7PSEUDOGENECONTR_60.3148.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.442.C.A</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.442.T.A</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.652.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.652.T.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.1185.G.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.1185.T.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.1316.A.A</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.1316.C.A</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.1671.A.A</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.1671.T.A</td>
<td></td>
</tr>
</tbody>
</table>

**FIG._8E**
<table>
<thead>
<tr>
<th>SNP_ID</th>
<th>PROBE ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D7B_60.3172.C.S</td>
<td></td>
</tr>
<tr>
<td>2D7B_60.3172.G.S</td>
<td></td>
</tr>
<tr>
<td>2D8_60.3181.A.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.3181.G.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4120.A.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4120.G.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4199.C.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4199.T.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4223.C.S</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4223.T.S</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4750.C.A</td>
<td></td>
</tr>
<tr>
<td>2D8_60.4750.G.A</td>
<td></td>
</tr>
</tbody>
</table>

FIG._8F
<table>
<thead>
<tr>
<th>GENE</th>
<th>RELATED / MAPPED TOGETHER</th>
<th>DNA / mRNA / PARTIALS</th>
<th>ACCESSION #</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D6</td>
<td>2D6</td>
<td>gDNA</td>
<td>M33388</td>
</tr>
<tr>
<td></td>
<td>2D7</td>
<td>gDNA</td>
<td>M33387</td>
</tr>
<tr>
<td></td>
<td>2D7A</td>
<td>gDNA</td>
<td>X58467</td>
</tr>
<tr>
<td></td>
<td>2D7B</td>
<td>gDNA</td>
<td>X58468</td>
</tr>
<tr>
<td></td>
<td>2D8</td>
<td>gDNA</td>
<td>M33387</td>
</tr>
<tr>
<td>2E1</td>
<td>2E1</td>
<td>gDNA</td>
<td>J02843</td>
</tr>
<tr>
<td>3A4</td>
<td>3A4</td>
<td>Partial gDNA</td>
<td>D11131</td>
</tr>
<tr>
<td></td>
<td>gDNA</td>
<td></td>
<td>AF209389</td>
</tr>
<tr>
<td>3A5</td>
<td>3A5</td>
<td>gDNA</td>
<td>AC005020</td>
</tr>
<tr>
<td></td>
<td>mRNA</td>
<td></td>
<td>NM_000777</td>
</tr>
<tr>
<td></td>
<td>Splice Variant (mRNA)</td>
<td></td>
<td>L26985</td>
</tr>
<tr>
<td>1A1</td>
<td>1A1</td>
<td>gDNA</td>
<td>X04300</td>
</tr>
<tr>
<td></td>
<td>gDNA</td>
<td></td>
<td>X02612</td>
</tr>
<tr>
<td></td>
<td>Partial</td>
<td></td>
<td>D12525</td>
</tr>
<tr>
<td>1A2</td>
<td>1A2</td>
<td>Partial gDNA</td>
<td>M31664</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA</td>
<td></td>
<td>M31665</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA</td>
<td></td>
<td>M31666</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA</td>
<td></td>
<td>M31667</td>
</tr>
<tr>
<td>1B1</td>
<td>1B1</td>
<td>gDNA</td>
<td>U56438</td>
</tr>
<tr>
<td>2C9</td>
<td>2C9</td>
<td>mRNA</td>
<td>M61855</td>
</tr>
<tr>
<td></td>
<td>2C9</td>
<td>Partial gDNA</td>
<td>L16877</td>
</tr>
<tr>
<td></td>
<td>2C9</td>
<td>Partial gDNA</td>
<td>L16878</td>
</tr>
<tr>
<td></td>
<td>2C9</td>
<td>Partial gDNA</td>
<td>L16879</td>
</tr>
<tr>
<td></td>
<td>2C9</td>
<td>Partial gDNA</td>
<td>L16880</td>
</tr>
<tr>
<td></td>
<td>2C9</td>
<td>Partial gDNA</td>
<td>L16881</td>
</tr>
<tr>
<td></td>
<td>2C9</td>
<td>Partial gDNA</td>
<td>L16882</td>
</tr>
<tr>
<td></td>
<td>2C9</td>
<td>Partial gDNA</td>
<td>L16883</td>
</tr>
<tr>
<td>2C19</td>
<td>2C19</td>
<td>mRNA</td>
<td>NM_000769</td>
</tr>
<tr>
<td></td>
<td>Draft gDNA (Exons 8-9)</td>
<td></td>
<td>AL133513</td>
</tr>
<tr>
<td>2C18</td>
<td>mRNA</td>
<td></td>
<td>M61856</td>
</tr>
<tr>
<td>2C8</td>
<td>mRNA</td>
<td></td>
<td>NM_000770</td>
</tr>
<tr>
<td>2C19</td>
<td>mRNA</td>
<td></td>
<td>NM_000769</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA (Exons 8-9)</td>
<td></td>
<td>AL133513</td>
</tr>
<tr>
<td>2C9</td>
<td>mRNA</td>
<td></td>
<td>M61855</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA</td>
<td></td>
<td>L16877</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA</td>
<td></td>
<td>L16878</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA</td>
<td></td>
<td>L16879</td>
</tr>
<tr>
<td></td>
<td>Partial gDNA</td>
<td></td>
<td>L16880</td>
</tr>
</tbody>
</table>

**FIG. 9A**
<table>
<thead>
<tr>
<th>GENE</th>
<th>RELATED / MAPPED TOGETHER</th>
<th>DNA / mRNA / PARTIALS</th>
<th>ACCESSION #</th>
</tr>
</thead>
<tbody>
<tr>
<td>2C9</td>
<td>Partial gDNA</td>
<td></td>
<td>L16881</td>
</tr>
<tr>
<td>2C9</td>
<td>Partial gDNA</td>
<td></td>
<td>L16882</td>
</tr>
<tr>
<td>2C9</td>
<td>Partial gDNA</td>
<td></td>
<td>L16883</td>
</tr>
<tr>
<td>2C18</td>
<td>mRNA</td>
<td></td>
<td>M61856</td>
</tr>
<tr>
<td>2C8</td>
<td>mRNA</td>
<td></td>
<td>NM_000770</td>
</tr>
</tbody>
</table>

**FIG._9B**

<table>
<thead>
<tr>
<th>AMPLICON</th>
<th>LENGTH</th>
<th>SNP_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A1_23F_22R</td>
<td>1127</td>
<td>1A1_6568</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1A1_6570</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1A1_7320</td>
</tr>
<tr>
<td>1A2_05F_03R</td>
<td>378</td>
<td>1A2_2640</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1A2_2866</td>
</tr>
<tr>
<td>1B1_02F_04R</td>
<td>1007</td>
<td>1B1_3947</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_3973</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_3976</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_3987</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_4160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_4646</td>
</tr>
<tr>
<td>1B1_08F_11R</td>
<td>1353</td>
<td>1B1_7930</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_7957</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_7973</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_7996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8131</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8147</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8184</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8195</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8242</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8587</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_8807</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1B1_9164</td>
</tr>
<tr>
<td>2C19_03F_06R</td>
<td>~6500</td>
<td>2C19_276</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2C19_395</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2C19_430</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2C19_636</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2C19_681</td>
</tr>
<tr>
<td>2D6_01F_01R</td>
<td>4522</td>
<td>2D6_1618</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2D6_1638</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2D6_1650</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2D6_1696</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2D6_1701</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2D6_1719</td>
</tr>
</tbody>
</table>

**FIG._10A**
<table>
<thead>
<tr>
<th>AMPLICON</th>
<th>LENGTH</th>
<th>SNP_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D6_1743</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_1757</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_2502</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_2578</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_2593</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_2603</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_2616</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_2642</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_2658</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3278</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3280</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3323</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3343</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3368</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3377</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3465</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3477</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3488</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3562</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3592</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3597</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_3598</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4168</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4206</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4232</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4472</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4554</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4557</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4558</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4802</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4817</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4896</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_4907</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_5447</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_5472</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_5496</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_5506</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_5661</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_5734</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D6_5799</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIG. 10B
The Codelink™ SNP Bioarray for Human Cytochrome P450 Genes

**FIG. 11**

---

Layout of Primer Pairs in Primer Plates

<table>
<thead>
<tr>
<th>Individual</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplicon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**NOTE:** Each sample individual will be aliquoted across 12 separate PCR reactions.

**FIG. 12**
Addition of Master Mix and Samples to Microfuge Tubes

**FIG. 13**

Master Mix (57.2 μL)

13 μL

27 μL of Sample 1

27 μL of Sample 2

27 μL of Sample 3

27 μL of Sample 4

Tip Orientation for Loading Reaction Mixtures into Chambers

**FIG. 14**

90°

Orientation of Sealing Strips Over Chamber Ports

**FIG. 15**
Prevention of Self-Extension Due to Base Additions

No Probe Extension

 FIG._17

| A | G*C | G*C | C*G | A*T | CxG | A |

 FIG._19

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[Bars]</td>
<td>[Bars]</td>
<td>[Bars]</td>
<td>[Bars]</td>
</tr>
</tbody>
</table>

- Target
+ Target
### Table

<table>
<thead>
<tr>
<th>Target-Independent Signal Strength</th>
<th>Probe</th>
<th>Probe Sequence (5' to 3')</th>
<th>Length / #GC</th>
<th>Predicted Stem-loop</th>
<th>Predicted Base Extended</th>
<th>Observed Base Extended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>50APOE321.T.A</td>
<td>TACACTGCGAGGCA</td>
<td>14/8</td>
<td>TGCCAGGCA</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Strong</td>
<td>60WIAF913.114.T.A</td>
<td>TCTCTGCTGCTCTCTGTGGCA</td>
<td>20/10</td>
<td>TGCTCTTGGCA</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Strong</td>
<td>60PMCC7111G.111.C.A</td>
<td>AAGTGCTCCCATGGAGTAGGAG</td>
<td>21/11</td>
<td>CTCC(8)GGAC</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Strong</td>
<td>70PMCC7111G.111.C.S</td>
<td>CGGAGGGCAAGGCG</td>
<td>15/12</td>
<td>GCG(8)GGG</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Strong</td>
<td>70LP21.150.C.A</td>
<td>CCCAGATGCCTACCCAGCTTG</td>
<td>21/13</td>
<td>CAGGGCT</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>Strong</td>
<td>60WIAF266.173.C.S</td>
<td>GCGAGGCAATTTATTTTG</td>
<td>19/8</td>
<td>GCAA(6)TTG</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>Strong</td>
<td>APOE182.A.A</td>
<td>CAGGGGCGGCT</td>
<td>12/10</td>
<td>GCGCGCGGCT</td>
<td>T</td>
<td>G&gt;C</td>
</tr>
</tbody>
</table>

Table: Examples of probe sequences that show a strong target-independent signal in the SBE assay. The predicted stem-loop region is underlined.

---

**FIG. 18**

- P = 1.0 For All Alleles
- One Allele Not
- Heterozygotes Not Calling
- P = 0.5-0.99 One Allele

**FIG. 20**
A. Target-dependent Extension

Surface-attached Probe

Target Binding, Polymerase Extension

Correct Signal

B. Target-independent Extension

Surface-attached Probe

Self-folding, Unstable, No Extension

Incorrect Signal Suppressed

FIG. 21
Natural A:T base pair, pairs equally well with target and itself

2-am-A:T base pair, (target-probe pair) forms a very stable base-pair

Non-natural 2-am-A:2-thioT base-pair, does not form a stable base-pair

A:2-thio-T base pair, (target-probe pair) forms a stable base-pair

Exo-cyclic amine modified A

Exo-cyclic amine modified G

Exo-cyclic amine modified C

Modified A

Modified G

Modified C
A. Target-dependent Extension

Surface-attached Probe

Target Binding, Polymerase Extension

Correct Signal

B. Target-independent Extension

Surface-attached Probe

Self-folding, Extension Disabled

By Terminator Base

Incorrect Signal Suppressed

FIG._24

"Terminator" Base

4-methyl-indole

FIG._25A

"Terminator" Nucleoside

2'-O-methyl-2-amino-A

FIG._25B
Table: List of modified bases / nucleosides used,

Q = abasic nucleotide: no base-pairing ability, no stacking energy; placed immediately downstream of putative stem-loop; expect A to be incorporated when Q is in template (the "A rule").

I = 4-methylindole: A analog; placed immediately downstream of putative stem-loop; terminates DNA polymerase activity.

K = 5-nitroindole: universal base; placed immediately downstream of putative stem-loop; does not form base pairs but contributes stacking energy.

Z = 2-amino-A: placed within the stem of putative stem-loop; forms 3 hydrogen-bonds with T; no base pairing with 2-thio-dT.

X = 2-thio-dT: placed with stem of putative stem-loop; stable base pairs with A; no base pairing with 2-amino-A.

FIG. 26

WIAF-1648.107.A.S = 5'-ATACACACATGTGCACACACA

R = 2.31  R = 3.18  R = 9.96  R = 4.85

Blank  No Probe  TA  XZ  I

FIG. 27

[Graph showing bar chart with different configurations and ratios]
1. WIAF-906.54.T.A = 5'-TGCCCTTGAGATTATTAGTGTTTATTA
   5'-TGCCCTTGAGATTATTAGTGTTTATTA

2. WIAF-1648.107.G.S = 5'-CACAGATGACACAGC
   5'-CACAGATGACACAGC

3. TSC1987.1400.358.A.S = 5'-TTTGGCTGATGGGTGACCTA
   5'-TTTGGCTGATGGGTGACCTA
   5'-TTTGGCTGATGGGTGACCTA
   5'-TTTGGCTGATGGGTGACCTA

FIG._28

FIG._29
Self-extension inhibited due to modifications on bases in the "stem" region that prevent extension enzyme from binding.

FIG._30

Extension due to probe-target hybrid is not inhibited.

FIG._31
Modified nucleotide bases reduce the binding affinity of the SBE enzyme or extension enzyme.

**FIG._32A**

**FIG._32B-1**

**FIG._32B-2**

**FIG._32B-3**

**FIG._32B-4**
Results From an Experiment Using Oligonucleotide Inhibitors to Prevent Self Extension of Probes in the SBE Assay.

**FIG. 33A**
No Targets / No Inhibitor

**FIG. 33B**
cRNA Targets / No Inhibitor

**FIG. 33C**
cRNA Targets / Apo E Inhibitor
Method For Uniplexed Target Prep for SNP Genotyping and Primer Extension without Self-extension

Genomic DNA

Primer Extension Preamplification with T7 Tailed Random Primer

IVT

cRNA

Hybridize to Array

Extend with RNA-dependent DNA Pol Activity of RT

Self-extender

Not Extended

FIG. 34
Controls

**Enzyme**
- rTth, Modified Reverse Transcriptase

**Results**
- The Self-extender Shows Signal Only in Presence of cRNA
- The Signal is Human Specific, Since Bacterial Probes Do Not Show Signal with the Human cRNA

**FIG. 35**

**Enzyme**
- Modified Reverse Transcriptase Enzyme (rTth)

**Results**
- Modified Enzyme Eliminates Self-extension

**FIG. 36**

**Enzyme**
- Regular Reverse Transcriptase (Superscript II)

**Results**
- Self-extension Seen with Regular Enzyme
Reaction of Amino Oligonucleotides on the SurModics Surface
(An Example of Acyl Substitution Reaction on the Polymer Backbone)

\[
\begin{align*}
\text{Oligo} & \quad \text{NH}_2 \\
\text{Oligo} & \quad \text{NH}
\end{align*}
\]

\[
\begin{align*}
\text{FIG. - 37}
\end{align*}
\]
P450 SINGLE NUCLEOTIDE POLYMORPHISM BIOCHIP ANALYSIS

FIELD OF THE INVENTION

[0001] The invention is generally directed to novel methods and compositions for the determination of single nucleotide polymorphisms (SNPs) in P450 genes using novel probes and methods that improve the specificity and efficiency of P450 SNP detection. This invention provides a unique collection of P450 SNP probes on one assay which can be performed on a variety of array platforms, the primer sequences for specific amplification of each of the seven P450 genes and amplicon control probes to evaluate whether the intended p450 gene targets were amplified successfully.

BACKGROUND OF THE INVENTION

[0002] The basic purpose of drug metabolism in the body is to make drugs more water soluble and thus more readily excreted in the urine or bile. One common way of metabolizing drugs involves the alteration of functional groups on the parent molecule (e.g. oxidation), via the cytochrome P450 enzymes. These enzymes are predominantly found in the liver. Cytochrome p450 enzymes are involved in numerous drug metabolism pathways as well as pathways used to make cholesterol, steroids, and other important lipids such as prostacyclins and thromboxane A2. Many drug interactions are a result of inhibition or induction of cytochrome P450 enzymes.

[0003] Mammalian cytochrome P450 genes encode a superfamily of hemeproteins that are active in the oxidative metabolism of endogenous and xenogenous compounds. Cytochrome P450 (referred to as CYP) are now classified into families on the basis of amino acid similarity; within families cytochrome P450 exhibit >40% similarity and >55% similarity within subfamilies. The cytochrome P450 enzymes are designated by the letters “CYP” followed by a numeral, a letter and another numeral (e.g. CYP2D6). In humans there are more than 20 different CYP enzymes. According to a recent compilation, the cytochrome P450 2C subfamily contains 36 distinct genes and pseudogenes, and is the largest cytochrome P450 subfamily that has been identified to date. It is generally accepted that mammalian cytochrome P450 genes are regulated primarily at the level of transcription, and there is little information known on the factors that regulate transcription.

[0004] There are a wide variety of polymorphisms (e.g. mutations) in p450 enzymes, and many of these polymorphisms result in (or are associated with) the inhibition or induction of enzymatic activity. These polymorphisms frequently occur in ethnic populations. The cytochrome P450 genes that have been studied most intensively in this regard are cytochrome P450 2D6, the debrisoquine hydroxylase gene, and cytochrome P450 2C19, which codes for the S-mephénylind hydroxylase. A variety of other P450 enzymes exhibit variation in expression levels; for example, cytochrome P450 3A4, the major cytochrome P450 present in human liver, varies over a several-fold range at both the protein and mRNA levels. In addition, many of the P450 enzymes are subject to induction by different drugs, including barbiturates, antibiotics, etc.

[0005] Of particular interest to the present invention are CYP1A1, CYP1A2, CYP1B1, CYP2C19, CYP3D6, CYP2E1 and CYP3A4. CYP2D6 has been studied extensively because it exhibits significant diversity, with roughly 7 to 10 percent of Caucasians being poor metabolizers of drugs metabolized by CYP2D6.

[0006] Patients with normal CYP2D6 activity are termed extensive metabolizers, with Asians and African Americans being less likely than Caucasians to be poor metabolizers. Poor metabolizers are at risk for drug accumulation and toxicity from drugs metabolized by this enzyme. While only 2 to 6 percent of total liver cytochrome P450 is CYP2D6, nearly 25 percent of clinically useful medications are metabolized by this enzyme. Poor metabolizers of CYP2D6 substrates are at risk for increased toxicity from medications that are metabolized by CYP2D6. Conversely, when formation of an active metabolite is essential for drug action, poor metabolizers of CYP2D6 can exhibit less response to drug therapy compared with extensive metabolizers. About 15 percent of clinically used medications are metabolized by CYP1A2, and it is the only CYP that is induced by tobacco and other polycyclic aromatic hydrocarbons.

[0007] CYP2E1 metabolizes a relatively small fraction of medications (although it has a significant role in the metabolism of acetaminophen), it plays a significant role in activation and inactivation of toxins. It is inducible by ethanol and metabolized primarily small organic molecules.

[0008] Members of the CYP3A family are the most abundant and most clinically significant cytochrome enzymes in humans, with CYP3A4 being the most common form and the most widely implicated in most drug interactions. The CYP3A family is located in the small intestine and in addition to drug, is also responsible for metabolizing most of the body’s endogenous steroids.

[0009] CYP2C19, along with CYP2D6, also exhibits genetic polymorphism, with 3 percent of Caucasians and 20 percent of Japanese lacking the enzyme completely. These individuals are at risk for more frequent and more severe adverse effects because of decreased elimination of drugs metabolized by CYP2C19.

[0010] Thus, detection of specific P450 SNPs is important in diagnostic medicine and molecular biology research and also, to understand the mechanism of action of many drugs and are likely to be the direct cause of therapeutically relevant variant and/or disease predispositions.

[0011] Accurate SNP detection requires good sensitivity in the SBE assays. Two major hurdles for highly parallel screening of SNPs on microarrays are: 1) the necessity to amplify DNA regions spanning the SNPs by PCR to achieve sufficient sensitivity and specificity of detecting a single-base variation in the complex human genome in a reproducible way; and, 2) the ability to distinguish unequivocally between homozygous and heterozygous allelic variants in the diploid human genome. Differential hybridization with allele-specific oligonucleotide (ASO) probes is most commonly used in the microarray format (Pastinen et al., Genome Research 2000). The requirement for sensitivity (i.e. low detection limits) has been greatly alleviated by the development of the polymerase chain reaction (PCR) and other amplification technologies which allow researchers to amplify exponentially a specific nucleic acid sequence before analysis (for a review, see Abramson et al., Current Opinion in Biotechnology, 4:41-47 (1993)). Multiplex PCR
amplification of SNP loci with subsequent hybridization to oligonucleotide arrays has been shown to be an accurate and reliable method of simultaneously genotyping at least hundreds of SNPs; see Wang et al., Science, 280:1077 (1998); see also Schafer et al., Nature Biotechnology 16:33-39 (1998).

[0012] Specificity, in contrast, remains a problem in many currently available gene probe assays. The extent of molecular complementarity between probe and target defines the specificity of the interaction. Variations in the concentrations of probes, of targets and of salts in the hybridization medium, in the reaction temperature, and in the length of the probe may alter or influence the specificity of the probe/target interaction.

[0013] It may be possible under some circumstances to distinguish targets with perfect complementarity from targets with mismatches, although this is generally very difficult using traditional technology, since small variations in the reaction conditions will alter the hybridization. New experimental techniques for mismatch detection with standard probes, as defined in greater detail below, include, but are not limited to, OLA, RCA, Invader®, single base extension (SBE) methods, allelic PCR, and competitive probe analysis. In SBE assays, a polynucleotide probe is attached to a support and hybridized to target DNA.

[0014] Generally, for SBE assays, probe sets are designed such that the nucleotide at the 3' end of the probe is either matched or mismatched with the queried base in the target. If the base matches and hybridizes, the DNA polymerase will extend the probe by one base in the presence of four labeled-terminator nucleotides. Alternately, if the 3' base is mismatched, the DNA polymerase does not extend the probe. Thus, the identity of the SNP or queried base in the target is determined by the probe set that is extended by the DNA polymerase.

[0015] Some probes form internal stem-loop structures resulting in target-independent self-extension of the probe thus giving a false positive signal that interferes with determination of the SNP base. The present invention aims to overcome such problems.

[0016] Accordingly, it is an object of the present invention to provide compositions and methods for evaluating samples from one or more patients, to ascertain the level and/or genotype of various P450 enzymes present. Another object of the present invention is to increase the sensitivity and specificity of the P450 SBE assays. The present invention uses a combination of amplification methods, and a variety of methods to prevent self-extension of capture probes in the absence of target thereby reducing the occurrence of false positive results.

SUMMARY OF THE INVENTION

[0017] In accordance with the objects outlined above the present invention provides a biochip comprising a solid substrate comprising an array comprising at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY1A1, at least one capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY1A2, at least one capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY1B1, at least one capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2C19, at least one capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2D6, at least one capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2E1, and at least one capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY3A4.

[0018] In addition, the invention provides a method of determining the identification of a nucleotide at a detection position in at least one target sequence selected from the group consisting of CPY1A1, CPY1A2, CPY1B1, CPY2C19, CPY2D6, CPY2E1 and CPY3A4, the method including providing an array comprising at least one first capture probe substantially homologous to a first portion of a nucleic acid encoding CPY1A1, wherein the first capture probe is directly adjacent to or includes at its terminus a detection position, at least one second capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY1A2, wherein the second capture probe is directly adjacent to or includes at its terminus a detection position, at least one third capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY1B1, wherein the third capture probe is directly adjacent to or includes at its terminus a detection position, at least one fourth capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2C19, wherein the fourth capture probe is directly adjacent to or includes at its terminus a detection position, at least one fifth capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2D6, wherein the fifth capture probe is directly adjacent to or includes at its terminus a detection position, at least one sixth capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2E1, wherein the sixth capture probe is directly adjacent to or includes at its terminus a detection position, at least one seventh capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY3A4, wherein the seventh capture probe is directly adjacent to or includes at its terminus a detection position.

[0019] In addition the invention provides a method of determining the identification of a nucleotide at a detection position in a target sequence. The method includes providing an array that includes a solid support with a first surface comprising a hydrogel layer comprising an array of capture probes, hybridizing the target sequence to at least one of the capture probes to form a hybridization complex and determining the nucleotide at the detection position.

[0020] In addition the invention provides a method of determining the identification of a nucleotide at a detection position in a target sequence. The method includes providing a solid support with a first surface comprising at least one extension probe that has been modified to form a non-self
extension probe, such that self extension of the non-self extension probe does not occur in the absence of the target and wherein, the non-self extension probe includes an interrogation nucleotide, hybridizing the target sequence to the non-self extension probe to form a hybridization complex, contacting the surface with an extension enzyme and at least one chain terminating nucleotide comprising a hapten under conditions whereby if the chain terminating nucleotide is perfectly complementary to the base of the target sequence immediately adjacent to the 3' end of the non-self extension probe in the hybridization complex, the chain terminating nucleotide is added to the non-self extension probe to form a modified extension probe. In addition the method includes contacting the modified extension probe with the binding partner of the hapten, wherein the binding partner is labeled and detecting the presence of the label to determine the nucleotide at the detection position.

0021 In addition the invention includes a method of determining the identification of a nucleotide at a detection position in a target sequence comprising amplifying the target DNA using random primers to generate DNA amplicons, transcribing the DNA amplicons to generate RNA target sequences (in vitro transcription), providing a solid support with a first surface comprising at least one one extension probe wherein the extension probe includes an interrogation nucleotide, hybridizing the RNA target sequence to the extension probe to form a hybridization complex, contacting the surface with a modified reverse transcriptase and at least one chain terminating nucleotide comprising a hapten under conditions whereby if the chain terminating nucleotide is perfectly complementary to the base of the target sequence immediately adjacent to the 3' end of the redesigned extension probe in the hybridization complex, the chain terminating nucleotide is added to the redesigned extension probe to form a modified extension probe contacting the modified extension probe with the binding partner of the hapten, wherein the binding partner is labeled and detecting the presence of the label to determine the nucleotide at the detection position.

0022 In addition the invention provides a method of determining the identification of a nucleotide at a detection position in a target sequence comprising providing a solid support with a first surface comprising a solid support with a first surface comprising a hydrogel layer comprising at least one extension probe, wherein the extension probe includes an interrogation nucleotide within two bases of the 3' end of the extension probe, hybridizing the target sequence to the extension probe to form a hybridization complex contacting the surface with an extension enzyme and at least one chain terminating nucleotide comprising a hapten under conditions whereby if the chain terminating nucleotide is perfectly complementary to the base of the target sequence immediately adjacent to the 3' end of the extension probe in the hybridization complex, the chain terminating nucleotide is added to the extension probe to form a modified extension probe, contacting the modified extension probe with the binding partner of the hapten, wherein the binding partner is labeled and detecting the presence of the label to determine the nucleotide at the detection position.

BRIEF DESCRIPTION OF THE DRAWINGS

0023 FIG. 1 depicts two different configurations of the SBE genotyping reaction. In FIG. 1A, target 5 with detection position 10 is hybridized to capture probe 25, attached to a solid support 15 via an attachment linker 20. The capture probe also serves as the SBE extension probe. FIG. 1B, a first portion of target 5 with detection position 10 is hybridized to capture probe 25, attached to a solid support 15 via an attachment linker 20. A second portion of target 5 is hybridized to an SBE extension primer 30 for a "sandwich" assay.

0024 FIG. 2A depicts a typical single base extension (SBE) assay. FIG. 2B demonstrates how self-extension of a probe in the absence of target sequence can occur in an SBE reaction due to stem-loop structures formed by the capture probe thus producing a false-positive result.

0025 FIGS. 3A and 3B is a worksheet describing the P450 polymorphisms (SNPs) and literature references for each of the genes included into our assay design. FIGS. 3C and 3D is a worksheet describing CYP allele haplotype families.

0026 FIG. 4 depicts a comparison table illustrating the relative homology shared by the genes in the CYP2D6 family. Comparison between corresponding introns and exons is depicted as well as their % similarity.

0027 FIG. 5 depicts the preferred P450 PCR primer list currently being used for amplification. In order to design primers specifically to the target gene of interest, regions were selected with base pair mismatches against subfamily related genes, focusing particularly at the 3' end of the primer. In this manner, the annealing hybridization specificity and the discrimination of the PCR polymerase which extends nucleotides from the 3' end of the primer is relied upon to confer specificity. Primer design was further restricted to a length greater than 19 bp and a balance in terms of GC content and Tm per pair. All primer candidates have been analyzed using the BLAST algorithm (homology analysis algorithm) against a compiled P450 sequence library (57 genes) to remove any potential cross-reactive primer candidates. In addition to homology analysis, the primer candidates were also screened against a database of repetitive sequences.

0028 FIG. 6A through D is the final P450 probe list on the chip for the current P450 assay.

0029 FIG. 7 depicts a summary of the Beta Validation Performance data and the discrimination capabilities of these probes.

0030 FIG. 8A through E depicts the probes designed to identify a particular SNP.

0031 FIGS. 9A and B depicts the relationship of the CYP genes.

0032 FIG. 10A through C summarizes the product amplicon characteristics.

0033 FIG. 11. The Codelink™ SNP Bioarray for human cytochrome P450 genes.

0034 FIG. 12. Layout of primer pairs in primer plates.

0035 FIG. 13. Addition of master mix and samples to microfuge tubes.

0036 FIG. 14. Tip orientation for loading reaction mixtures into chambers.
[0037] FIG. 15. Orientation of sealing strips over chamber ports.


[0039] FIG. 17. Prevention of self-extension due to base additions. Increasing the length of the probe by one or more bases at the 3′ end of the probe creates a mismatch at the end of the stem-loop structure. DNA polymerase will not extend the probe when the 3′ end is mismatched ad hence, there is no self-extension.

[0040] FIG. 18. Shows examples of probe sequences that show strong target-independent self extension in the SBE assay. The predicted stem-loop is underlined.

[0041] FIG. 19. None, one, two or three bases were added to the end of the 2WIAF-913.114.T.A.A probe. The addition of one base reduced the self-extension signal but did not reduce target-dependent signal; this probe is ‘repaired’ and will call the SNP base correctly. The addition of two or three additional bases created new stem-loop structures that resulted in self-extension. The frequency at which additional bases create new self-extenders varies with probe sequence. The sequences of the probes are: 2WIAF-913.114_0_T.A.A; TCTCTGCTGTCTCTTTGGA 2WIAF-913.114_1_T.A.A; TCTCTGCTGTCTCTTTGGACA 2WIAF-913.114_2_T.A.A; TCTCTGCTGTCTCTTTGGACA 2WIAF-913.114_3_T.A.A; TCTCTGCTGTCTCTTTGGA. The putative stem-loop structures are shown in underlined. Note the mismatched 3′ base of 2WIAF-913.114_0_T.A.A.

[0042] FIG. 20. Results from the attempted redesign of 35 probes for the P540 gene family. For 35 probe sets that called the SNP base incorrectly, 23 (66%) the SNP base called correctly 100% of the time when one or two bases were added to the 3′ end of the self-extending probe. P is the product of call rate and accuracy; P=1 indicates that the probe set calls the SNP base correctly 100% of the time.

[0043] FIG. 21(A) depicts target-dependent extension by incorporating labeled nucleotides corresponding to the complementary base in the target sequence. X and Y are the modified bases which have minimal effect on the target-dependent signal generation, as their base-pairing abilities with the natural bases are unaffected (B) The modified base-pair X, Y effectively suppresses the target-independent extension as they cannot form a stable base-pair with each other. Due to lack of stability, no extension occurs and false positive signals are suppressed in the SBE assay.

[0044] FIG. 22. Modified nucleic acids that are used to synthesize novel probes used in preferred embodiments of this invention. This base-pair effectively suppresses the target-independent extension, as they cannot form a stable base-pair with each other as it thermodynamically destabilizes the self-folded product.

[0045] FIG. 23. Exemplary structures of some exo-cyclic amine modified bases.

[0046] FIG. 24. A modified “terminator-type base”. In (A) the modified “terminator-type base” Z is able to bind target nucleic acid and to be extended in an SBE assay. But in (B) the modified “terminator-type base” Z prevents DNA polymerases from extending past the modified base position. Thus, only target-dependent probe-extension takes place and the target-independent extension is suppressed.


[0049] FIG. 27. Graphs of experiments demonstrating the effectiveness of use of probes with modified nucleic acids. In this experiment, probe sequence 5′-ATACACACATGTGGCACACACA was used. This shows that the modified base has its intended effect.

[0050] FIG. 28. Experiments with probe sequence 5′-GCAGCCACTTATTTGCC, which also forms a stable hairpin loop. This result clearly shows that the modified base-pair is having its intended effect. Placement of the modified base-pair closer to the extension site (3′-end) can have an even more dramatic impact.

[0051] FIG. 29. Experiments with three different probe sequences with natural bases. The probes are extended during an SBE assay in the absence of any target. But, when natural bases are replaced with modified bases/nucleosides, target-dependent signal from each one of them increases 5-fold.

[0052] FIG. 30. Modifications of bases in the “stem” region of capture probes inhibits self extension by preventing the polymerase enzyme from binding to the stem-loop structure.

[0053] FIG. 31. The thermodynamic stability of the probe-target duplex allows polymerase enzyme activity and extension is not compromised.

[0054] FIG. 32(A) Modified nucleotide bases used in the “stem” region that reduce the binding affinity of the SBE enzyme (polymerase). (B) depicts chiral phosphodiester analogues that may be used in the “stem” structure of probes to prevent polymerase extension of stem-loop forming probes. (a) Regular phosphodiester bond demonstrating the positions of the pro-R and pro-S non-bridging oxygens which are involved in hydrogen bond formation with enzyme protein contributing to binding affinity and specificity of the enzyme; and three chiral phosphodiester analogues that reduce enzyme binding (b) an H-phosphonate; (c) a phosphorothionate; and (d) a methylphosphonate.

[0055] FIG. 33. Results from and SBE assay using “oligonucleotide inhibitors” to prevent self-extension of probes. A: modified SBE assay performed on RCA 1 SNP chip without both single stranded RNA targets and inhibitor. The false positive signal of APO E321.T.A are indicated in framed box. B: Single-stranded RNA were applied on RCA SNP chip. The intensities of false positive signal in APO E are the same as that of in A: C: The single-strand RNA targets plus short oligo inhibitor were applied on RCA chip, the false positive signal of APO E321 was inhibited, and the signal is now similar in intensity to other positive signals present on the chip.

[0056] FIG. 34. Method for uniplexed target preparation for SNP genotyping and primer extension without self-extension. This combination of methods and modified enzyme defines one of the preferred embodiments of the invention.

[0057] FIG. 35. Results obtained from control probes wherein no PCR or other amplification technology was used
for SNP genotyping. Primer extension preamplification followed by in vitro translation and then probe extension using a modified reverse transcriptase was used in this method.

[0058] FIG. 36. Demonstrates the advantage of the modified reverse transcriptase over the regular reverse transcriptase. The modified enzyme eliminates self-extension.

[0059] FIG. 37. Reaction of amino oligonucleotides on the SurModics™ surface. Here, an example of an acyl substitution reaction on the polymer backbone is shown.

[0060] FIG. 38. Possible structure of a SurModics™ Gel Matrix, a polymer used on a solid substrate in a preferred embodiment.

DETAILED DESCRIPTION OF THE INVENTION

[0061] The present invention is directed to methods and compositions, including biochips, comprising a particularly useful combination of probes designed to elucidate the genotypes (including the presence of single nucleotide polymorphisms, or SN Ps) of a variety of p450 enzymes. The compositions and methods of the invention rely on the use of PCR primers (although, as will be appreciated by those in the art, other types of amplification reactions may be done as well), that provide unique and advantageous specificity between several of the p450 genes, to allow the specific amplification of each of the relevant genes. The resulting amplicons are then analyzed for SNPs using novel probes that improve the specificity and efficiency of the single base extension (SBE) reaction, to allow discrimination and detection of p450 enzyme variants, which allows correlation to disease states and drug susceptibility and resistance.

[0062] Accordingly, it is an object of the present invention to provide compositions and methods for analyzing and evaluating samples from one or more patients, to ascertain the level and/or genotype of various P450 enzymes present. As will be appreciated by those in the art, the sample solution may comprise any number of things, including, but not limited to, bodily fluids (including, but not limited to, blood, urine, serum, lymph, saliva, anal and vaginal secretions, perspiration and semen, of virtually any organism, with mammalian samples being preferred and human samples being particularly preferred). As will be appreciated by those in the art, the sample may be the product of an amplification reaction, including both target and signal amplification as is generally described in PCT/US99/01705, such as PCR, etc., amplification reactions and outlined below. As will be appreciated by those in the art, virtually any experimental manipulation may have been done on the sample.

[0063] The compositions and methods of the invention are directed to the detection of SNPs on target DNA sequences. The term “target sequence” or “target nucleic acid” or “target analyte” or grammatical equivalents herein means a nucleic acid sequence, that may be a portion of a gene, a regulatory sequence, genomic DNA, cDNA, RNA including mRNA and rRNA, or others. As is outlined herein, the target sequence may be a target sequence from a sample, or a secondary target such as a product of an amplification reaction such as PCR etc. It may be any length, with the understanding that longer sequences are more specific. As will be appreciated by those in the art, the complementary target sequence may take many forms. For example, it may be contained within a larger nucleic acid sequence, i.e. all or part of a gene or mRNA, a restriction fragment of a plasmid or genomic DNA, among others. As is outlined more fully below, probes are made to hybridize to target sequences to determine the presence or absence of SNPs in the target sequence of a sample. Generally speaking, this term will be understood by those skilled in the art. The target sequence may also be comprised of different target domains; for example, a first target domain of the sample target sequence may hybridize to a first capture probe, a second target domain may hybridize to a portion of a capture probe, etc. The target domains may be adjacent or separated as indicated. Unless specified, the terms “first” and “second” are not meant to confer an orientation of the sequences with respect to the 5'-3' orientation of the target sequence. For example, assuming a 5'-3' orientation of the complementary target sequence, the first target domain may be located either 5' to the second domain, or 3' to the second domain.

[0064] Single base changes that are inherited are generally referred to as polymorphisms or SNPs. As is more fully outlined below, preferred embodiments of the invention comprise target sequences comprising a SNP (single nucleotide polymorphism) or a plurality of SNPs for which sequence information is desired, generally referred to herein as “SNP site” or “SNP position” or the “queried base” or the “interrogation position” or the “detection position”. In a preferred embodiment, the SNP position is a single nucleotide, although in some embodiments, it may comprise a plurality of nucleotides, either contiguous with each other or separated by one or more nucleotides. By “plurality” as used herein is meant at least two. As used herein, the capture probe comprises the “interrogation position”, usually, but not always, at the 3' end or towards the 3' end.

[0065] For the purposes of this invention, sequences are referred to herein as “perfectly matched” or “mismatched”. It should be noted in this context that “mismatch” is a relative term and meant to indicate a difference in the identity of a base between two sequences in two different probes at the particular SNP position. In general, sequences that differ from wild type sequences are referred to as mismatches. However, and particularly in the case of SNPs, what constitutes “wild type” may be difficult to determine as multiple alleles can be relatively frequently observed in the population, and thus “mismatch” in this context requires the artificial adoption of one sequence as a standard. Thus, for the purposes of this invention, sequences are referred to herein as a “perfect match”, or a “mismatch” or a “mutant” with respect to a wild type sequence standard. “Mismatches” sometimes also refer to “allelic variants”. The term “allele”, which is used interchangeably herein with “allelic variant” refers to alternative forms of a gene or portions thereof. Alleles occupy the same locus or position on homologous chromosomes. When a subject has two identical alleles of a gene, the subject is said to be “homozygous” for the gene or allele. When a subject has two different alleles of a gene, the subject is said to be “heterozygous” for the gene. Alleles of a specific gene can differ from each other in a single nucleotide, or several nucleotides, and can include substitutions, deletions, and insertions of nucleotides. An allele of a gene can also be a form of a gene containing a mutation. The term “allelic variant of a polymorphic region of a gene” refers to a region of a gene having one of several nucleotide sequences found in that region of the gene in
other individuals of the same species. Thus the above terms have to be used in context to derive the full meaning of the term.


[0067] As will be appreciated by those in the art, all of these nucleic acid analogs may find use in the present invention. In addition, mixtures of naturally occurring nucleic acids and analogs can be made. Alternatively, mixtures of different nucleic acid analogs, and mixtures of naturally occurring nucleic acids and analogs may be made.

[0068] The nucleic acids may be ‘single stranded’ or ‘double stranded’, as specified, or contain portions of both double stranded or single stranded sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA or a hybrid, where the nucleic acid contains any combination of deoxyribo- and ribo-nucleotides, and any combination of bases, including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine, iso guanine, etc. A preferred embodiment utilizes isocytosine and isoguanine in nucleic acids designed to be complementary to other probes, rather than target sequences, as this reduces non-specific hybridization, as is generally described in U.S. Pat. No. 5,681,702. As used herein, the term “nucleoside” includes nucleotides as well as nucleoside and nucleotide analogs, and modified nucleosides such as amino modified nucleosides. In addition, “nucleoside” includes non-naturally occurring analog structures. Thus for example the individual units of a peptide nucleic acid, each containing a base, are referred to herein as a nucleoside.

[0069] If required, the target sequence is prepared using known techniques. For example, the sample may be treated to lyse the cells, using known lysis buffers, electroporation, etc., with purification and/or amplification as needed, as will be appreciated by those in the art. Suitable amplification techniques to amplify target nucleic acids are outlined in PCT US99/01705, hereby expressly incorporated by reference and outlined below.

[0070] General techniques for target sequence amplification are discussed below. The primers or probes used herein permit amplification of the target DNA, as is well understood in the art.

[0071] Target amplification involves the amplification (replication) of the target sequence such that the number of copies of the target sequence is increased. Suitable target amplification techniques include, but are not limited to, the polymerase chain reaction (PCR), allele PCR, strand dis-placement amplification (SDA), and nucleic acid sequence based amplification (NASBA), oligonucleotide ligation assay (OLA), rolling circle amplification (RCA), Invader™ technology, using scissile primers (CPT) etc., defined below.

[0072] Different amplification techniques may have further requirements of primers or probes, as is more fully described below. The size of the primer nucleic acid may vary, as will be appreciated by those in the art, in general varying from 5 to 500 nucleotides in length, depending on the use and amplification technique.

[0073] Once the complex between the primer and the target sequence has been formed, an enzyme, sometimes termed an “amplification enzyme”, is used to modify the primer. As for all the methods outlined herein, the enzymes may be added at any point during the assay, either prior to, during, or after the addition of the primers. The identification of the enzyme will depend on the amplification technique used, as is more fully outlined below.

[0074] In a preferred embodiment, the target amplification technique is PCR. The polymerase chain reaction (PCR) is widely used and described, and involve the use of primer extension combined with thermal cycling to amplify a target sequence; see U.S. Pat. Nos. 4,683,195 and 4,683,202, and PCR Essential Data, J. W. Wiley & sons, Ed. C. R. Newton, 1995, all of which are incorporated by reference. In addition, there are a number of variations of PCR which can be used in the invention, including “quantitative competitive PCR” or “Q-PCR”, “arbitrarily primed PCR” or “AP-PCR”,

US 2004/0229222 A1
Nov. 18, 2004
“immuno-PCR”, “Alu-PCR”, “PCR single strand conformational polymorphism” or “PCR-SSCP”, “reverse transcriptase PCR” or “RT-PCR”, “biotin capture PCR”, “vectorette PCR”, “panhandle PCR”, and “PCR select cDNA subtraction”, among others. The particulars of PCR are well known, and include the use of a thermostable polymerase such as Taq I polymerase and thermal cycling. Accordingly, the PCR reaction requires at least one PCR primer and a polymerase.

[0075] Multiplex PCR amplification of SNP loci with subsequent hybridization to oligonucleotide arrays has been shown to be an accurate and a reliable method of simultaneously genotyping at least hundreds of SNPs; see Wang et al., Science, 280:1077 (1998); see also Schafer et al., Nature Biotechnology 16:33-39 (1998). Multiplex PCR reactions involving different primer sets for different regions of the target DNA can be used for amplifying target DNA. In a preferred embodiment, a 7-multiplex PCR reaction is performed to amplify seven different regions, each region corresponding to a different SNP position of the target DNA, using appropriate primer sets. Amplicons are pooled and the total pool is used for analysis. Thus, for example, if 10 SNP positions were amplified by PCR, 7×10=70 SNP positions can be interrogated at a time for each amplified DNA sample; for example, on a 450 chip.

[0076] An additional amplification method is allelic PCR, described in Newton et al., Nucl. Acid Res. 17:2503 (1989), hereby expressly incorporated by reference. Allelic PCR allows single base discrimination based on the fact that the PCR reaction does not proceed well if the terminal 3'-nucleotide is mismatched, assuming the DNA polymerase being used lacks a 3'-exonuclease proofreading activity.

[0077] In a preferred embodiment, the target amplification technique is SDA. Strand displacement amplification (SDA) is generally described in Walker et al., in Molecular Methods for Virus Detection, Academic Press, Inc., 1995, and U.S. Pat. Nos. 5,455,166 and 5,130,238, all of which are hereby expressly incorporated by reference in their entirety. In general, SDA may be described as follows. A single stranded target nucleic acid, usually a DNA target sequence, is contacted with an SDA primer generally with a length of 25-100 nucleotides. An SDA primer is substantially complementary to a region at the 3' end of the target sequence, and the primer has a sequence at its 5' end (outside of the region that is complementary to the target) that is a recognition sequence for a restriction endonuclease, sometimes referred to herein as a “nicking enzyme” or a “nicking endonuclease”. The SDA primer then hybridizes to the target sequence. The SDA reaction mixture also contains a polymerase (an “SDA polymerase”) and a mixture of all four deoxynucleoside-triphosphates (also called deoxynucleotides or dNTPs, i.e. dATP, dTTP, dCTP and dGTP), at least one species of which is a substituted or modified dNTP; thus, the SDA primer is modified, i.e. extended, to form a modified primer, sometimes referred to herein as a “newly synthesized strand”. The substituted dNTP is modified such that it will inhibit cleavage in the strand containing the substituted dNTP but will not inhibit cleavage on the other strand. Examples of suitable substituted dNTPs include, but are not limited, 2'deoxyadenosine 5'-O-(1-thiotriphosphate), 5-methyldioxycytidine 5'-triphosphate, 2'deoxyuridine 5'-triphosphate, and 2'deoxyguanosine 5'-triphosphate. In addition, the substitution of the dNTP may occur after incorporation into a newly synthesized strand; for example, a methylyase may be used to add methyl groups to the synthesized strand. In addition, if all the nucleotides are substituted, the polymerase may have 5'-3' exonuclease activity. However, if less than all the nucleotides are substituted, the polymerase preferably lacks 5'-3' exonuclease activity. A chart depicting suitable enzymes, and their corresponding recognition sites and the modified dNTP to be used is found in U.S. Pat. No. 5,455,166, hereby expressly incorporated by reference.

[0078] Accordingly, the SDA reaction requires, in no particular order, an SDA primer, an SDA polymerase, a nicking endonuclease, and dNTPs, at least one species of which is modified.

[0079] In general, SDA does not require thermocycling. The temperature of the reaction is generally set to be high enough to prevent non-specific hybridization but low enough to allow specific hybridization; this is generally from about 37° C. to about 42° C., depending on the enzymes.

[0080] In a preferred embodiment, the target amplification technique is nucleic acid sequence based amplification (NASBA). NASBA is generally described in U.S. Pat. No. 5,409,818; Soeknunan et al., Nucleic Acid Sequence-Based Amplification, Ch. 12 (pp. 261-285) of Molecular Methods for Virus Detection, Academic Press, 1995; and “Profiling from Gene-based Diagnostics”, CIB International Publishing Inc., N.J., 1996, all of which are incorporated by reference. NASBA is very similar to both TMA and QBR. Transcription mediated amplification (TMA) is generally described in U.S. Pat. Nos. 5,399,491, 5,888,779, 5,705,365, 5,710,029, all of which are incorporated by reference. The main difference between NASBA and TMA is that NASBA utilizes the addition of RNase H to effect RNA degradation, and TMA relies on inherent RNase H activity of the reverse transcriptase.

[0081] In general, these techniques may be described as follows. A single stranded target nucleic acid, usually an RNA target sequence (sometimes referred to herein as “the first target sequence” or “the first template”), is contacted with a first primer, generally referred to herein as a “NASBA primer” (although “TMA primer” is also suitable). Starting with a DNA target sequence is described below. These primers generally have a length of 25-100 nucleotides, with NASBA primers of approximately 50-75 nucleotides being preferred. The first primer is preferably a DNA primer that has at its 3' end a sequence that is substantially complementary to the 3' end of the first template. The first primer also has an RNA polymerase promoter at its 5' end (or its complement (antisense), depending on the configuration of the system). The first primer is then hybridized to the first template to form a first hybridization complex. The reaction mixture also includes a reverse transcriptase enzyme (an “NASBA reverse transcriptase”) and a mixture of the four dNTPs, such that the first NASBA primer is modified, i.e. extended, to form a modified first primer, comprising a hybridization complex of RNA (the first template) and DNA (the newly synthesized strand).

[0082] By “reverse transcriptase” or “RNA-directed DNA polymerase” herein is meant an enzyme capable of synthesizing DNA from a DNA primer and an RNA template. Suitable RNA-directed DNA polymerases include, but are not limited to, avian myeloblastosis virus reverse tran-
scriptase ("AMV RT") and the Moloney murine leukemia virus RT. When the amplification reaction is TMA, the reverse transcriptase enzyme further comprises a RNA degrading activity.

[0083] In a preferred embodiment, the "reverse transcriptase" is a modified reverse transcriptase (RT) in which the DNA-dependent DNA polymerase activity has been eliminated. Here, self-extending probes are not extended by the modified RT because these are DNA-dependent extensions and only probe bound to RNA targets will be extended by the RNA-dependent DNA polymerase activity.

[0084] In addition to the components listed above, the NASBA reaction also includes an RNA degrading enzyme, also sometimes referred to herein as a ribonuclease, that will hydrolyze RNA of an RNA:DNA hybrid without hydrolyzing single- or double-stranded RNA or DNA. Suitable ribonucleases include, but are not limited to, RNase H from E. coli and calf thymus.

[0085] The ribonuclease activity degrades the first RNA template in the hybridization complex, resulting in a dissociation of the hybridization complex leaving a first single stranded newly synthesized DNA strand, sometimes referred to herein as "the second template".

[0086] In addition, the NASBA reaction also includes a second NASBA primer, generally comprising DNA (although as for all the probes herein, including primers, nucleic acid analogs may also be used). This second NASBA primer has a sequence at its 3′ end that is substantially complementary to the 3′ end of the second template, and also contains an antisense sequence for a functional promoter and the antisense sequence of a transcription initiation site. Thus, this primer sequence, when used as a template for synthesis of the third DNA template, contains sufficient information to allow specific and efficient binding of an RNA polymerase and initiation of transcription at the desired site. Preferred embodiments utilize the antisense promoter and transcription initiation site that of the T7 RNA polymerase, although other RNA polymerase promoters and initiation sites can be used as well, as outlined below.

[0087] The second primer hybridizes to the second template, and a DNA polymerase, also termed a "DNA-directed DNA polymerase", also present in the reaction, synthesizes a third template (a second newly synthesized DNA strand), resulting in second hybridization complex comprising two newly synthesized DNA strands.

[0088] Finally, the inclusion of an RNA polymerase and the required four ribonucleoside triphosphates (ribonucleotides or NTPs) results in the synthesis of an RNA strand (a third newly synthesized strand that is essentially the same as the first template). The RNA polymerase, sometimes referred to herein as a "DNA-directed RNA polymerase", recognizes the promoter and specifically initiates RNA synthesis at the initiation site. In addition, the RNA polymerase preferably synthesizes several copies of RNA per DNA duplex. Preferred RNA polymerases include, but are not limited to, T7 RNA polymerase, and other bacteriophage RNA polymerases including those of phage T3, phage qIII, Salmonella phage sp6, or Pseudomonas phase gh-1.

[0089] In some embodiments, TMA and NASBA are used with starting DNA target sequences. In this embodiment, it is necessary to utilize the first primer comprising the RNA polymerase promoter and a DNA polymerase enzyme to generate a double stranded DNA hybrid with the newly synthesized strand comprising the promoter sequence. The hybrid is then denatured and the second primer added.

[0090] Accordingly, the NASBA reaction requires, in no particular order, a first NASBA primer, a second NASBA primer comprising an antisense sequence of an RNA polymerase promoter, an RNA polymerase that recognizes the promoter, a reverse transcriptase, a DNA polymerase, an RNA degrading enzyme, NTPs and dNTPs, in addition to the detection components outlined below.

[0091] These components result in a single starting RNA template generating a single DNA duplex; however, since this DNA duplex results in the creation of multiple RNA strands, which can then be used to initiate the reaction again, amplification proceeds rapidly.

[0092] Accordingly, the TMA reaction requires, in no particular order, a first TMA primer, a second TMA primer comprising an antisense sequence of an RNA polymerase promoter, an RNA polymerase that recognizes the promoter, a reverse transcriptase with RNA degrading activity, a DNA polymerase, NTPs and dNTPs, in addition to the detection components outlined below.

[0093] These components result in a single starting RNA template generating a single DNA duplex; however, since this DNA duplex results in the creation of multiple RNA strands, which can then be used to initiate the reaction again, amplification proceeds rapidly.

[0094] In a preferred embodiment, the target amplification technique is the oligonucleotide ligation assay (OLA), sometimes referred to as the ligation chain reaction (LCR). The method can be run in two different ways; in a first embodiment, only one strand of a target sequence is used as a template for ligation (OLA); alternatively, both strands may be used (OLA). Oligonucleotide ligation amplification ("OLA", sometimes referred to herein as the ligation chain reaction (LCR)) involves the ligation of two smaller probes into a single long probe, using the target sequence as the template. See generally U.S. Pat. Nos. 5,185,243 5,679,524 and 5,573,907; EP 0 320 308 B1; EP 0 336 731 B1; EP 0 439 182 B1; WO 90/01069; WO 89/12696; and WO 97/31256, WO 89/09835, and U.S. S. Nos. 60/078,102 and 60/073,011, all of which are incorporated by reference.

[0095] A variation of LCR utilizes a "chemical ligation" of sorts, as is generally outlined in U.S. Pat. Nos. 5,616,464 and 5,767,259, both of which are hereby expressly incorporated by reference in their entirety. In this embodiment, similar to LCR, a pair of primers are utilized, wherein the first primer is substantially complementary to a first domain of the target and the second primer is substantially complementary to an adjacent second domain of the target (although, as for LCR, if a "gap" exists, a polymerase and dNTPs may be added to "fill in" the gap). Each primer has a portion that acts as a "side chain" that does not bind the target sequence and acts one half of a stem structure that interacts non-covalently through hydrogen bonding, salt bridges, van der Waal’s forces, etc. Preferred embodiments utilize substantially complementary nucleic acids as the side chains. Thus, upon hybridization of the primers to the target sequence, the side chains of the primers are brought into spatial proximity, and, if the side chains comprise nucleic acids as well, can also form side chain hybridization complexes.
[0096] At least one of the side chains of the primers comprises an activatable cross-linking agent, generally covalently attached to the side chain, that upon activation, results in a chemical cross-link or chemical ligation. The activatable group may comprise any moiety that will allow cross-linking of the side chains, and include groups activated chemically, photoinitically and thermally, with photoactivatable groups being preferred. In some embodiments a single activatable group on one of the side chains is enough to result in cross-linking via interaction to a functional group on the other side chain; in alternate embodiments, activatable groups are required on each side chain.

[0097] Once the hybridization complex is formed, and the cross-linking agent has been activated such that the primers have been covalently attached, the reaction is subjected to conditions to allow for the dissociation of the hybridization complex and the release of the target to serve as a template for the next ligation or cross-linking. In this way, signal amplification occurs, and can be detected as outlined herein.

[0098] In a preferred embodiment the target amplification technique is RCA. A variation of OLA which can also be used for genotyping is termed “rolling circle amplification” or RCA. Rolling circle amplification utilizes a single probe that hybridizes to a target such that each terminus of the probe hybridizes adjacent to each other (or, alternatively, the intervening nucleotides can be “filled in” using a polymerase and dNTPs). Then, upon ligation of the two termini of the probe, a circular probe is formed, also referred to as a “padlock probe” or the “RCA probe”. Then, a primer and a polymerase is added such that the primer sequence is extended. But as the circular probe has no terminus, the polymerase repeatedly extends the circular probe resulting in concatamers of the circular probe. As such, the probe is amplified. The resultant amplicon may be cleaved in a variety of ways for further use in assays. Rolling-circle amplification is generally described in Bander et al. (1998) Nuc. Acids Res. 26:5073-5078; Barany, F. (1991) Proc. Natl. Acad. Sci. USA 88:189-193, Lazzari et al. (1998) Nat. Genet. 19:225-232; Zhang et al., Gene 211:277 (1998) and Daubendiek et al., Nature Biotech. 15:273 (1997); all of which are incorporated by reference in their entirety.

[0099] In a preferred embodiment, the RCA probes comprise a cleavage site, such that either after or during the rolling circle amplification, the RCA concatamer may be cleaved into amplicons. In some embodiments, this facilitates the detection, since the amplicons are generally smaller and exhibit favorable hybridization kinetics on the surface. As will be appreciated by those in the art, the cleavage site can take on a number of forms, including, but not limited to, the use of restriction sites in the probe, the use of ribozyme sequences, or through the use or incorporation of nucleic acid cleavage moieties.

[0100] In a preferred embodiment, the padlock probe or RCA probe contains a restriction site. The restriction endonuclease site allows for cleavage of the long concatamers that are typically the result of RCA into smaller individual units that hybridize either more efficiently or faster to surface bound capture probes. Thus, following RCA (or in some cases, during the reaction), the product nucleic acid is contacted with the appropriate restriction endonuclease. This results in cleavage of the product nucleic acid into smaller fragments. The fragments are then hybridized with the capture probe that is immobilized resulting in a concentration of product fragments onto the detection electrode.

[0101] In a preferred embodiment, the cleavage site is a ribozyme cleavage site as is generally described in Daubendiek et al., Nature Biotech. 15:273 (1997), hereby expressly incorporated by reference. In this embodiment, by using RCA probes that encode catalytic RNAs, NTPs and an RNA polymerase, the resulting concatamer can self-cleave, ultimately forming monomeric amplicons.

[0102] In a preferred embodiment, cleavage is accomplished using DNA cleavage reagents. For example, as is known in the art, there are a number of intercalating moieties that can effect cleavage, for example using light.

[0103] Thus, in a preferred embodiment the OLA/RCA is performed in solution followed by restriction endonuclease cleavage of the RCA product. The cleaved product is then applied to an array as described herein. The incorporation of an endonuclease site allows the generation of short, easily hybridizable sequences. Furthermore, the unique capture sequence in each rolling circle padlock probe sequence allows diverse sets of nucleic acid sequences to be analyzed in parallel on an array, since each sequence is resolved on the basis of hybridization specificity.

[0104] In a preferred embodiment, the signal amplification technique is CPT. CPT technology is described in a number of patents and patent applications, including U.S. Pat. Nos. 5,011,769, 5,403,711, 5,660,988, and 4,876,187, and PCT published applications WO 95/05480, WO 95/1416, and WO 95/00667, and U.S. Ser. No. 09/014,304, all of which are expressly incorporated by reference in their entirety. A CPT primer (also sometimes referred to herein as a “scissile primer”), comprises two probe sequences separated by a scissile linkage. The CPT primer is substantially complementary to the target sequence and thus will hybridize to it to form a hybridization complex. The scissile linkage is cleaved, without cleaving the target sequence, resulting in the two probe sequences being separated. The two probe sequences can thus be more easily dissociated from the target, and the reaction can be repeated any number of times. The cleaved primer is then detected. By “scissile linkage” herein is meant a linkage within the scissile probe that can be cleaved when the probe is part of a hybridization complex, that is, when a double-stranded complex is formed. It is important that the scissile linkage cleave only the scissile probe and not the sequence to which it is hybridized (i.e. either the target sequence or a probe sequence), such that the target sequence may be reused in the reaction for amplification of the signal.

[0105] In a preferred embodiment, Invader™ technology is used. Invader™ technology is based on structure-specific polymerases that cleave nucleic acids in a site-specific manner. Two probes are used: an “invader” probe and a “signaling” probe, that adjacentably hybridize to a target sequence with a non-complementary overlap. The enzyme cleaves at the overlap due to its recognition of the “tail”, and releases the “tail”. This can then be detected. The Invader™ technology is described in U.S. Pat. Nos. 5,846,717; 5,614,402; 5,719,028; 5,541,311; and 5,843,669, all of which are hereby incorporated by reference.

[0106] By “extension enzyme” herein is meant an enzyme that will extend a sequence by the addition of NTPs. As is
well known in the art, there are a wide variety of suitable extension enzymes, of which polymerases (both RNA and DNA, depending on the composition of the target sequence and probe) are preferred. Preferred polymerases are those that lack strand displacement activity, such that they will be capable of adding only the necessary bases at the end of the probe, without further extending the probe to include nucleotides that are complementary to a targeting domain and thus preventing circularization. Suitable polymerases include, but are not limited to, both DNA and RNA polymerases, including the Klenow fragment of DNA polymerase I, SEQUENASE 1.0 and SEQUENASE 2.0 (U.S. Biochemical), T5 DNA polymerase, Ph29 DNA polymerase and various RNA polymerases such as from Thermus sp., or Q beta replicase from bacteriophage, as well as T4 and T7 RNA polymerases can be used, among others.

[0107] Even more preferred polymerases are those that are essentially devoid of a 5’ to 3’ exonuclease activity, so as to assure that the probe will not be extended past the 5’ end of the probe. Exemplary enzymes lacking 5’ to 3’ exonuclease activity include the Klenow fragment of the DNA polymerase and the Stoffel fragment of DNAShm Polymerase. For example, the Stoffel fragment of Taq DNA polymerase lacks 5’ to 3’ exonuclease activity due to genetic manipulations, which result in the production of a truncated protein lacking the N-terminal 289 amino acids. (See e.g., Lawyer et al., J. Biol. Chem., 264:6427-6437 [1989] and Lawyer et al., PCR Meth. Appl., 2:275-287 [1993]). Analogous mutant polymerases have been generated for polymerases derived from I. maritima, Tsp41, T7, T205, 1th and Taf.

[0108] In the above embodiments, the polymerases can be any polymerase with unique features as explained in each case as outlined herein, preferably one lacking 3’ exonuclease activity (3’ exo’).

[0109] Examples of suitable polymerase include but are not limited to exonuclease minus DNA Polymerase I large (Klenow) Fragment, Ph29 DNA polymerase, Taq DNA Polymerase, Deep vent (exo’), thermostable and the like. In addition, in some embodiments, a polymerase that will replicate single-stranded DNA (i.e. without a primer forming a double stranded section] can be used.

[0110] In a preferred embodiment, the polymerase creates more than 100 copies of the circular DNA. In more preferred embodiments the polymerase creates more than 1000 copies of the circular DNA; while in a most preferred embodiment the polymerase creates more than 10,000 copies or more than 50,000 copies of the template, thus amplifying the target sequence.

[0111] In certain preferred embodiments, terminal transferase can be used to add nucleotides comprising separation labels such as biotin to any linear molecules, and then the mixture run through a streptavidin system to remove any linear nucleic acids, leaving only the closed circular probes. For example, as in the RCA method, when genomic DNA is used as the target, the DNA may be biotinylated using a variety of techniques, and precircle probes added and circularized. Since the circularized probes are catalyzed on the genomic DNA, the linear unreacted precircle probes can be washed away. The closed circle probes can then be cleaved, such that they are removed from the genomic DNA, collected and amplified.

[0112] Thus, using the amplification methods outlined above, a number of target molecules are made for hybridization to probes in the assays of the invention. Generally, the amplification steps are repeated for a period of time to allow a number of cycles, depending on the number of copies of the probe, original target sequence and the sensitivity of detection, with cycles ranging from 1 to thousands, with from 10 to 100 cycles being preferred and from 20 to 50 cycles being especially preferred. As is more fully outlined below, the products of these reactions can be detected in a number of ways, as is generally outlined in U.S. Ser. Nos. 09/28,553; 09/458,501; 09/572,187; 09/495,992; 09/341,217; WO00/31148; Ser. Nos. 09/439,889; 09/438,209; 09/344,620; PCTUS00/17422; Ser. No. 09/478,727, all of which are expressly incorporated by reference in their entirety. Also, when the binding ligand or probe is a nucleic acid, preferred compositions and techniques outlined in U.S. Pat. Nos. 5,591,578; 5,824,473; 5,705,348; 5,780,234 and 5,770,369; U.S. Ser. Nos. 08/873,598 08/911,589; WO 99/10035; WO98/12430; WO98/51758; WO 99/160897; WO99/57317; WO99/67425; WO00/24941; PCT US00/100; WO00/38836; WO99/37819; WO99/37319 and PCTUS00/20476; and related materials, are expressly incorporated by reference in their entirety.

[0113] In a preferred embodiment, the amplification technique is signal amplification. Signal amplification involves the use of limited number of target molecules as templates to either generate multiple signaling probes or allow the use of multiple signaling probes. Signal amplification strategies include PCR, CPT, Invader™, and the use of amplification probes in sandwich assays.

[0114] In most cases, double stranded target nucleic acids are denatured to render them single stranded so as to permit hybridization of the probes described below. For denaturing the target DNA, preferred embodiment utilizes a thermal step, generally by raising the temperature of the reaction to about 95°C, although pH changes and other techniques such as the use of extra probes or nucleic acid binding proteins may also be used.

[0115] Amplified target DNA is then contacted with the various types of probes described below to form a hybridization complex.

[0116] By “probe nucleic acid” is meant an oligonucleotide that will hybridize to some portion or a domain of the target sequence. The probe is also referred to as a “capture probe”, “SBE probe” or a “microarray probe” or sometimes, an “extension probe”. Depending on whether the probe is complementary to the WT or the mutant target sequence, the probe may sometimes be referred to as the “WT probe” or the “mutant probe”. Probes of the present invention are designed to be complementary to a target sequence or an amplicon of the target sequence such that hybridization of the target sequence and the probes of the present invention occurs. As is outlined above, this complementarity need not be perfect, that is, there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the capture probes of the present invention. In preferred embodiments, the capture probe is single stranded. In another preferred embodiment, the capture probe is modified, as described below, wherein the probe is referred to as a “modified capture probe” or a “modified probe”. As is more fully outlined below, in some embodiments, the capture probe comprises additional bases (usually one additional base) to prevent self-extension.
When the probe or primer is modified to prevent self-extension, it is referred to herein as a “non-self extension probe”.

[0117] In a preferred embodiment, contacting is done to probes attached to a biochip and are designed to be “substantially complementary” to a target sequence, such that hybridization of the target sequence and the probes of the present invention occurs. As outlined below, this complementarity need not be perfect; there may be any number of base pair mismatches which will interfere with hybridization between the target sequence and the capture probe of the present invention. However, if the number of mutations is so great that no hybridization can occur under even the least stringent of hybridization conditions, the sequence is not a complementary target sequence. Thus, by “substantially complementary” herein is meant that capture probes are sufficiently complementary to the target sequences to hybridize under normal reaction conditions, particularly high stringency conditions, as outlined herein. The term “complementary”, in the context of a nucleic acid sequence, means a nucleic acid sequence having a sequence relationship to a second nucleic acid sequence such that there is perfect alignment of Watson-Crick base pairs along the entire length of both nucleic acid sequences.

[0118] A variety of hybridization conditions may be used in the present invention, typically classified by the degree of “stringency” of the conditions, including high, moderate and low stringency conditions; see for example Maniatis et al., Molecular Cloning: A Laboratory Manual, 2d Ed., 1989, and Short Protocols in Molecular Biology, ed. Ausubel, et al., hereby incorporated by reference. Stringent conditions are sequence-dependent and will be different in different circumstances. Stringency can be controlled by altering a step parameter that is a thermodynamic variable, including, but not limited to, temperature, formamide concentration, salt concentration, chaotropic salt concentration pH, organic solvent concentration, etc. These parameters may also be used to control non-specific binding, as is generally outlined in U.S. Pat. No. 5,681,697. Thus it may be desirable to perform certain steps at higher stringency conditions to reduce non-specific binding.

[0119] The Tm is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at Tm, 50% of the probes are occupied at equilibrium). For example, “maximum stringency” typically occurs at about Tm-5°C (50 below the Tm of the probe); “high stringency” at about 5-10° below the Tm; “intermediate stringency” at about 10-20° below the Tm of the probe; and “low stringency” at about 20-25° below the Tm. In general, hybridization conditions are carried out under high ionic strength conditions, for example, using 6xSSC or 6xSSPE. Under high stringency conditions, hybridization is followed by two washes with low salt solution, for example 0.5xSSC, at the calculated temperature. Under medium stringency conditions, hybridization is followed by two washes with medium salt solution, for example 2xSSC. Under low stringency conditions, hybridization is followed by two washes with high salt solution, for example 6xSSC. Functionally, maximum stringency conditions may be used to identify nucleic acid sequences having strict identity or near-strict identity with the hybridization probe; while high stringency conditions are used to identify nucleic acid sequences having about 80% or more sequence identity with the probe. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes, “Overview of principles of hybridization and the strategy of nucleic acid assays” (1993).

[0120] Thus, stringent conditions will be those in which the salt concentration is less than about 1.0 sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g. 10 to 50 nucleotides) and at least about 60°C for long probes (e.g. greater than 50 nucleotides).

[0121] Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. The hybridization conditions may also vary when a non-ionic backbone, i.e. PNA is used, as is known in the art. In addition, cross-linking agents may be added after target binding to cross-link, i.e. covalently attach, the two strands of the hybridization complex. Preferred embodiments of the present invention may use “competitors” on arrays, to reduce non-specific binding.

[0122] Thus, the assays of the present invention are generally run under stringency conditions which allows formation of the hybridization complex only in the presence of target, but occasionally, less stringent conditions may be used. For example, enzymatic extension may require less stringent hybridization conditions as assay selectivity is enhanced by the enzyme, as is generally understood in the art. Details of conditions used in preferred embodiments of this invention are further described below. Furthermore, preferred embodiments use arrays such that the capture probes bind to the target DNA in a sequence-specific manner, and permits the unbound material to be washed away.

[0123] Preferred embodiment of the present invention uses the single base extension (SBE; sometimes also called minisequencing) assay and its variations, described herein. In a general SBE assay, two capture probes (wild type (WT) and mutant) that differ only in their terminal bases are synthesized such that, one of the two capture probes is perfectly complementary to the “queried base” or the “interrogation position” of the target sequence to be analyzed. That is, the capture probes include the interrogation or SNP base. Under normal hybridization conditions, both the WT and the mutant probe can hybridize to the target sequence. But enzymes such as ligases or polymerases, defined above, can distinguish a perfect or an imperfect match such that, enzymatic extension by a single chain terminating base (sometimes carrying the label for detection), occurs only if a perfect duplex is present. That is, the DNA polymerase can extend the capture probe by one base in the presence of four labeled-terminator nucleotides, only if there is perfect hybridization. Else, a 3' base mismatch results in breathing and dissolves primer extension. By “terminator nucleotides” is meant that the nucleotide is derivatized such that no further extensions can occur, so only one nucleotide is added. Preferred embodiments utilize dideoxy-triphosphate nucleotides (ddNTPs) and generally a set of nucleotides comprising ddATP, ddCTP, ddGTP and ddTTP is used, at least one of which includes a label, and preferably all four.
The labels may be different or the same, as specified. Once the labeled nucleotide is added, detection of the label proceeds as outlined below. See generally Sylven et al., Genomics 8:684-92 (1990); U.S. Pat. Nos. 5,846,710 and 5,888,819; Pastinen et al., Genomics Res. 7(6):606-14 (1997); all of which are expressly incorporated herein by reference. Thus, the identity of the SNP or queried base in the target is determined by the probe set that is extended by the DNA polymerase. In a variation of this type of assay, the probes terminate at a position one base upstream of the queried base in the target.

[0124] A limitation of the SBE method is that unless the target nucleic acid is in sufficient concentration, the amount of unextended primer in the reaction greatly exceeds the resultant extended-labeled primer. The excess of unextended primer competes with the detection of the labeled primer in the assays described herein. Accordingly, when SBE is used, preferred embodiments utilize methods for the removal of unextended primers as outlined herein.

[0125] One method to overcome this limitation is ther- mcycling minisequencing in which repeated cycles of annealing, primer extension, and heat denaturation using a thermocycler and theremo-stable polymerase allows the amplification of the extension probe which results in the accumulation of extended primers. For example, if the original unextended primer to target nucleic acid concentra- tion is 100:1 and 100 thermocycles and extensions are performed, a majority of the primer will be extended.

[0126] As will be appreciated by those in the art, the configuration of the SBE system can take on several forms. The SBE reaction may be done in solution, and then the newly synthesized strands, with the base-specific detectable labels, can be detected. For example, they can be directly hybridized to capture probes that are complementary to the extension primers, and the presence of the label is then detected.

[0127] Alternatively, the SBE reaction can occur on a surface. For example, a target nucleic acid may be captured using a first capture probe that hybridizes to a first target domain of the target, and the reaction can proceed at a second target domain. The extended labeled primers are then bound to a second capture probe and detected.

[0128] Thus, the SBE reaction requires, in no particular order, an extension primer, a polymerase and dNTPs, at least one of which is labeled.


[0130] Specificity remains a problem in many currently available SBE assays. The extent of molecular complementarity between probe and target defines the specificity of the interaction. Variations in the concentrations of probes, of targets and of salts in the hybridization medium, in the reaction temperature, and in the length of the probe may alter or influence the specificity of the probe/target interaction.

[0131] Two major hurdles for highly parallel screening of SNPs on microarrays are 1) the necessity to amplify DNA regions spanning the SNPs by PCR to achieve sufficient sensitivity and specificity of detecting a single-base variation in the complex human genome in a reproducible way; and, 2) the ability to distinguish unequivocally between homozygous and heterozygous allicic variants in the diploid human genome. Differential hybridization with allele-spe- cific oligonucleotide (ASO) probes is most commonly used in the microarray format (Pastinen et al., Genome Res. 10:1031-42 (2000) hereby expressly incorporated by reference).

[0132] A problem associated with the SBE assay is that, some probes used in the assay form internal "stem-loop structures" resulting in target-independent "self-extension" of the probe, thus giving a false positive signal that interferes with the determination of the SNP base. To overcome these problems, this invention recites several preferred embodi- ments described herein involving novel designs of capture probes, and/or, the use of a novel combination of methods involving unplexed target preparation and primer extension with modified polymerases to avoid self-extension.

[0133] That is, the present invention provides methods for preventing self extension of primers on an array. The present invention also provides an array with modified primers. Preferably the primers are modified as described herein. In some embodiments, the arrays include primers that include modified nucleotides. Modified nucleotides include exo- cyclic amine modified bases like 2-thio thymine, 2-amino adenine, amine modified cytosine, amine modified guanine, or terminator bases like 4-methylindole. In some methods of the invention, the modified nucleotides alter the polymerase binding or protein binding to the stem region of the non self-extension probe, wherein the modified nucleotides are present. Modified nucleotides also comprises a sugar and phosphate modifications.

[0134] Some preferred phosphate modifications include, but are not limited to phosphorothioates, phosphoramidates, methyl phosphonates, methyl phosphates, H-phosphonates.

[0135] In another preferred embodiment, to prevent self- extension, short complementary oligonucleotides are used whereby self-extension is inhibited.

[0136] Details of the modified nucleotides used in preferred embodiments of the invention is described herein.

[0137] In one preferred embodiment, the capture probe is designed such that it terminates at one or more bases downstream of the queried SNP site in the probe. Generally, a stem of three or more base pairs is required for self- extension. The bases added downstream are designed such that they are complementary to the bases in the target but are mismatched at the 3' end of the stem-loop structure.

[0138] If there is a mismatch between bases in the stem- loop, a bubble formed which induces breathing and prevents the DNA polymerase from extending the capture probe thus preventing false signal detection. In practice, adding four or more bases is not feasible since the SNP position will be too
far upstream and may induce short duplex formation and polymerase dependent extension and thus, a false positive result. In a preferred embodiment, the number of additional bases added downstream of the SNP site on the probe is one, or two or three bases. Target dependent extension is not affected by the additional bases as the added bases are complementary to the target. Accordingly, the interrogation position is the 3' terminal nucleotide. Alternatively, the interrogation position is the penultimate nucleotide of the primer.

[0139] In another preferred embodiment to destabilize hairpin structures in probes, “modified base pairs” are incorporated into the probe that prevent self-extension but do not interfere with target hybridization. Such modified base pairs have been used before in applications of gene-therapy wherein each modified base of the “pair” are found in each complementary single stranded nucleic acid involved in gene therapy. Such modified base pairs include but are not limited to a) 2-amino-A:2-thio-T, b) 2-aminopurine:2-thio-T, c) 6-thio-G, d) 2-thio-C, e) hydrophobic bases such as 4-methylindole, difluorotoluene, etc.

[0140] In another preferred embodiment, the capture probe comprises modified phosphates or sugars in the stem region of the probe to reduce stem-loop structures and thus false positives. It is well known in the art that the phosphate backbone of the nucleic acid plays an important role in nucleic acid-protein interactions. Electrostatic interactions between the positively charged amino acids of a protein and the negatively charged phosphate backbone and the formation of hydrogen bonds between the phosphate oxygen and protein contribute to the binding affinity of a particular protein or enzyme to a nucleic acid. In this preferred embodiment, by modifying the phosphate or the sugar rings of the probe backbone, one can decrease the binding affinity of the polymerase enzyme for the stem region thereby decreasing the probability of non-specific nucleotide incorporation. Such modifications are not performed at the queued base of the probe and hence, the thermodynamic stability of the probe-target duplex is not altered. Examples of such modified nucleic acids are described below.

[0141] In the current invention, several surface charge modifications for probes are proposed herein, including, but not limited to, phosphorothionate, phosphoramidate, methyl phosphate and methyl phosphate modifications of the phosphate backbone and 2' O-methyl modifications on the sugar ring. Phosphorothionate (sulfur substitution) and phosphoramidate (nitrogen substitution) substitutions alter the charge distribution, hydrophobicity and the ability for an enzyme to form efficient hydrogen bonding with the phosphate backbone. Methyl phosphate and methyl phosphate eliminate the phosphate charge altogether, thus inhibiting the binding of the enzyme to the DNA altogether (see Smith, SA and McLaughlin, Biochemistry 36: 6046-58 (1997) and Deringer et al., Biochemistry 39: 55-63 (2000), hereby expressly incorporated by reference).

[0142] In yet another preferred embodiment, inhibitory oligonucleotides are used. The invention makes use of complementary short oligonucleotides that create a blunt end on the probe oligonucleotides and prevent generation of false signals that are generated by enzymatic self-extension of probes. The signal of interest is only created in the presence of target and all other times the signal remains in the off mode (see FIG. 33). Short complementary APO E321.T.A oligo to APO E321.T.A SNP probe can inhibit APO E321.T.A SNP probe self-extension.

[0143] In another preferred embodiment, a combination of technologies is used wherein the combined result produces a marked reduction in false positive results due to self-extension. Here, three different technologies, PEP (primer extension preamplification), IVT (in vitro transcription) and probe extension with a modified reverse transcriptase are used that allow genome-wide SNP genotyping without multiple PCRs, without RCA (or other signal amplification technologies), and without problems from primer extension. The method involves performing a PEP (primer extension preamplification) reaction (known in U.S. Pat. No. 6,183,958; Zhang et al., Proc. Natl. Acad. Sci. 89:5847-51 (1992); Casas and Kirkpatrick, Biotechniques, 20: 219-25 (1996)) with random primers to amplify the genomic DNA in one reaction, followed by an IVT reaction (if one of the primers had a polymerase promoter sequence). The product, cRNA, is hybridized to the probe which is extended using a modified reverse transcriptase (RT) in which the DNA-dependent DNA polymerase activity has been eliminated. This means that self-extensions will not be extended by the RT because these are DNA-dependent extensions and only probe bound to RNA targets will be extended because these rely on the RNA-dependent DNA polymerase activity. Hence, self-extension is reduced.

[0144] A preferred embodiment of the present invention comprises an “array” of capture probes, including the WT and the mutant probe to a queried base, as described above, that are attached or immobilized to a solid support. The spatial location of the label on the solid support indicates the array element that shares a complementary sequence with the target, that is, whether the WT or the mutant base occurs in the DNA target. If both the WT and the mutant probes give a signal on detection, it indicates the presence of a heterozygote in the DNA sample.

[0145] By “array” herein is meant a plurality of probes in an array format; the size of the array will depend on the composition and end use of the array. Arrays containing from about 2 to about 100,000, with from about 5 to about 100 being preferred, and from about 10 to about 100 being particularly preferred. In addition, each array also comprises a first chromosome control probe and a second chromosome control probe. In some embodiments, the compositions of the invention may not be in array format; that is, for some embodiments, compositions comprising a single capture ligand may be made as well. In addition, in some arrays, multiple substrates may be used, either of different or identical compositions. Thus for example, large arrays may comprise a plurality of smaller substrates.

[0146] Accordingly, as described above, the present invention provides arrays wherein the probes or primers includes “non self-extension probes or primers”. That is, the probes or primers of the array are modified or designed such that they do not self-extend during primer extension reactions and thus minimize false positive reactions in the SBE assay.
[0147] In preferred embodiments, the invention includes a method of contacting the array that includes the non-self extendable primers described above with a target sample comprising P50 SNPs.

[0148] The devices of the invention describe a substrate with at least one surface comprising an array, and in a preferred embodiment, an array of electrodes. By “electrode” herein is meant a composition, which, when connected to an electronic device, is able to sense a current or charge and convert it to a signal. Alternatively an electrode can be defined as a composition which can apply a potential to and/or pass electrons to or from species in the solution. Preferred electrodes are known in the art and include, but are not limited to, certain metals and their oxides, including gold; copper; silver; chromium; titanium; platinum; palladium; silicon; aluminum; metal oxide electrodes including platinum oxide, titanium oxide, tin oxide, indium tin oxide, palladium oxide, silicon oxide, aluminum oxide, molybdenum oxide (MoO3), ruthenium oxides, and zinc oxide and tungsten oxide (WO3), both of which are transparent; conductive plastics (such as polymers like polythiophenes, polycrylamide, polyanilines, polypyrroles, and metal impregnated polymers); and carbon (including glassy carbon electrodes, graphite and carbon paste). Preferred electrodes include gold, silicon, carbon and metal oxide electrodes, with gold being particularly preferred.

[0149] The electrodes described herein are depicted as a flat surface, which is only one of the possible configurations of the electrode. The conformation of the electrode will vary with the detection method used. For example, flat planar electrodes may be preferred for optical detection methods or when arrays of nucleic acids are made, thus requiring addressable locations for detection. That is, each electrode has an interconnection attached to the electrode at one end and to a device that can control the electrode, on the other end thereby making each electrode independently addressable.

[0150] Alternatively, the electrode may be in the form of a tube comprising polymers, as will be described, and nucleic acids bound to the inner surface. This allows a maximum of surface area containing the nucleic acids to be exposed to a small volume of sample.

[0151] In preferred embodiments where polymer layers are used on the detection surface of the electrode, the electrode comprises polymers that can help prevent electrical contact between the electrodes and the TEMs, or between the electrode and charged species within the solvent.

[0152] All of these techniques rely on the formation of assay complexes on the surface of an electrode as a result of hybridization of a target sequence (either the target sequence of the sample or a sequence generated in the assay) to a capture probe on the surface. The assay complex further comprises a detection label to aid detection of complex formation. In preferred embodiments, the detection label is either an electron transfer moiety (ETM), or a fluorescent moiety or any other label that is either directly or indirectly attached to the target. Labels and various detection systems are described below.

[0153] In addition, the present invention is directed to a novel invention that capitalizes on novel properties of surface-bound arrays, and uses “competitors” to reduce non-specific binding.

[0154] Nucleic acids arrays are well known in the art, and can be classified in a number of ways; both ordered arrays (e.g. the ability to resolve chemistries at discrete sites), and random arrays are included. Ordered arrays include, but are not limited to, those made using photolithography techniques (Affymetrix GeneChip™), spotting techniques (Synteni and others), printing techniques (Hewlett Packard and Rosetta), three dimensional “gel pad” arrays, etc.

[0155] By “substrate” or “solid support” or other grammatical equivalents herein is meant any material that can be modified to contain discrete individual sites appropriate for the attachment or association of nucleic acids. As will be appreciated by those in the art, the “substrates” outlined herein can be made from a wide variety of materials, including, but not limited to, silicon such as silicon wafers, silicon dioxide, silicon nitride, glass, fused silica, modified silicon, carbon, gallium arsenide, indium phosphide, aluminum, ceramics, polymide, quartz, plastics, resins and polymers including polymethylmethacrylate, acrylics, polyybutylene, polyurethanes, polyethylene terephthalate, polycarbonate, polystyrene and other styrene copolymers, polypropylene, polytetrafluoroethylene, leflun, nylon or nitrocellulose, etc., polysaccharides, metal surfaces such as superalloys, zircaloy, steel, gold, silver, copper, tungsten, molybdenum, tantalum, Kovar, Kevlar, Kapton, Mylar, brass, sapphire, etc. Preferred embodiments utilize glass, silicon and ceramic materials, depending on the reagents utilized. As will be appreciated by those in the art, the material comprising the substrate should be compatible with the reagents outlined herein.

[0156] As will be appreciated by those in the art, nucleic acid probes can be attached or immobi-lized to a solid support in a wide variety of ways. By “immobilized” and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can be covalent or non-covalent. By “non-covalent binding” and grammatical equivalents herein is meant one or more of either electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as, streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By “covalent binding” and grammatical equivalents herein is meant that the two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds. Covalent bonds can be formed directly between the probe and the solid support or can be formed by a cross linker or by inclusion of a specific reactive group on either the solid support or the probe or both molecules. Immobilization may also involve a combination of covalent and non-covalent interactions. In general, the probes are attached to the biochip in a wide variety of ways, as will be appreciated by those in the art. As described herein, the nucleic acids can either be synthesized first, with subsequent attachment to the biochip, or can be directly synthesized on the biochip.

[0157] In a preferred embodiment, as is more fully outlined below, the substrate comprises a number of different layers, including electrodes, insulating layers and polymer layers, including, but not limited to, conductive polymers, polyacrylamide, polypyrrole, SunModic™ gel Matrices, etc.
for attachment of probes to the support, as is generally described in WO 98/20162 and WO 99/57317, both of which are hereby expressly incorporated herein by reference in their entirety. In alternative embodiments, probes are attached to polymers that are in contact with the microelectrodes.

[0158] In a preferred embodiment, the polymer used for attaching the oligonucleotide probes of the array is the SurModics™ gel matrix. This is a polyacrylamide-based gel that is derivatized to enable covalent attachment of oligonucleotides onto the polymer. For example, in FIG. 37, an example of an acyl substitution reaction is described through which an activated oligonucleotide can be attached to the polymer. The structure of the SurModics™ gel matrix is shown in FIG. 38. The method and platform include a substrate, the SurModics™ matrix and a number of oligonucleotide probes. Most commonly the substrate is a glass slide. Covalently attached to the glass slide is a thin polyacrylamide matrix. The polyacrylamide is functionalized with oligonucleotide attachment groups and photo-crosslinking groups. Different oligonucleotides of roughly 20-30 nucleotides are synthesized using standard phosphoramidite chemistry with an AminoLink™ terminal nucleotide, which comprises a (CH₂)₃—NH— linker. These oligonucleotides are spotted at discrete locations on the matrix. The probes are covalently attached to the matrix via an interaction with the AminoLink™ amino group and the oligonucleotide attachment groups present within the matrix. Once formed, the photo-crosslinked acrylamide matrix, with covalently attached oligonucleotide probes, is used in assaying a sample.

[0159] In this embodiment, the probes of the invention are arrayed onto slides coated with a film of hydrogel containing activated (NHS) esters ("3D-Link slides; cat.# DN01-0025; Surmodics, Inc.; Eden Prairie, Minn.) employing a modified Badged II dispense robot (Packard, Meridian, Conn.; Motorola Life Sciences, Tempe, Ariz.).

[0160] By “insulating layer” herein is meant a layer of material that will not substantially transport electrons. Preferably, the insulating layer is a layer of insulative dielectric material, including, but not limited to, ceramics, plastics, printed circuit board materials, polymers, metal oxide or nitrides such as SiO₂, Si₃N₄, or Al₂O₃.

[0161] In a preferred embodiment, the polymer layer uses polymers including, but not limited to, polypropylene, polyethylene, polyvinylidene fluoride, polystyrene, polyacrylamide, agarose gel, polyethylene glycol, cellulose, gel, dendrimers, metallic nanoparticles, carbon nanotubes, their derivatives, their copolymers, and combinations thereof, to increase the amount of probe concentration at a particular site. In preferred embodiments, the material comprises a neutral pyrurode matrix. To increase the probe loading capacity, porous matrix such as polyacrylamide, agarose, or sol gels are preferred. In these embodiments, probe molecules are attached onto a supporting matrix on the surface of the electrodes using the functional chemistry mentioned below. In alternative embodiments, probes are attached to polymers that are in contact with microelectrodes.

[0162] Furthermore, substrates are also referred to as “biochips”. By “biochip” or equivalents herein is meant a substrate comprising an array of distinct biomolecules, particularly nucleic acids.

[0163] Preferred substrates also include printed circuit board (PCB) materials. Reference is made to U.S. Ser. No. 09/796,077; PCT US00/34145; PCT US01/02664; PCT/US00/33499; PCT/US00/33497; PCT/US99/23324; WO 01/34302; WO 98/20162; WO 98/112430; WO 00/16899; WO 99/57317; WO 99/67425; WO 01/35100; WO 00/62931; WO 01/06016; WO 01/07665; and PCT/US01/01150, all of which are expressly incorporated by reference in their entirety.

[0164] In a preferred embodiment, the substrate also includes “array locations”. Accordingly, the present invention provides compositions comprising substrates with a plurality of array locations. By “array locations” or “pads” or “sites” herein is meant a location on the substrate that comprises a covalently attached nucleic acid probe. Additionally, the present system finds particular utility in array formats, further described below, wherein there is a matrix of addressable detection electrodes (herein generally referred to as “pads”, “addresses” or “micro-locations”).

[0165] The electrodes, in some preferred embodiments of this invention, comprise self-assembled monolayers (SAMs). By “monolayer” or “self-assembled monolayer” or “SAM” herein is meant a relatively ordered assembly of molecules spontaneously chemisorbed on a surface, in which the molecules are oriented approximately parallel to each other and roughly perpendicular to the surface. A majority of the molecules includes a functional group that adheres to the surface, and a portion that interacts with neighboring molecules in the monolayer to form the relatively ordered array. A “mixed” monolayer comprises a heterogeneous monolayer, that is, where at least two different molecules make up the monolayer. SAMs can also comprise conductive oligomers, described below.

[0166] As outlined herein, the efficiency of target sequence binding (for example, oligonucleotide hybridization) may increase when the sequence is at a distance from the detection electrode. Similarly, non-specific binding of biomolecules, including the target sequences, to a detection electrode is generally reduced when a monolayer is present. Thus, a monolayer facilitates the maintenance of the sequence away from the electrode surface. In addition, a monolayer serves to keep charged species away from the surface of the electrode. Thus, this layer helps to prevent electrical contact between the electrodes and the ETMs, or between the electrode and charged species within the solvent. Such contact can result in a direct “short circuit” or an indirect short circuit via charged species which may be present in the sample. Accordingly, the monolayer is preferably tightly packed in a uniform layer on the electrode surface, such that a minimum of “holes” exist. The monolayer thus serves as a physical barrier to block solvent accessibility to the detection electrode.

[0167] By “conductive oligomer” herein is meant a substantially conducting oligomer, preferably linear, some embodiments of which are referred to in the literature as “molecular wires”. By “substantially conducting” herein is meant that the oligomer is capable of transferring electrons at 100 Hz. Generally, the conductive oligomer has substan-
tially overlapping π-orbitals, i.e., conjugated π-orbitals, as between the monomeric units of the conductive oligomer, although the conductive oligomer may also contain one or more sigma (σ) bonds. Additionally, a conductive oligomer may be defined functionally by its ability to inject or receive electrons into or from an associated ETM. Furthermore, the conductive oligomer is more conductive than the insulators as defined herein. Additionally, the conductive oligomers of the invention are to be distinguished from electroactive polymers, that themselves may donate or accept electrons. In other preferred embodiments, the monolayer comprises electroconduit-forming species. By “electroconduit-forming species” or “EFS” herein is meant a molecule that is capable of generating sufficient electroconduits in a monolayer, generally of insulators to allow detection of electrons or ETMs at the surface.

[0168] In a preferred embodiment, the conductive oligomers have a conductivity, $S$, of from between about $10^{-6}$ to about $10^7$ Ω$^{-1}$cm$^{-1}$, with from about $10^{-5}$ to about $10^9$ Ω$^{-1}$cm$^{-1}$ being preferred, with these $S$ values being calculated for molecules ranging from about 20 Å to about 200 Å. As described below, insulators have a conductivity $S$ of about $10^{-7}$ Ω$^{-1}$cm$^{-2}$ or lower, with less than about $10^{-8}$ Ω$^{-1}$cm$^{-1}$ being preferred. See generally Gardner et al., Sensors and Actuators A 51 (1995) 57-66, incorporated herein by reference.

[0169] Desired characteristics of a conductive oligomer include high conductivity, sufficient solubility in organic solvents and/or water for synthesis and use of the compositions of the invention, and preferably chemical resistance to reactions that occur i) during nucleic acid synthesis (such that nucleosides containing the conductive oligomers may be added to a nucleic acid synthesizer during the synthesis of the compositions of the invention), ii) during the attachment of the conductive oligomer to an electrode, or iii) during hybridization assays. In addition, conductive oligomers that will promote the formation of self-assembled monolayers are preferred. The oligomers of the invention comprise at least two monomeric subunits, as described herein. As is described more fully below, oligomers include homo- and hetero-oligomers, and include polyamides.

[0170] In general, EFS have one or more of the following qualities: they may be relatively rigid molecules, for example as compared to an alkyl chain; they may attach to the electrode surface with a geometry different from the other monolayer forming species (for example, alkyl chains attached to gold surfaces with thiol groups are thought to attach at roughly 45° angles, and phenyl-acetylene chains attached to gold via thiols are thought to go down at 90° angles; they may have a structure that sterically interferes or interrupts the formation of a tightly packed monolayer, for example through the inclusion of branching groups such as alkyl groups, or the inclusion of highly flexible species, such as polyethylene glycol units; or they may be capable of being activated to form electroconduits, for example, photoactivatable species that can be selectively removed from the surface upon photolysis, leaving electroconduits.

[0171] Preferred EFS include conductive oligomers, as defined below, and phenyl-acetylene-polyethylene glycol species, as well as asymmetrical SAM-forming disulide species such as depicted the figures of U.S. Ser. No. 60/145, 912 filed Jul. 27, 1999, hereby expressly incorporated by reference. However, in some embodiments, the EFS is not a conductive oligomer.

[0172] As will be appreciated by those in the art, nucleic acid probes can be attached or immobilized to a solid support in a wide variety of ways. By “immobilized” and grammatical equivalents herein is meant the association or binding between the nucleic acid probe and the solid support is sufficient to be stable under the conditions of binding, washing, analysis, and removal as outlined below. The binding can be covalent or non-covalent. By “non-covalent binding” and grammatical equivalents herein is meant one or more of either electrostatic, hydrophilic, and hydrophobic interactions. Included in non-covalent binding is the covalent attachment of a molecule, such as streptavidin to the support and the non-covalent binding of the biotinylated probe to the streptavidin. By “covalent binding” and grammatical equivalents herein is meant that two moieties, the solid support and the probe, are attached by at least one bond, including sigma bonds, pi bonds and coordination bonds.

[0173] The method of attachment of the capture probe to the detection surface can be done in a variety of ways, depending on the composition of the “capture binding ligand” or “capture probe” and the composition of the detection surface. Both direct attachment or indirect attachment can be used. Indirect attachment is done using an attachment linker. In general, both ways utilize functional groups on the capture probe, the attachment linker or spacer, and the detection surface for covalent attachment. Preferred functional groups for attachment are amino groups, carboxy groups, oxo groups and thiol groups. These functional groups can then be attached, either directly or indirectly through the use of a linker, sometimes depicted herein as “Z.” “Linkers” or “spacers” or “anchoring groups” are well known in the art; for example, homo- or hetero-bifunctional linkers as are well known (see 1994 Pierce Chemical Company catalog, technical section on cross-linkers, pages 155-200, incorporated herein by reference). Preferred modifications useful in the practice of the invention include, but are not limited to, —OH, —NH$_2$, —SH, —COOR (where R=H, lower (C$_{1,12}$) alkyl, aryl, heterocyclic alkyl or aryl, or a metal ion), —CN, or —CHO. Immobilization of such derivatized probes is accomplished by direct attaching of the probe molecules on the detection surface through a functional group such —OH, —SH, —NH$_2$. In a preferred embodiment, probes are covalently attached to a SurModics™ matrix which is a polycrylamide based gel that is derivatized to enable covalent attachment of oligonucleotides onto the polymer. For example, in FIG. 37, an example of an acyl substitution reaction is described through which an activated oligonucleotide can be attached to the polymer. The structure of the SurModics™ gel matrix is shown in FIG. 38.

[0174] The present system finds particular utility in array formats, wherein there is a matrix of addressable detection electrodes (herein generally referred to “pads”, “addresses” or “micro-locations”).

[0175] Some array configurations are described herein. In a preferred embodiment CodeLink™ array technology is used, CodeLink™ technology provides an apparatus for performing high-capacity biological reactions on a biochip
comprising a substrate having an array of biological binding sites. It provides a hybridization chamber having one or more arrays, preferably comprising arrays consisting of hydrophilic, 3-dimensional gel and more preferably comprising arrays consisting of 3-dimensional polyacrylamide gels, wherein nucleic acid hybridization is performed by reacting a biological sample containing a target molecule of interest with a complementary oligonucleotide probe immobilized on the gel. Nucleic acid hybridization assays are advantageously performed using probe array technology, which utilizes binding of target single-stranded DNA onto immobilized oligonucleotide probes. Preferred arrays include those outlined in U.S. Ser. Nos. 09/458,501, 09/459, 685, 09/464,490, 09/605,766, PCT/US00/34145, Ser. No. 09/492,013, PCT/US01/02664, WO 01/54814, Ser. Nos. 09/458,533, 09/344,217, PCT/US99/27783, Ser. No. 09/439, 889, PCT/US00/42053 and WO 01/34292 all of which are hereby incorporated by reference in their entirety.

[0176] In another preferred embodiment, eSensor™ array technology is used. eSensor™ technology uses self-assembled monolayers (SAMs) on surfaces for binding and detection of biological molecules. SAMs are alky chains that protect an electrode from solution electronically active agents (e.g. salts). Electrochemical labels (e.g. ferrocene), which are initially bound to the label probe, flow to the electrode and back producing a detectable signal. See for example WO98/20162; PCT US98/12430; PCT US98/12082; PCT US99/01705; PCT/US99/21683; PCT/US99/10104; PCT/US99/01703; PCT/US00/31233; U.S. Pat. Nos. 5,620,850; 6,197,515; 6,013,459; 6,013,170; and 6,005,573; and references cited therein. In other preferred embodiments, electronic array technology is used, as is further described below.

[0177] The present invention also has devices that allow for simultaneous multiple biochip analysis. In particular, the devices are configured to hold multiple cartridges comprising nucleic acid arrays, and allow for high throughput analysis of samples.

[0178] By “cartridge” herein is meant a casing or housing for the biochip. As outlined herein, and as will be appreciated by those in the art, the cartridge can take on a number of configurations and can be made of a variety of materials. Suitable materials include, but are not limited to, fiberglass, teflon, ceramics, glass, silicon, mica, plastic (including acrylics, polystyrene and copolymers of styrene and other materials, polypropylene, polyethylene, polybutylene, polycarbonate, polyurethanes, Teflon™, and derivatives thereof, etc.), etc. Particularly preferred cartridge materials are plastic (including polycarbonate and polypropylene) and glass.

[0179] As will be appreciated by those in the art, the cartridge can comprise a number of components, including reaction chambers, inlet and outlet ports, heating elements including thermoelectric components, RF antennae, electromagnetic components, memory chips, scaling components such as gaskets, electronic components including interconnects, multiplexers, processors, etc.

[0180] The devices comprise a number of cartridge stations that are configured to receive the biochips, either same or different types of biochips, allowing analysis of different types of biochips. The stations can include a wide variety of different components, including but not limited to, thermocontrollers, signaling systems, sensors for leak detection, alphanumeric displays, and detectors. When preferred embodiments include the use of biochips comprising electrodes that rely on electrochemical detection, the devices and/or stations can comprise devices boards and processors used in the detection process. The biochip cartridges comprising the arrays of biomolecules, and can be configured in a variety of ways. For example, the chips can include reaction chambers with inlet and outlet ports for the introduction and removal of reagents. In addition, the cartridges can include caps or lids that have microfluidic components, such that the sample can be introduced, reagents added, reactions done, and then the sample removed for detection, as described in detail in U.S. Ser. No. 09/904,175 filed Jul. 11, 2001, hereby incorporated by reference in its entirety.

[0181] In a preferred embodiment, the cartridge comprises a reaction chamber. Generally, the reaction chamber comprises a space or volume that contains the contacting sample to the biochip array. The volume of the reaction chamber can vary depending on the size of the array and the assay being done. In general, reaction chamber ranges from 1 nl to about 1 ml, with from about 1 to about 250 μl being preferred and from about 10 to about 100 μl being especially preferred. In some embodiments, to avoid the introduction of air bubbles into the reaction chamber (which can be disruptive to detection), the reaction chamber is less than the size of the sample to be introduced, to allow a slight overflow and thus ensure that the reaction chamber contains little or no air.

[0182] The reaction chamber of the cartridge comprises an inlet port for the introduction of the sample to be analyzed. The inlet port may optionally comprise a seal to prevent or reduce the evaporation of the sample or reagents from the reaction chamber. In a preferred embodiment the seal comprises a gasket, through which a pipette or syringe can be pushed. The gasket can be rubber or silicone or other suitable materials, such as materials containing cellulose.

[0183] The reaction chamber can be configured in a variety of ways. In a preferred embodiment, the reaction chamber is configured to minimize the introduction or retention of air bubbles or other sample impurities. Thus, in a preferred embodiment, the reaction chamber further comprises an outlet port to allow air or excess sample to exit the reaction chamber. Thus the fluid sample flows up into the reaction chamber and contacts the array. In some embodiments, the outlet port vents to either a waste storage well, to an external surface of the chip or cartridge, or, in a preferred embodiment, back into the inlet port. Thus for example a preferred embodiment utilizes a system wherein the exit port vents to the inlet port, preferably above the point of loading. For example, when a pipette is used to load the cartridge, the tip of the pipette extends below the exit port, such that air from the exit port is not introduced into the reaction chamber. In addition, the materials of the cartridge housing and biochip can be chosen to be similar to hydrophobicity or hydrophlicity, to avoid the creation of air bubbles.

[0184] In addition, in a preferred embodiment, the reaction chamber/inlet and/or outlet ports optionally include the use of valves. For example, a semi-permeable membrane or filter may be used, that preferentially allows the escape of gas but retains the sample fluid in the chamber. For example, porous teflons such as Gortex™ allow air but not fluids to penetrate.
[0185] As will be appreciated by those in the art, there are a variety of reaction chamber geometries which can be used in this way. Generally having the intersection of the inlet port and the reaction chamber be at the “bottom” of the cartridge, with a small aperture, with the reaction chamber widening, is preferred. In addition, the “top” of the reaction chamber may narrow, as well. Thus, preferred embodiments for the size and shape of the reaction chamber allow for smooth loading of the reaction chamber. Preferred embodiments utilize reaction chamber geometries that avoid the use of sharp corners or other components that serve as points for bubble formation.

[0186] In addition, in some embodiments, the reaction chamber can be configured to allow mixing of the sample. For example, when a sample and a reagent are introduced simultaneously or separately into the chamber, the inlet port and/or the reaction chamber can comprise weirs, channels or other components to maximize the mixing of the sample and reagent. In addition, as is outlined below, the reaction may utilize magnetic beads for mixing and/or separation.

[0187] In a preferred embodiment, the cartridge comprises a sealing mechanism to prevent leakage of the sample or reagents onto other parts of the substrate, particularly (in the case of electronic detection) onto electronic interconnects. As will be appreciated by those in the art, this may take on a variety of different forms. In one embodiment, there is a gasket between the biochip substrate comprising the array and the cartridge, comprising sheets, tubes or strips. Alternatively, there may be a rubber or silicone strip or tube used; for example, the housing may comprise an indentation or channel into which the gasket fits, and then the housing, gasket and chip are clamped together. Furthermore, adhesives can be used to attach the gasket to the cartridge, for example, a double sided adhesive can be used; for example, silicone, acryle and combination adhesives can be used to attach the gasket to the biochip, which is then clamped into the cartridge as described herein.

[0188] In some embodiments, the reaction chamber and biochip substrate are configured such that a separate sealing mechanism is not required. For example, the biochip substrate can serve as one “half” of the reaction chamber, with the array on the inside, and the reaction chamber housing can serve as the other “half”. Depending on the materials used, there may be an optional adhesive to attach the two. Alternatively, when there are arrays on both sides of the substrate, the housing may encompass the substrate.

[0189] Thus, in these embodiments, the volume of the reaction chamber can be set either by forming a well in the cartridge, such that the addition of the biochip substrate forms a reaction chamber around the array, or by using a flat cartridge and using a gasket of a defined depth, or by combinations of the two.

[0190] In a preferred embodiment, the cartridge comprises a cap or lid. The cap may be functional, as outlined below when it comprises microfluidic components. In addition, the cap may be designed for safety purposes, to prevent the leakage of biological materials or cross-contamination. Additionally, the cap can be designed to be removable. As will be appreciated by those in the art, the cap can take on a wide variety of configurations. For example, in one embodiment, the cap merely seals the inlet port to prevent evaporation of the sample during the assay. In a preferred embodiment, the cap may comprise a number of additional elements for use in sample handling and reagent storage, to allow for a variety of different sample reactions. For example, a variety of microfluidic components can be built into the cap to effect a number of manipulations on a sample to ultimately result in target analyte detection or quantitation. See generally PCT US00/10903, and references outlined therein, all of which are expressly incorporated by reference. These manipulations can include cell handling (cell concentration, cell lysis, cell removal, cell separation, etc.), separation of the desired target analyte from other sample components, chemical or enzymatic reactions on the target analyte, detection of the target analyte, etc. The devices of the invention can include one or more wells for sample manipulation, waste or reagents; microchannels (sometimes referred to as flow channels) and/or between these wells, including microchannels containing electrophoretic separation matrices; valves to control fluid movement; on-chip pumps such as electrostatic, electrohydrodynamic, or electrokinetic pumps. In addition, as outlined herein, portions of the internal surfaces of the device may be coated with a variety of coatings as needed, to reduce non-specific binding, to allow the attachment of binding ligands, for biocompatibility, for flow resistance, etc. These microfluidic caps can be made in a variety of ways, as will be appreciated by those in the art. See for example references described in PCT US00/10903, and references outlined therein, all of which are expressly incorporated by reference.

[0191] When the cap of the cartridge is used as part of the assay, it may be configured to include one or more of a variety of components, herein referred to as “modules”, that will be present on any given device depending on its use, and are connected as required by microchannels. These modules include, but are not limited to: sample introduction or collection modules; cell handling modules (for example, for cell lysis, cell removal, cell concentration, cell separation or capture, cell growth, etc.); separation modules, for example, for electrophoresis, dielectrophoresis, gel filtration, ion exchange-affinity chromatography (capture and release) etc.; reaction modules for chemical or biological alteration of the sample, including amplification of the target analyte (for example, when the target analyte is nucleic acid, amplification techniques are useful, including but not limited to polymerase chain reaction (PCR), oligonucleotide ligation assay (OLA); strand displacement amplification (SDA), and nucleic acid sequence based amplification (NASBA) and other techniques outlined in WO 99/37819 and PCT US00/19889), chemical, physical or enzymatic cleavage or alteration of the target analyte, or chemical modification of the target; fluid pumps (including, but not limited to, electrostatic, electrohydrodynamic, or electrokinetic pumps; fluid valves; thermal modules for heating and cooling; storage modules for assay reagents; mixing chambers; and detection modules.

[0192] In addition, while these microfluidic components are described herein as being associated with the cap of the cartridge, as will be appreciated by those in the art, these modules and channels (as well as other components outlined herein) may be located anywhere in the cartridge or device. In addition, some components may be in the device; for example, “off chip” pumps may be located within one or more stations of the device.
[0193] The cartridge comprises at least one biochip, with some embodiments utilizing one or more biochips per cartridge.

[0194] Detection of a label indicates the presence of the target sequence. All the techniques described below rely on the formation of assay complexes on a surface, as a result of hybridization of a target sequence comprising the sequence complementary to the capture probe or the SBE probe sequence.

[0195] There are three general ways in which the assays of the invention are run. That is, the hybridization can be either direct or indirect (sandwich type). In a first embodiment, the “target sequence” or “target analyte” is labeled; binding of the target sequence thus provides the label at the surface of the solid support. Alternatively, in a second embodiment, unlabeled target sequences are used, and a sandwich” format is utilized; in this embodiment, there are at least two binding ligands used per target sequence molecule; a “capture” or “anchor” binding ligand (also referred to herein as a “capture probe”, particularly in reference to a nucleic acid binding ligand) that is attached to the detection surface as described herein, and a soluble binding ligand (frequently referred to herein as a “signaling probe”, “label probe”, or “electron transfer moiety”), that binds independently to the target sequence, and either directly or indirectly comprises at least one label. In a third embodiment, as further outlined below, none of the compounds comprises a label, and the system relies on changes in electronic properties for detection.

[0196] In another preferred embodiment, the detection technique comprises a “sandwich” assay, as is generally described in U.S. Ser. No. 60/073,011 and in U.S. Pat. Nos. 5,681,702, 5,597,909, 5,545,730, 5,594,117, 5,591,584, 5,591,670, 5,590,731, 5,575,670, 5,591,584, 5,624,802, 5,635,352, 5,594,118, 5,359,100, 5,124,246 and 5,681,697, all of which are hereby incorporated by reference. Although sandwich assays do not result in the alteration of primers, sandwich assays can be considered signal amplification techniques since multiple signals (i.e. label probes) are bound to a single target, resulting in the amplification of the signal. Sandwich assays are used when the target sequence does not comprise a label; that is, when a secondary probe, comprising labels, is used to generate the signal. As discussed herein, it should be noted that the sandwich assays can be used for the detection of primary target sequences (e.g. from a patient sample), or as a method to detect the product of an amplification reaction as outlined above.

[0197] A variety of detection methods may be used, including, but not limited to, optical detection (as a result of spectral changes upon changes in redox states), which includes fluorescence, phosphorescence, luminiscence, chemiluminescence, electrochemiluminescence, and refractive index; and electronic detection, including, but not limited to, amperometry, voltammetry, capacitance and impedance; and electrochemical detection that include, but are not limited to, transition metal complexes, organic ETMs, and electrodes. Detection of electron transfer is generally initiated electronically, with voltage being preferred. A potential is applied to the assay complex. Precise control and variations in the applied potential can be via a potentialstat and either a three electrode system (one reference, one sample (or working) and one counter electrode) or a two electrode system (one sample and one counter electrode). This allows matching of applied potential to peak potential of the system which depends in part on the choice of ETMs (when reporters are used) and in part on the other system components, the composition and integrity of the monolayer, and what type of reference electrode is used.

[0198] In some embodiments, the detection module is configured to allow for optical detection of target sequences. Here, the detection surface may comprise any surface suitable for the attachment of capture probes described above. Generally, optical detection of target sequences involve providing a colored or luminescent dye as a ‘label’ on the target sequence. Preferred labels include, but are not limited to, fluorescent lanthanide complexes, including those of Europium and Terbium, fluorescein, rhodamine, tetramethylrhodamine, eosin, erythrosin, coumarin, methyl-coumarins, pyrene, Malachite green, stilbene, Lucifer Yellow, Cascade Blue™, Texas Red, 1,1’-diisopropyl-3,3’-dimethyl-imino-3,1-propanediylbis[3-[3-methyl-2(3H)-benzoxazolylidene)methyl]-, tetraiodide, which is sold under the name YOYO-1, and others described in the 6th Edition of the Molecular Probes Handbook by Richard P. Haugland, hereby expressly incorporated by reference.

[0199] In a preferred embodiment, the streptavidin-Alexa system is used for detection. Here, a sample is initially processed by amplification with biotinylated primers, resulting in biotinylated amplicons. The array matrix is then contacted with these biotinylated amplicons. Hybridization of the amplicons to the oligonucleotide probes is measured by addition of Streptavidin that has been conjugated with a fluorescent dye, an Alexa dye purchased from Molecular Probes, Inc. Due to the biotin-streptavidin interaction, the dye is localized to those discrete locations of the array matrix where an oligonucleotide probe is hybridized to a biotinylated amplicon. All other locations are free of biotinylated amplicons, and thus free of the dye. The presence of dye may be detected in any of a variety of ways, and detection may be automated to allow for automated data processing.

[0200] After binding, a variety of techniques allow for the detection of radiation emitted by the fluorescent labels. These techniques include using fiber optic sensors with nucleic acid probes in solution or attached to the fiber optic. Fluorescence is monitored using a photomultiplier tube or other light detection instrument attached to the fiber optic.

[0201] In addition, scanning fluorescence detectors such as the FluorImager sold by Molecular Dynamics are ideally suited to monitoring the fluorescence of modified nucleic acid molecules arrayed on solid surfaces. The advantage of this system is the large number of electron transfer probes that can be scanned at once using chips covered with thousands of distinct nucleic acid probes.

[0202] Further, photodiodes, CCD cameras, or an active pixel system may be used to image the radiation emitted by fluorescent labels.

[0203] These methods also include time or frequency dependent methods based on AC or DC currents, pulsed methods, lock-in techniques, filtering (high pass, low pass, band pass), and time-resolved techniques including time-resolved fluorescence.

[0204] For electrochemical detection, the target sequence can comprise an electrochemically active reporter (also
referred to herein as an electron transfer moiety (ETM), such as a transition metal complex, defined below. ETMs can be attached to either nucleic acids, target sequences, or soluble binding ligands as is generally outlined in WO "08/20162, hereby expressly incorporated by reference in its entirety.

[0205] Once the assay complexes are formed, the presence or absence of the ETMs are detected as is described below and in U.S. Pat. Nos. 5,591,578; 5,824,473; 5,770,309; 5,705,348 and 5,780,234; U.S. Ser. Nos. 08/911,589; 09/135,183; 09/306,653; 09/134,058; 09/295,691; 09/238, 351; 09/245,105 and 09/338,726; and PCT applications WO09/20162; WO 00/16898; PCT US99/01705; PCT US99/01703; PCT US00/10093 and PCT US99/10104, all of which are expressly incorporated herein by reference in their entirety.

[0206] The terms "electron donor moiety," "electron acceptor moiety," and "ETMs" (or grammatical equivalents herein refers to molecules capable of electron transfer under certain conditions. It is to be understood that electron donor and acceptor capabilities are relative, that is, a molecule which can lose an electron under certain experimental conditions, will be able to accept an electron under different experimental conditions. It is to be understood that the number of possible electron donor moieties and electron acceptor moieties is very large, and that, one skilled in the art of electron transfer compounds, will be able to utilize a number of compounds in the present invention. Preferred ETMs include, but are not limited to, transition metal complexes, organic ETMs, and electrodes.

[0207] In a preferred embodiment, the ETMs are transition metal complexes. Transition metals are those whose atoms have a partial or complete d shell of electrons. Suitable transition metals for use in the present invention include, but are not limited to, cement (Cu), copper (Cu), cobalt (Co), palladium (Pd), zinc (Zn), iron (Fe), ruthenium (Ru), rhodium (Rh), osmium (Os), rhenium (Re), platinum (Pt), scandium (Sc), titanium (Ti), Vanadium (V), chromium (Cr), manganese (Mn), nickel (Ni), Molybdenum (Mo), technetium (Tc), tungsten (W), and iridium (Ir). That is, the first series of transition metals, the platinum metals (Ru, Rh, Pd, Os, Ir and Pt), along with Fe, Re, W, Mo and Tc, are preferred. Particularly preferred are ruthenium, rhenium, osmium, platinum, cobalt and iron.

[0208] As will be appreciated in the art, the co-ligands can be the same or different. Suitable ligands fall into two categories: ligands which use nitrogen, oxygen, sulfur, carbon or phosphorus atoms (depending on the metal ion) as the coordination atoms (generally referred to in the literature as sigma (σ) donors) and organometallic ligands such as metal-cyclopentadiene ligands (generally referred to in the literature as pi (π) donors, and depicted herein as Lm). Suitable nitrogen donating ligands are well known in the art and include, but are not limited to, NH₂⁻, NR⁻R; NHR; pyridine; pyrazine; imidazol; bipyridine and substituted derivatives of bipyridine; terpyridine and substituted derivatives; phenanthrolines, particularly 1,10-phenanthroline (abbreviated phen) and substituted derivatives of phenanthrolines such as 4,7-dimethylphenanthroline and dipyrrolid [3,2-a:2',3'-c:phenazine abbreviated dpp]; dipyrrolophenazine; 1,4,5,8,9,12-hexaaazatriphenylene (abbreviated ha); 9,10-phenanthrenequinione diimine (abbreviated phi); 1,4,5,8-tetraazaphenanthrene (abbreviated nap); 1,4,8,11-tetraazaacyclotetradecane (abbreviated cyclam), EDTA, EGTA and isocyanide. Substituted derivatives, including fused derivatives, may also be used. In some embodiments, porphyrins and substituted derivatives of the porphyrin family may be used. See, for example, Comprehensive Coordination Chemistry, Ed. Wilkinson et al., Pergamon Press, 1987, Chapters 13.2 (pp73-98), 21.1 (pp. 813-898) and 21.3 (pp 915-957), all of which are hereby expressly incorporated by reference.

[0209] In addition to transition metal complexes, other organic electron donors and acceptors may be covalently attached to the nucleic acid for use in the invention. These organic molecules include, but are not limited to, riboflavin, xanthene dyes, azo dyes, acridine orange, N,N,N-dimethyl-2,7-diazapyrenium dichloride (DAP⁺), methylviologen, ethidium bromide, quinones such as N,N-dimethylanthracene, 1,9-def 6,5,10-dejdiisooquinoline dichloride (ADIQ⁺), porphyrins [meso-tetrakis[N,N-methyl-p-xyridinium]porphyrin tetrachloride], varalime blue B hydrochloride, Bind- scheller’s green; 2,6-dichlorophenol, 2,6-dibromophenolindophenol; Brilliant creast blue (3-amino-9-dimethylamino-10-methylphenoxazone chloride), methylene blue; Nile blue A (aminoaphthodimethylaminophenoxazone sulfate), indigo-5,5',7,7'-tetrasulfonic acid, indigo-5,5',7-trisulfonic acid; phenosafranine; indigo-5-monosulfonic acid; safranine T; bis(dimethylglyoximato)-iron(II) chloride; indoline scarlet, neutral red, anthracene, corone, pyrene, 9-phenylanthracene, rubrene, binaphthyl, DPA, phenothiazine, fluoranthene, phanthenone, chrysene, 1,8-diphenyl-1,3,5,7-octatetraene, naphthalene, acenaphthene, perylene, TMPO and analogs and substituted derivatives of these compounds.

[0210] The choice of the specific ETMs will be influenced by the type of electron transfer detection used, as is generally outlined below. Preferred ETMs are metalloccenes, with ferrocene being particularly preferred.

[0211] In a preferred embodiment, a plurality of ETMs are used.


[0213] Detection of electron transfer is generally initiated electronically, with voltage being preferred. A potential is applied to the assay complex. Precise control and variations in the applied potential can be measured via a potentiostat and either a three electrode system (one reference, one sample (or working) and one counter electrode) or a two electrode system (one sample and one counter electrode). This allows matching of applied potential to peak potential of the system which depends in part on the choice of ETMs (when reporters are used) and in part on the other system components, the composition and integrity of the monolayer, and what type of reference electrode is used.


[0215] In a preferred embodiment, electronic detection is used, including amperometry, voltammetry, capacitance, and impedance. Suitable techniques include, but are not limited to, electrograviometry; coulometry (including controlled potential coulometry and constant current coulometry); voltammetry (cyclic voltammetry, pulse voltammetry (normal pulse voltammetry, square wave voltammetry, differential pulse voltammetry, Osteryoung square wave voltammetry, and countostatic pulse techniques); anodic stripping analysis, cathodic stripping analysis, square wave stripping voltammetry); conductance measurements (electrotytic conductance, direct analysis); time-dependent electrochemical analyses (chronoamperometry, chronopotentiometry, cyclic chronopotentiometry and amperometry, AC polargraphy, chronogalvametry, and chronocoulometry); AC impedance measurement, capacitance measurement; AC voltammetry, and photoelectrochemistry.

[0216] In a preferred embodiment, monitoring electron transfer is via amperometric detection. This method of detection involves applying a potential (as compared to a separate reference electrode) between the nucleic acid-conjugated electrode and a reference (quartz) electrode in the sample containing target genes of interest. Electron transfer of differing efficiencies is induced in samples in the presence or absence of target nucleic acid; that is, the presence or absence of the target nucleic acid, and thus the label probe, can result in different currents.

[0217] The device for measuring electron transfer amperometrically involves sensitive current detection and includes a means of controlling the voltage potential, usually a potentiostat. This voltage is optimized with reference to the potential of the electron donating complex on the label
probe. Possible electron donating complexes include those previously mentioned with complexes of iron, osmium, platinum, cobalt, rhenum and ruthenium being preferred and complexes of iron being most preferred.

[0218] In a preferred embodiment, alternative electron detection modes are utilized. For example, potentiometric (or voltammetric) measurements involve non-faradic (no net current flow) processes and are utilized traditionally in pH and other ion detectors. Similar sensors are used to monitor electron transfer between the ETM and the electrode. In addition, other properties of insulators (such as resistance) and of conductors (such as conductivity, impedance and capacitance) could be used to monitor electron transfer between ETM and the electrode. Finally, any system that generates a current (such as electron transfer) also generates a small magnetic field, which may be monitored in some embodiments.

[0219] In a preferred embodiment, electron transfer is initiated using alternating current (AC) methods.

[0220] Without being bound by theory, it appears that ETMs, bound to an electrode, generally respond similarly to an AC voltage across a circuit containing resistors and capacitors.

[0221] Alternatively, reporterless or labelless systems are used. In this embodiment, two detection electrodes are used to measure changes in capacitance or impedance as a result of target sequence binding. See generally U.S. Ser. No. 09/458,533, filed Dec. 9, 1999 and CPT US00/33497, both of which are expressly incorporated by reference.

[0222] In this embodiment, using a labelless system, the surface of the two detection electrodes is covered with a layer of polymer matrix.

[0223] When labels such as ETMs are not used, other initiation/detection systems may be preferred. In this embodiment, molecular interactions between immobilized probe molecules and target molecules in a sample mixture are detected by detecting an electrical signal using AC impedance. In other embodiments, such molecular interactions are detected by detecting an electrical signal using an electrical or electrochemical detection method selected from the group consisting of impedance spectroscopy, cyclic voltammetry, AC voltammetry, pulse voltammetry, square wave voltammetry, AC voltammetry, hydrodynamic modulation voltammetry, conductance, potential step method, potentiometric measurements, amperometric measurements, current step method, other steady-state or transient measurement methods, and combinations thereof.

[0224] In one embodiment of the apparatus of the present invention, the means for producing electrical impedance at each test electrode is accomplished using a Model 1260 Impedance/Gain Phase Analyzer with Model 1287 Electrochemical Interface (Solartron Inc., Houston, Tex.). Other electrical impedance measurement means include, but are not limited to, transient methods using AC signal perturbation superimposed upon a DC potential applied to an electrochemical cell such as AC bridge and AC voltammetry. The measurements can be conducted at any particular frequency that specifically produces electrical signal changes that are readily detected or otherwise determined to be advantageous. Such particular frequencies are advantageously determined by scanning frequencies to ascertain the frequency producing, for example, the largest difference in electrical signal. The means for detecting changes in impedance at each test site electrode as a result of molecular interactions between probe and target molecules can be accomplished by using any of the above-described instruments.

[0225] The following examples serve to more fully describe the manner of using the above-described invention, as well as to set forth the best modes contemplated for carrying out various aspects of the invention. It is understood that these examples in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All references cited herein are incorporated by reference.

EXAMPLES

Example 1

Human P450 SNP Assay

[0226] A list of mutations (alleles) that are of possible clinical interest in the P450 genes have been identified. The worksheet in FIG. 3A through D describes the polymorphisms (SNPs) and literature references for each of the genes that have been incorporated into the initial assay design. The P450 assay is designed to discriminate the different polymorphic states (alleles) of the selected SNPs present in a given individual DNA sample. The assay protocol can be broken down into two major process areas; target preparation and signal detection. Signal detection is accomplished through the technique known as single base extension (SBE). Target preparation is based on the specific amplification of human genome regions if interest using polymerase chain reaction (PCR) technology.

[0227] Amplification of DNA through PCR technology is highly dependent on the design of the short primer sequences used for extension and amplification. For the P450 assay, the difficulty in design of these primer sequences lies in the nature of the closely related P450 gene family. All of the P450 genes belong to a superfamily of 50-60 genes and 25+ pseudogenes and may share significant levels of (regional) homology, particularly those in the same subfamily (i.e. subfamily 2C contains genes 2C8, 2C9, 2C18, 2C19, etc.). Gene members in the same subfamily are located on the same chromosome and usually contain an identical number of introns and exons. The attached table illustrates the relative homology shared by the genes in the 2D6 family (see FIG. 4).

[0228] The approach to designing primers specific to only the target gene of interest has been to select regions with base pair mismatches against subfamily related genes focusing particularly at the 3' end of the primer. In this manner, both the annealing hybridization specificity and on the discrimination of the PCR polymerase which extends nucleotides from the 3' end of the primer are relied upon. Primer design was further restricted to a length greater than 19 bp and a balance in terms of GC content and Tm per pair. Whenever possible, primers have also been selected to end with an -AA-3' or -CA-3' at the 3' end to prevent primer-dimer formation. All primer candidates have been analyzed using the BLAST algorithm (homology analysis algorithm) against a compiled P450 sequence library (57 genes) to
remove any potential cross-reactive primer candidates. In addition to homology analysis, the primer candidates were also screened against a database of repetitive sequences.

FIG. 5 characterizes all the P450 primers currently being used for amplification.

[0229] Primers were mainly designed in the exonic regions of the gene sequences due to the principle that these areas are relatively more conserved. In many cases, the exon regions that contain discriminating mismatches were far apart resulting in amplicons varying in length from 300-6500 bp. This variance in amplicon length dictates the amplification format for long PCR conditions. Optimization of PCR conditions to reduce any non-specific products and maintain yield was focused on the following areas: annealing temperature, extension time, Mg2+ concentration, primer concentration, and enzyme concentration. The Human P450 CodeLink protocol reflects the current optimal PCR conditions (see example 2 below).

[0230] The two genes, CYP2D6 and CYP2C19, are of the greatest clinical interest and also have the highest homology to its relative subfamily genes. To assess the specificity of the CYP2D6 and CYP2C19 primers, primers specific to 2D6, 2D7A, 2D7B and 2C9 were designed. Experiments were run with all the primers using genomic DNA target (to assess that the primers are working) and followed by reamplification using all the primers with CYP2D6 and CYP2C19 diluted amplicon as the target. It is assumed that if any cross-species amplification was occurring in the PCR of CYP2D6 and CYP2C19, these cross-products would be seen upon reamplification of CYP2D6/2C19 diluted ampli-

[0231] The P450 amplicons amplified from a single DNA sample are prepared for SNP detection by combining all amplicons, purifying the pooled solution (removal of excess dNTPs and primers) and fragmenting of the PCR product. The entire protocol and process is outlined in example 2. Once the target preparation is complete, the next part of the process is signal detection via a single base extension (SBE).

[0232] The detection of each single nucleotide polymorphism is accomplished via SBE of the 3’ end of our probes bound to a polymer surface on the 5’ end. The probes are designed in pairs with the 5’ end at the SNP position or SNP+1. The perfect match to the target present will be kinetically favored during hybridization and give the most signal intensity. In the case of a heterozygote mutation, both probes in a pair will demonstrate relatively equal intensity. The quality of detection is therefore reliant on the specificity and quality of the probe design. All initial probe designs were processed using probe designing software packages, e.g., “Probe Design”. The identical fixed reference sequences with the SNPs identified using IUPAC nomenclature used for primer design was used for probe design.

[0233] For each SNP, probe candidates were designed from both sequence directions, sense and anti-sense. In addition to direction, probes were designed at 2 different Tm’s, 60° and 70° C. For cases where the probe sequences overlap multiple SNPs in close proximity, dITP (deoxyinosine tri-phosphate) is used in places where a polymor-

[0234] To increase confidence in the assay analysis, 2 types of control probes were also designed. Probes to detect the successful amplification of each of the 10 unique ampli-

[0235] All of the probe candidates were passed through iterations of initial performance screening through actual chip builds and assay tests. The probe set that discriminate the mutations the best were selected (sequencing data used as the gold standard). After initial performance screening, many of the probes which demonstrated signal above background due to possible self-extension were redesigned with addition of 1-2 base overhang. FIG. 6 depicts the final probe list for the current P450 assay. Performance data on the discrimination capabilities of these probes is summarized in FIG. 7.

Example 2

Human P450 CodeLink Protocol: A Preferred Protocol for Running the Chips of the Invention

(See FIGS. II to 16).

[0236] Product Description

[0237] The CodeLink™ SNP Bioarrays: Human P450 is a tool for genotyping 96 single-nucleotide polymorphisms (SNPs) related to Cytochrome P450 genes. The CodeLink SNP Bioarrays: Human P450 contains single-stranded oligonucleotide probes for a subset of seven human Cyto-

[0238] Four independent assays can be performed on each CodeLink SNP Bioarrays: Human P450.

[0239] Performance Specifications

[0240] This product will have a call rate of 95% with accuracy of 98%.

[0241] Storage and Handling

[0242] Upon receipt, the CodeLink SNP Bioarrays: Human P450 should be stored at room temperature in original packaging.

[0243] The Motorola SBE Kit and Uniplex PCR primer plates should be stored at ~20° C.

[0244] Product Use Limitations

[0245] The CodeLink SNP Bioarrays: Human P450 is for research use only and is not to be used for diagnostic purposes.

[0246] All biological specimens and materials should be handled as if capable of transmitting infection and disposed
of with proper precautions in accordance with federal, state, and local regulations. These include adherence to the OSHA Bloodborne Pathogens.

[0247] Standard (29 CFR 1910.1030) for blood-derived and other samples governed by this act.

[0248] Precautions

[0249] Exercise care to avoid cross-contamination of samples, reagents and arrays during all steps of this procedure.

[0250] Avoid microbial contamination. Avoid areas of high particulate content.


[0252] 1.0 Getting Started

[0253] 1.1 Motorola Reagents/Kits

[0254] CodeLink SNP Bioarrays: P450 Beta Kit, Motorola, Part# 20002

[0255] Data Disk CD, Motorola, Part# 20006

[0256] P450 Beta PCR primer plates, Motorola, Part# 30007

[0257] SBE Kit, Motorola, Part# 20006

[0258] Biotinylated Acyclo Terminator™ Nucleotide Mix

[0259] SBE Buffer

[0260] 1.2 Prescribed Reagents/Kits

[0261] Roche Biochemicals

[0262] PCR-Grade Deoxy-Nucleotide Set, Roche Biochemicals, Part# 1960964

[0263] Applied Biosystems

[0264] GeneAmp® XL DNA Polymerase (contains 3.3×XL Buffer, 25 mM Magnesium acetate (Mg(OAc)₂), Applied Biosystems, Part# N808-0187 [400 units] or Part# N808-0188 [2400 units]

[0265] Life Technologies, Inc. (LTI)

[0266] DNase I Amplification Grade (1U/µL) shipped with DNase I 10× Buffer, LTI, Part# 1808-015

[0267] Ultrapure ddH₂O (Nuclease free), Life Technologies, Inc. (LTI), Part# 15230-170

[0268] Aldrich

[0269] Ethyl alcohol, anhydrous 99.5+%, Aldrich, Part# 15, 199-4, or equivalent

[0270] Qiagen

[0271] QIAQuick™ 8 PCR Purification Kit, Qiagen, Part# 28142

[0272] Amersham

[0273] Thermo Sequenase™ (Enzyme only), Amersham, Part# E79000Y [1000 units] or Part# E79000Z [10,000 units]

[0274] Sigma

[0275] 1 M Tris-HCl pH 7.6 stock solution, Sigma, Part# T2788

[0276] 5 M sodium chloride solution, Sigma, Part# S5150

[0277] 20×SSC, Sigma, Part# S6639

[0278] Molecular Probes

[0279] Streptavidin, Alexa Flouros® 532 conjugate, Molecular Probes, Part# S-11224

[0280] Fisher

[0281] Tween 20 (polyoxyethylene sorbitan mono-taurate) enzyme grade, Fisher, Part# BP337-100

[0282] 1.3 Genomic DNA (gDNA) Sample Quality

[0283] The sample genomic DNA (gDNA) should be at a concentration of 0.450 mg/mL to 0.300 mg/mL. The recommended buffer is 10 mM Tris, 1 mM EDTA, pH 8.0. The OD 260/280 should be between 1.71 and 1.79 and protein concentration should be less than 0.9 mg/mL DNA.

[0284] 1.4 Supplies Required

[0285] Wire slide rack, VWR, Part# 25461-014, or equivalent

[0286] Microcentrifuge

[0287] 1.7 mL microcentrifuge tubes and racks

[0288] Solution basis, 55 mL, sterile, VWR, Part# 21007-972, or equivalent

[0289] Micropipettors (2 µL, 20 µL, 200 µL, 1000 µL)

[0290] Sterile aerosol resistant pipette tips

[0291] Wide orifice tips, Rainin, Part# HR250W

[0292] MicroAmp® Clear Adhesive Plate seals, Applied Biosystems, Part# 430611

[0293] 15 mL Falcon tubes, VWR, Part# 21008-918, or equivalent

[0294] 96-well collection microplates RB, Qiagen, Part# 19581

[0295] 0.2 mL 96-well plate, Marsh Bio Products, Part# T0296PE

[0296] Chamber removal fixture, Motorola, Part# 100017

[0297] Powder-free gloves

[0298] Slide storage box

[0299] 1.5 Licenses Required

[0300] License to practice PCR from Applied Biosystems or Hoffman-LaRoche and F. Hoffman- LaRoche Ltd.

[0301] 1.6 Instruments Required

[0302] Thermal cycler, Applied Biosystems GeneAmp® PCR System 9700, Part# N805-0001

[0303] 96-well plate clinical centrifuge, Qiagen, Part# 81010
[0304] QIAVac 6S, Qiagen, Part# 19503
[0305] QIAVac plate holder, Qiagen, Part# 19548
[0306] Filtered ddH₂O source, Barnstead or equivalent
[0307] Hybaid Omnislide Thermal Cycler, Motorola, Part# 600002
[0308] Hybaid Omnislide Wash Module, Motorola, Part# 600003
[0309] Hybaid Ambient Rack, Motorola, Part# 600074
[0310] GenePix™ 4000 Axon Scanner, Motorola, Part# 600000
[0311] Computer, Motorola, Part# 600001
[0312] CodeLink™ System Software:
[0313] Analysis, Motorola, Part# 20004
[0314] Scanning, Motorola, Part# 200003

[0315] 1.7 Reagent Preparation

[0316] The following reagents should be prepared before beginning the assay, with the exception of the Staining Solution.

[0317] Target Preparation Reagents

[0318] Water used in all target preparations should be Ultrapure ddH₂O (nuclease free) from LTi.
[0319] 25 mM dNTP solution
[0320] Using the PCR-grade deoxy-nucleotide set, make 20 mL of solution using the following:
[0321] 0.5 mL 100 mM dTTP
[0322] 0.5 mL 100 mM dATP
[0323] 0.5 mL 100 mM dGTP
[0324] 0.5 mL 100 mM dCTP
[0325] 18 mL of ultrapure ddH₂O
[0326] Divide into 5, 4 mL aliquots and store dNTP mix at -20°C in 15 mL tubes.
[0327] PCR Master Mix
[0328] To make 41 mL (sufficient for 1000 PCR reactions or 83 samples):
[0329] 20.0 mL ultrapure ddH₂O
[0330] 15.0 mL GeneAmp XL 3.3× Buffer II (Vortex before use)
[0331] 2.0 mL Mg(OAc)₂
[0332] 4.0 mL 2.5 mM dNTP solution
[0333] Divide into 5, 8 mL aliquots and store at -20°C. (there will be 1 mL excess).
[0334] DNase 1× Buffer
[0335] To make 1 mL:
[0336] 900 µL ultrapure ddH₂O
[0337] 100 µL DNase 10× Buffer
[0338] Store at -20°C.
[0339] Diluted Genomic DNA (gDNA)
[0340] To make 60 µL:
[0341] 58 µL ultrapure ddH₂O
[0342] 2 µL gDNA

[0343] Post Reaction Processing Reagents

[0344] To make 1 L Washing Solution:
[0345] Add 250 mL of 20XSSC to 750 mL of filtered ddH₂O. Mix well. Store at room temperature and preheat in Hybaid Omnislide Wash Module to 60°C before use.
[0346] To make 1 L Staining Diluent:
[0347] Combine 918 mL of filtered ddH₂O with 30 mL of 5 M NaCl and 50 mL of 1 M Tris-Cl pH 7.6. Add 2 mL of Tween 20 and mix well. Store at room temperature.
[0348] To make 2 mLs of Streptavidin Alexa Diluent:
[0349] Combine 100 µL of 1 M Tris-Cl pH 7.6 with 60 µL of 5 M NaCl and 1840 µL of ultrapure ddH₂O.
[0350] To make 1 mL Streptavidin Alexa-Fluor 532 conjugate:
[0351] Add one mL of Streptavidin Alexa dixuent to a vial containing crystalline Streptavidin Alexa-Fluor 532 conjugate. Mix well by inversion and store at 4°C, protected from light.
[0352] To make Staining Solution:
[0353] Prepare immediately prior to use.
[0354] Determine the volume of staining solution needed based on the amount of slides being processed (300 mL for 1-10 slides, 600 mL for 11-20 slides). For 50 mL of staining solution, mix 300 µL of Streptavidin Alexa-Fluor 532 conjugate solution with 300 mL of staining diluent. For 600 mL of staining solution, mix 600 µL of Streptavidin Alexa-Fluor 532 conjugate solution with 600 mL of TNT. Mix well by inversion.
[0355] To make 1 L Destaining Solution:
[0356] Combine 870 mL of filtered ddH₂O with 30 mL of 5 M NaCl and 100 mL of 1 M Tris-Cl pH 7.6. Mix well and store at room temperature.

[0357] 1.8 Instrument Preparation

[0358] Refer to the original manufacturer’s user manual for specific instructions.

[0359] Thermal Cycler GeneAmp 9700

[0360] Enter the following PCR program into the thermal cycler:

[0361] a. 94°C for 1:00 minute
[0362] b. 94°C for 0:30 minutes
[0363] c. 68°C for 4:00 minutes
[0364] d. Repeat b and c, 29 times for a total of 30 cycles
[0365] c. 72°C for 10:00 minutes
[0366] f. 4°C Forever

[0367] Program will end automatically after it completes the last step entered.

[0368] Save program as “PCR”.

[0369] Thermal Cycler GeneAmp 9700

[0370] Enter the following program for DNase target fragmentation into the thermal cycler:

[0371] a. 37°C for 10:00 minutes
[0372] b. 95°C for 10:00 minutes
[0373] c. 4°C Forever

[0374] Program will end automatically after it completes the last step entered.

[0375] Save program as “Fragment”.

[0376] Hybaid Omnislide Thermal Cycler

[0377] Choose the simulated slide program type and enter the following SBE reaction program into the thermal cycler:

[0378] a. 85°C for 0:00 minute
[0379] b. 85°C for 0:30 minutes
[0380] c. 60°C for 10:00 minutes
[0381] d. Go to step b for 7 more cycles for a total of 8 cycles.

[0382] Program will end automatically after it completes the last step entered.

[0383] Save program as “01” when prompted.

[0384] 2.0 Target Preparation

[0385] 2.1 PCR Reactions

[0386] For each sample, 45 µL of gDNA at approximately 10 ng/µL will be necessary.

[0387] Each primer plate will accommodate 8 samples at 12 PCR reactions per sample.

[0388] 1. Prepare the reaction mix for each sample in a new 1.7 mL microcentrifuge tube using the appropriate volumes (Table 1).

[0389] 2. Aliquot 45 µL of reaction mix into each well of the 12 well layout in the primer plate allocated for each sample (FIG. 12).


[0391] 4. Spin in centrifuge for 1 minute at 1000 rpm.

[0392] 5. Immediately place plate in GeneAmp PCR System 9700. Run program “PCR” entered in section 1.8.

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume for 1 PCR reaction</th>
<th>Volume for 12 PCR reactions w/Excess</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR Master Mix</td>
<td>41.0 µL</td>
<td>615.0 µL</td>
</tr>
<tr>
<td>GeneAmp XL enzyme</td>
<td>1.0 µL</td>
<td>15.0 µL</td>
</tr>
<tr>
<td>gDNA (10 ng/µL)</td>
<td>3.0 µL</td>
<td>45.0 µL</td>
</tr>
<tr>
<td>TOTAL</td>
<td>45.0 µL</td>
<td>675.0 µL</td>
</tr>
</tbody>
</table>

[0393] 2.2 PCR Pooling and Purification

[0394] Use the QIAquick 8 PCR Purification Kit to remove primers and nucleotides from the PCR reactions according to the following protocol:

[0395] 1. Remove plate from the thermal cycler within 12 hours of PCR program completion.

[0396] 2. Pool each sample’s PCR reactions (12) into separate new 15 mL conical tubes for each sample. (Total volume=600 µL in each tube).

[0397] 3. Add 3,000 µL of PB buffer provided in kit to each of the 15 mL tubes containing pooled target. Mix well by vortexing.

[0398] 4. Place a new QIAquick 8-well strip on the QIAvac 65 vacuum manifold without a collection plate.

[0399] 5. Pipet entire contents of one 15 mL conical tube (3600 µL at a time) into one of the wells of the QIAquick 8-well strip under house vacuum (150 mbar).

[0400] 6. Repeat for 7 other sample tubes.

[0401] 7. Add 1 mL of PE buffer provided in kit to each well containing sample.

[0402] 8. Repeat PE buffer addition for a total of 2 mLs per well.


[0404] 10. Remove the 8-well QIAquick trip and smack and dab on a stack of paper towels to remove excess PE buffer.

[0405] 11. Replace the 8-well QIAquick strip on the manifold and apply house vacuum for 5 minutes to dry columns. Remove cover.

[0406] 12. Place a new QIAquick collection plate on the plate holder and insert under the 8-well QIAquick strip. Replace cover.

[0407] 13. Add 40 µL of DNase 1× Buffer to the center of each QIAquick well and incubate at room temperature for 1 minute.

[0408] 14. Turn on house vacuum for 1 minute to collect purified PCR product in the collection plate.

[0409] NOTE: Do not use the provided EB buffer for elution.
[0410] 2.3 DNA Fragmentation

[0411] Digest the PCR products according to the following protocol:

[0412] 1. Transfer each well of purified PCR product from the QiAquick collection plate into a new 0.2 mL Marsh 96-well plate maintaining the integrity and identity of the individual samples.

[0413] 2. Make fresh 1:300 dilution of DNase I in DNase 1x Buffer.

[0414] 3. Add 10 μL of diluted DNase I to each well containing purified PCR product.

[0415] 4. Immediately, place the 0.2 mL PCR 96-well plate onto the thermal cycler and run the "Fragment" protocol programmed in section 1.8.

[0416] NOTE: DO NOT use EDTA to stop the fragmentation; the enzyme is heat-killed during thermal cycling.

[0417] The samples are now ready for immediate use in the SBE reaction, or store samples at -20°C in a sealed plate until ready for the SBE reaction.

[0418] 3.0 SBE Reaction

[0419] A maximum of 16 slides can be run on the Hybaid Omnislide thermal cycler at one time.

[0420] 3.1 Prepare SBE Reaction Master Mix

[0421] 1. Thaw all reaction components (except enzyme) at room temperature. This includes PCR products (if frozen above), 10x Nucleotide Mix, SBE Buffer, and ultrapure ddH₂O.

[0422] 2. When reagents are completely thawed, vortex each thoroughly and spin for 15 seconds in the microcentrifuge before preparing the reaction mix.

[0423] 3. Prepare the SBE Reaction Master Mix with the appropriate volumes (except enzyme) in a new 1.7 mL microcentrifuge tube (Table 2). Adjust volumes based on the total number of arrays being run. (A maximum of 26 arrays can be prepped in one 1.7 mL tube.)

[0424] 4. Remove the enzyme from the freezer and add the appropriate volume (Table 2). Mix well by pipetting.

[0425] 5. Immediately return all reagents to -20°C.

[0426] 6. Gently vortex the tube after the addition of all reagents. Spin in the microcentrifuge for 15 seconds to consolidate sample and to remove any bubbles or foam.

TABLE 2

<table>
<thead>
<tr>
<th>Reagent</th>
<th>1 Array</th>
<th>4 Arrays W/Excess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motorola SBE Buffer</td>
<td>8.0 μL</td>
<td>35.2 μL</td>
</tr>
<tr>
<td>10x Nucleotide Mix</td>
<td>4.0 μL</td>
<td>17.6 μL</td>
</tr>
</tbody>
</table>

[0427] 3.2 Prepare Individual SBE Reactions

[0428] 1. After mixing and spinning, transfer 13 μL of the SBE reaction master mix to each clean, labeled microfuge tube (FIG. 13).

[0429] 2. Add 27 μL of the purified PCR target to each tube.

[0430] 3.3 Load Reaction Mixtures into Flex Chambers

[0431] The time between loading of flex chambers and placement on the thermal cycler should not exceed 45 minutes. The slides should not be placed on ice at any time.

[0432] 1. For each flex chamber, aspirate 40 μL of reaction mixture in a wide orifice tip (making sure there are no bubbles in the tip).

[0433] 2. Place the pipette tip at a 90° angle over the appropriate array port and apply downward pressure until the tip forms a seal with the adhesive on the port. (FIG. 14.) This will minimize leaking outside of the port and reduces the risk of blocking the introduction channel.

[0434] WITHOUT using the blowout feature of the pipettor, slowly eject the sample out of the tip and into the array chamber. You will see the fluid moving across the array area. Use smooth, even pressure to maintain a uniform rate during filling.

[0435] 4. After fluid has completely filled the chamber and a slight excess has appeared in the opposite port, maintain pressure and remove the pipette tip from the input port.

[0436] 5. After filling a flex chamber, check for the presence of air bubbles. If there are any large (>1.5 mm) bubbles present or if any bubbles appear in the center of the array, the array cannot be used. Small bubbles near the ports should not affect performance.

[0437] 6. Repeat this loading procedure using fresh tips with the other array chambers and slides.

[0438] NOTE: Do not attempt to withdraw fluid from flex chamber in the event of misloading.

[0439] 3.4 Seal Flex Chambers

[0440] 1. After all flex chambers on each slide have been loaded, they are ready for sealing.

[0441] 2. Select a pre-cut sealing strip and remove the backing using forceps to expose the adhesive. Handling carefully, apply the strip to the edge of the slide near one end. Slowly place the strip so that it covers all four ports on one side of the slide (FIG. 15). NOTE: Adhesive will become firmly attached at first contact with the chamber. Use caution to ensure
that the sealing strip is in the proper orientation before placing it on the chamber. Do not attempt to remove a misaligned sealing strip. Instead, place a second strip to cover any open ports.

[0442] 3. Repeat the application of a second strip to cover the remaining 4 ports and for each slide until all ports are sealed.

[0443] 4. When all strips have been applied, ensure that each port is sealed securely by pressing firmly on the strips. Be careful not to apply pressure over the array itself as this may cause volume loss.

[0444] 3.5 Thermal Cycling

[0445] 1. Add 50 mL of filtered ddH₂O to the plastic chamber surrounding the heat blocks of the Hybid Omnislide thermal cycler.

[0446] 2. Place each loaded and sealed slide (flex chamber acing up) onto the heat blocks. Do not use the rack normally used for this step. Lay the slides directly on the heat blocks, 8 slides per block. (FIG. 16.)

[0447] 3. Place the cover over the heat blocks and secure by turning the lock tabs.

[0448] 4. Press the “Menu” button on the Omnislide control panel. Select “Run” from the menu and then press “Enter”.

[0449] 5. Enter program “01” (programmed in section 1.8). Once you have selected the “01” program, press the “Enter” button.

[0450] NOTE: The appropriate calibration factor of 100 will be displaced after the “Enter” button is pressed.

[0451] 4.0 Post-Reaction Slide Processing

[0452] For the post-reaction slide processing steps, the volume of each solution used is dependent on the number of slides being processed. For 1-10 slides, use 300 mL of the designated solution. For 11-20 slides, use 600 mL of the designated solution.

[0453] NOTE: Between every step, the Hybid wash station sleeves should be rinsed with filtered ddH₂O three times.

[0454] 4.1 Prepare Hybid Wash Station

[0455] NOTE: Two sleeves and one rack for the Hybid washer will be necessary for slide processing.

[0456] 1. Fill Sleeve 1 with the appropriate volume of preheated Washing Solution. Place the sleeve in the station and set the temperature to 60° C.

[0457] 2. Take Sleeve 2 and a rack to the sink with filtered ddH₂O source and fill the sleeve with the appropriate volume.

[0458] 4.2 Remove Flex Chambers

[0459] 1. Remove the slides from the thermal cycler within 15 minutes of program completion.

[0460] NOTE: DO NOT ALLOW THE SLIDE TO DRY DURING THIS PROCESS. If drying is noticed, make a notation of the occurrence as it may lead to an increase in background noise.

[0461] 2. Take the slides to the sink where your rack and sleeve are located. Turn on the filtered ddH₂O source. (A 500 mL squirt bottle filled with filtered ddH₂O can also be used.)

[0462] 3. Wearing powder-free gloves, place a slide in the chamber removal fixture. Grasp the removal tab portion of the flex chamber and, while holding the slide securely, slowly pull back the flex chamber tab until the first array area is exposed.

[0463] 4. Immediately rinse the exposed array area with filtered ddH₂O to rinse away the SBE reaction solution by holding the chamber removal fixture in the filtered ddH₂O stream.

[0464] 5. After the first array has been rinsed, continue to pull back the flex chamber tab until the second array is exposed and then immediately rinse it with filtered ddH₂O.

[0465] 6. Repeat this process until four arrays are sequentially exposed and rinsed as quickly as possible. This minimizes the risk of cross contamination between arrays.

[0466] 7. After the flex chamber has been removed, rinse the entire slide with filtered ddH₂O, then quickly place it into a slot in the Hybid rack and immediately submerge the rack into the sleeve of filtered ddH₂O (Sleeve 2).

[0467] 8. Repeat this process for each slide until all the slides have had their chambers removed and been placed in the rack.

[0468] NOTE: The first slides will be submerged in filtered ddH₂O while subsequent slides have their chambers removed.

[0469] 4.3 Wash and Rinse Slides

[0470] 1. Place the rack containing slides from the sleeve with filtered ddH₂O (Sleeve 2) into the preheated sleeve (Sleeve 1) containing Washing Solution in the wash station.

[0471] 2. Allow the slides to incubate for 30 minutes at 60° C. During this time, open the wash sleeve and gently agitate the Washing Solution by moving the rack up and down 5 times within the sleeve every 10 minutes.

[0472] 3. Rinse and fill Sleeve 2 with the appropriate volume of filtered ddH₂O. When the incubation is complete, transfer the rack from the Washing Solution (Sleeve 1) to the filtered ddH₂O (Sleeve 2).

[0473] CAUTION: The rack and Washing Solution will be HOT, handle with care.

[0474] 4. Using both sleeves and rinsing between uses, rinse the rack of slides in filtered ddH₂O two additional times with agitation. Leave the rack in the last filtered ddH₂O rinse in the sleeve.

[0475] While the slides are incubating in the Washing Solution, prepare the appropriate amount of Staining Solution using the dilution instructions in section 1.7.
4.4 Stain Slides

- Fill the available sleeve (Sleeve 1) with the appropriate amount of Staining Solution.
- Transfer the rack from the last filtered ddH₂O wash to the sleeve with the Staining Solution.
- Place sleeve 1 (containing rack and Staining Solution) in the NON-heated slot on the wash station.
- Allow the slides to incubate in the Staining Solution for 30 minutes at room temperature. During this time, open the wash sleeve and gently agitate the Staining Solution by moving the rack up and down 5 times within the sleeve every 10 minutes.
- While the slides are incubating in the Staining Solution, prepare the appropriate amount of Destaining Solution using the dilution instructions in section 1.7.

4.5 Destain Slides

- Fill the available sleeve (Sleeve 2) with the appropriate amount of Destaining Solution.
- When the slides have completed the incubation described in section 4.4, remove the sleeve and take it to the sink.
- Transfer the rack from the Staining Solution to the sleeve with the Destaining Solution.
- Place the sleeve (containing rack and Destaining Solution) into the NON-heated slot on the wash station.
- Incubate the slides at room temperature for five minutes and repeat one additional time with fresh Destaining Solution (Sleeve 1).

4.6 Rinse Slides

- After the last incubation in the Destaining Solution, remove Sleeve 1 and take it to the sink.
- Fill the available sleeve (Sleeve 2) with the appropriate amount of filtered ddH₂O. Transfer the rack from the Destaining Solution (Sleeve 1) to the filtered ddH₂O sleeve (Sleeve 2).
- Fill a 1 L graduated cylinder with the appropriate amount of filtered ddH₂O.
- Holding the rack, decant the filtered ddH₂O. Immediately refill the sleeve with water from the cylinder making sure that all slides are completely covered.
- Repeat the decant and refill process three additional times using filtered ddH₂O. Leaving the rack in the last filtered ddH₂O rinse.

4.7 Dry Slides

- Carefully remove a rinsed slide from its slot in the wash rack. Replace the wash rack in the filtered ddH₂O filled sleeve.

NOTE: Slides should be handled by their label ends or edges only. The wet polymer substrate is fragile and easily damaged.

2. Allow the slides to air-dry. If available, a dry stream of clean nitrogen (no oil) can be blown over the surface to facilitate drying. To avoid damage to the wet polymer, keep the nitrogen source at least 6 inches away from the slide surface. When the slide is dry, place it in the slide storage box.

3. Repeat drying procedure for each slide.

4. When all slides are dry, inspect them individually. There should be no sign of salt crystals or water spots. If there are salt crystals or water spots, rinse the individual slide with filtered ddH₂O and dry.

5. Store the slides in a slide storage box.


Example 3

Probes With Multiple Base Additions

Target-independent self extension through stem-loop forming probes that result in false-positives have been discussed. The intensity of the signal is roughly proportional to the stability of the stem-loop structure. A survey of about 500 probes indicated that about 15% show a self-extension signal. Several examples of probe sequences that self-extend are shown in FIG. 18. In most cases the base incorporated is the one predicted by the template sequence immediately downstream of the stem-loop structure; this result strongly suggests that the DNA polymerase is extending the duplex formed by the stem-loop structure. Generally, a stem of three or more base pairs is required for self-extension.

The present invention describes a simple strategy in which one or more bases are added to the 3' end of the probe that self-extends (see FIG. 17). The probes used herein are novel in that the additional bases are not self-complementary within the probe but are complementary to the target. When such an additional base(s) is/are added, it creates a mismatch along the stem-loop structure and hence, the DNA polymerase does not extend the self annealed probe and hence no false positive detection occurs resulting in correct detection of the SNP base.

Target dependent extension is not affected by the additional base(s) added to the probe because the additional bases are complementary to the target. In practice adding four or more bases is not feasible. The SNP base will be too far upstream of the 3' end of the probe and a mismatch will create an internal bulge. If the short duplex downstream of the internal bulge is extended by the polymerase, the mismatch will not be detected.

All experiments were done with the DNA microarrays described in examples 1 and 2. The probes are poly-nucleotides (generally from 10-40 nt in length) attached at the 5' end to a support material on the slide. The SBE reaction consists of buffer, hapten-labeled acyclo terminator
nucleotides and DNA polymerase. Target (PCR amplified genomic DNA containing the SNP of interest) were either added or not, to the reaction.

[0507] FIG. 19 shows the SBE assay results for a probe that shows strong self-extension in the absence of target. When a single base is added to the 3' end of the probe, the target-independent signal is significantly reduced. The target-dependent signal was not altered indicating that the probe functions properly in the presence or absence of target. The addition of two or more bases created new stem-loop structures; these probes show strong target-independent signal as expected. FIG. 20 shows the overall performance of a set of rescued probes for the P450 gene. Of the 35 probe sets that made incorrect calls, 26 (60%) sets repaired by adding one or two bases to the probe called correctly 100% of the time.

Example 4

Probes With Modified Nucleotides

[0508] Single Base Extension (SBE) assay and self extending probe problems have been explained (FIG. 2). Here, we present a set of modified bases, when incorporated into a probe, inhibits it from folding and/or forming a stable secondary structures (FIGS. 21 and 24). The modified base does not form a stable base-pair with itself but individual bases hybridize well with natural DNA/RNA bases to form non-natural base-pairs that are very stable. Such a base-pair (FIGS. 22, 23, 25 and 26) thermodynamically destabilizes a self-folded structure but has minimal effect on the stability of a probe-target duplex. Thus, they selectively inhibit the formation and/or extension of self-folded molecules.


[0510] Self-folded structure may also be destabilized by using only one of the two bases in the base-pair, which would decrease the stability of both the self-folded molecule as well as the probe-target duplex. In that case, the melting temperature of the self-folded structure can be lowered substantially more than that of the probe-target duplex. Some example base-pairs/bases that can have such an effect include but are not limited to a) 2-amino-A:2-thio-T, b) 2-aminopurine:2-thio-T, c) 6-thio-G, d) 2-thio-C, e) hydrophobic bases such as 4-methylindole, difluorotoluene, etc. (Moran, S, Ren R X, Sheils C J, Runnym S 4th, Kool E T, Nucleic Acids Res (1966) 1:24(11):2044-52), 4-thio-T (FIG. 22).

[0511] Some solutions to the primer-dimer issues can also be applied to the self-folding issue. For example, a set of oligonucleotides carrying analogs of natural bases, which have a modification at the exo-cyclic amine positions (FIG. 23), are also inhibited from self-base-pairing (U.S. Pat. No. 6,001,611, PCT WO 00/07779). This kind of base can also be used in the same fashion as described in FIG. 21.

[0512] Another way of suppressing the signal from self-folded structures is by extending only the probe-target duplexes. This can be achieved by using certain modifications in the probe molecules that are not replicated by DNA polymerases (FIG. 24). That way, only the natural, target strands are used as templates, and thus extended, by the polymerases. There are a few known examples of such modifications known in literature (see above and U.S. Pat. No. 6,001,611; Wang M D, Chen J L, Weiss A., Nucleic Acids Res. (1999) 1:27(23), 4642-4. Again, such modifications have been previously used only for the primer-dimer issues and not the kind of applications proposed herein. For example, a hydrophobic base, 4-methylindole, when present in a DNA template, terminates the DNA polymerization at that site. This base has been called a “terminator”. Incorporation of 2'-O-methyl RNA nucleosides has also been shown to inhibit DNA polymerization at the residue sites. It is proposed that these types of “terminator” bases and nucleosides can also be used to prevent extension of self-folded molecules, thus selectively enhancing the signal from probe-target duplexes (FIG. 25).

[0513] Experiments were carried out on the DNA microarrays described above using our standard SBE Assay, to test these hypotheses. Probes containing either no modification or with modified bases described above were obtained from Operon and attached to gel-slabs using amino-NHS ester chemistry (Gen3 chemistry). Sequences of the probes used are shown on top of each graph (FIGS. 27-29). All data is from slides run in Ti chambers, standard SBE CYCLE parameters, 4 TAMRA nucleotides, Tris-HCl pH8.5 plus 10 mM KCl (buffer), no DMSO. When present, target was used at 10 ng per array at 60 ul volume. Data is summed from 8 arrays for “without target” assays and 7 arrays for “with target” assays.

[0514] Results are shown in the following graphs (FIGS. 27-29) and clearly show that incorporation of such modifications increases the signal in the presence of target vs in the absence of target many fold, proving that the hypotheses worked as intended. Thus, a signal from a self-extending probe in the presence of target is only threefold higher than in the absence of target (R=3.18). But, when an A:T base-pair in the step-loop is replaced with its modified analog base-pair X:Z, this ratio is increased to almost 10-fold (FIG. 27). This clearly shows that the modified base-pair is having the intended effect. Placement of the modified base-pair closer to the extension site (3'-end) can have an even more dramatic impact. Similarly, when adenosine immediately adjacent to the stem-loop is replaced with its “terminator-type" analog 4-methylindole, the signal increases from 3-fold to almost 5-fold (FIG. 27).

[0515] In FIG. 28 a signal with a self-extending probe in the presence of target is only three-fold higher than in the absence of target (R=2.42). But, when either one or both of A:T base-pairs in the step-loop are replaced with their modified analog base-pairs X:Z, this ratio is increased to almost 5-fold. This clearly shows that the modified base-pair is having the intended effect. FIG. 29 shows similar results with three different probe sequences.
Example 5

Probes With a Modified Stem That Inhibits Extension Enzyme Binding for Hairpin Structures

[0516] Non-specific incorporation of nucleotides due to hair-pin loop structures or palindromic sequence can often lead to false positive results in the single base extension assay. Here, by altering the affinity of the enzyme to bind to the “stem” region of the capture probes, non-specific nucleotides incorporation in the absence of target DNA can be reduced or eliminated (FIGS. 30 and 31). Modifications include but are not limited to, phosphorothioate, phosphoromimetic, chiral phosphodiester analogues and methyl phosphate on the “stem” region of the capture probe. The phosphate in the nucleic acid plays a crucial role in the nucleic acid-protein interaction. The electrostatic interactions between the positively charged amino acid and the negatively charged phosphate backbone, and formation of hydrogen bonds between the phosphate oxygen and protein contribute to the binding affinity and specificity of the enzyme (FIG. 32B (a)).

[0517] By modifying the phosphate or the sugar ring at the “stem” region, we can decrease the binding affinity of the enzyme reducing the probability of non-specific nucleotides incorporation. Since modifications are not performed at the base, thermodynamic duplex stability between the capture probe and the target has not been altered (FIGS. 30 and 31).

[0518] Several modifications are proposed but are not limited to those. Phosphorothioate (sulfur substitution) and phosphoroamidate (nitrogen substitution) can alter the charge distribution, hydrophobicity and the ability for the enzyme to form an efficient hydrogen bond to the phosphate backbone. Methyl phosphonate and methyl phosphate eliminate the phosphate charge altogether, which can inhibit the enzyme to bind to the DNA (see FIGS. 32A and 32B (b to d)).

[0519] Other modifications on the sugar ring can also be introduced such as the 2’ O-methyl RNA and LNA.

Example 6

Use of Inhibitory Oligonucleotides to Prevent Self-Extension

[0520] The invention makes use of complementary short oligonucleotides that create a blunt end on the probe oligonucleotides and prevent generation of false signals that are generated by enzymatic self-extension of probes. The signal of interest is only created in the presence of target and all other times the signal remains in the off mode (see FIG. 33). Short complementary APO E321.TA oligo to APO E321.TA SNP probe can inhibit APO E321.TA SNP probe self-extension. APO E321.TA probe (5’AAGCACTGAGG-C GGA 3’) is a strong self-extension producing strong self-extension false signal under all conditions. The assay used for these studies was the modified SBE assay by using genome-wide single strand RNA as targets for SBE reaction.

Example 7

Combination of Technologies and Use of Modified Reverse Transcriptase to Prevent Self-Extension

[0521] The invention is based upon a novel method (comprising three different technologies) which allows genome-wide SNP genotyping without multiple PCRs, without RCA (or other signal amplification technologies), and without problems from primer extension. The method involves performing a PEP (primer extension preamplification) reaction (known in U.S. Pat. No. 6,183,958; Zhang et al., Proc. Natl. Acad. Sci. 89:5847-51 (1992); Casas and Kirkpatrick, Bio-techniques, 20: 219-25 (1996)) with random primers to amplify the genomic DNA in one reaction, followed by an in vitro transcription (IVT) reaction (if one of the primers had a polymerase promoter sequence). The product, cRNA, is hybridized and the probe is extended using a reverse transcriptase (RT). This would serve the same purpose as PCR of individual SNP loci in the sense that we would be reducing the complexity of the genome through the PEP and amplifying the target with the IVT. Thus, one target prep is performed for all of the SNPs loci (as we do in the expression assay). Finally, the use of a modified RT (in which the DNA-dependent DNA polymerase activity has been eliminated) will mean that self-extensions will not be extended by the RT because these are DNA-dependent extensions and only probe bound to RNA targets will be extended because these rely on the RNA-dependent DNA polymerase activity (see FIG. 34).

[0522] This disclosure presents a novel methodology for SNP genotyping which does not require multiplexed PCR or thousands of PCR reactions. Furthermore, self-extension problems would not be evident because these self-extensions would not be extended by an RT which only has RNA-dependent DNA polymerase activity.

[0523] The technical feasibility of the approach has been demonstrated in three separate pieces. The feasibility of a primer extension preamplification (PEP) has been demonstrated in many molecular biology labs. The feasibility of an IVT has been demonstrated. The feasibility of an RT extending oligo probe on arrays after hybridization to RNA targets has been demonstrated (Pastinen, T., et al., (2000) Genome Research, 10:1031-42).

[0524] This method is differentiated from other methods in that all three technologies for uniplexed target prep and primer extension are combined which should not be complicated by self-extension problems that other primer extension technologies are subject to.

[0525] The novelty of the approach is based on the fact that we are using a modified RT to direct only RNA-dependent DNA polymerase activity and the fact that we are combining PEP and IVT for a new method of uniplexed target prep and applying this on oligonucleotide arrays for primer extension.

[0526] Strategic benefits are cost reduction in PCR, competitive advantages of a uniplexed target prep, discrimination, and better data reliability due to the elimination of self-extension.

[0527] Methodology

[0528] Step 1

[0529] Primer Extension using human genomic DNA and primer RRN017, which was designed to have 8 degenerate nucleotides at the 3’ end and a T7 RNA polymerase sequence at the 5’ end. It had the sequence GGGCCAGTMGT- TAAACAGACTCATAAGGGAGCCC- GNNNNNNNCCAGA. Primer extension with this primer should result in fragments of human DNA with T7 RNA polymerase sequences on the 5’ ends. The assay was performed with 2 ug of human genomic DNA, and 5 U of
AmpliTaq DNA polymerase, and 100 ng of RRNOT7 primer using PCR Amplification Buffer I (Perkin Elmer) in a final volume of 60 µl. The reaction was run for 50 cycles at 92°C, 1', 37°C, 2', with a slow ramp of 10 sec/degree to 55°C for 4'.

[0530] Step 2

[0531] In vitro Transcription using these standard conditions: A total of 325 µg of cRNA was obtained from 2 µg of starting human genomic DNA. The cRNA was quantitated using UV absorbition, and visualized on an agarose gel.

[0532] Step 3

[0533] The following oligonucleotides were diluted to 18 nM, and manually spotted on blank Surmodics slides. The first is a non-self-extender, whereas the second is a self-extender, since it loops back on itself.

GTTCTTAATTCGATGTTGGAAATTG
GTTCTTAATTCGATGTTGGAAATTG

[0534] Slides were blocked and processed according to the standard Surmodics protocol. Additionally, a bacterial oligonucleotide, YJ8EK, was also spotted and process on blank Surmodici slides, to serve as a negative control.

[0535] Step 4

[0536] 25 µg of cRNA were hybridized per oligonucleotide spot, in 50% formamide, 2xSSPE, at 37°C in a humidity chamber overnight (18 hrs). Slides were washed and dried.

[0537] Step 5

[0538] Each hybridized probe was extended in situ using rTh RNA-dependent DNA polymerase (Perkin Elmer), in the presence of Mn2+ and cy5-UTP, in a final volume of 20 µl, at 30°C, for 5'. Slides were washed, dried and scanned on an Axon scanner, using the Cy5 channel (635 nm) at 600 PMT.

[0539] Results

[0540] After performing this assay in the presence of rTh, a modified reverse transcriptase which only extends DNA off an RNA template in the presence of Mn2+ ions, self-extension is not observed off an SNP oligonucleotide which normally self-extends (see FIGS. 35 and 36).

REFERENCES

[0541] The following references are hereby incorporated by reference in their entirety.


[0547] 6. PCT WO 00/06779.

[0548] 7. U.S. Pat. No. 6,001,611.


---continued---

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 1

ccacagaagcc ttgcaaggtctca 23

<210> SEQ ID NO 2
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 2

cctacccaggg aaagcaagagacct 25

<210> SEQ ID NO 3
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 3
tgaagcctct gcagaggccc aa 22

<210> SEQ ID NO 4
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 4
gagagctgct ctgtgtactc cagca 25

<210> SEQ ID NO 5
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 5
caccttcctac cagcctgtggt tgas 24

<210> SEQ ID NO 6
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 6
cacctgtgctctcctgggttgaa 23

<210> SEQ ID NO 7
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 7
cctgacctct gcagccttttc tattaa 27

<210> SEQ ID NO 8
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 8

aggagcttggg gcgtgctctca aa 22
<210> SEQ ID NO 9
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 9
ctgtaggtgt ggtggtgtgg gatgaa

<210> SEQ ID NO 10
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 10
caccacctcc actggccccc ttgtga

<210> SEQ ID NO 11
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 11
cctctggttaa aggaagctat gggtcag

<210> SEQ ID NO 12
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 12
catgcaagct caatgcaagct taatgcagaa

<210> SEQ ID NO 13
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 13
cctctccccc aacggcccct c

<210> SEQ ID NO 14
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 14
gcagagaga taaggcgttc catca

<210> SEQ ID NO 15
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 15
atcctgatat gcagactcca gtgca

<210> SEQ ID NO 16
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 16
caagcansag aggtaconca tcacotgga

<210> SEQ ID NO 17
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 17
cagctotca gattotgta tcctcna

<210> SEQ ID NO 18
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 18
gggtcagcag atgcogagcc aa

<210> SEQ ID NO 19
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 19
cctgctattt tggcttggas gcctcag

<210> SEQ ID NO 20
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 20
tccaggggt tqttgtatct cattc

<210> SEQ ID NO 21
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 21
gcaaqcgaa ggtatcggt ggaag

<210> SEQ ID NO 22
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 22
gcaaqcgga ggtatcggt ggaac

<210> SEQ ID NO 23
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 23
tocacagcgg caat
<210> SEQ ID NO: 24
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 24
tccacgccgg caac 14

<210> SEQ ID NO: 25
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 25
ataaggtct tccacgccg 20

<210> SEQ ID NO: 26
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 26
ataaggtct tccacgccg a 21

<210> SEQ ID NO: 27
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 27
cacttacct gcgtgctggtg 20

<210> SEQ ID NO: 28
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 28
acttacagt gcgtgctggg 19

<210> SEQ ID NO: 29
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 29
tggcctgtgc cactttct 18

<210> SEQ ID NO: 30
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 30
tggcctgtgc cactttcg 17

<210> SEQ ID NO: 31
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 31
-continued

tggctctgc catctctc 18
<210> SEQ ID NO 32
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 32
tggctctgc catcttgt 18
<210> SEQ ID NO 33
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 33
ccatgcctgg gacgag 17
<210> SEQ ID NO 34
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 34
ccatgcctgg gacgaga 17
<210> SEQ ID NO 35
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 35
gagcgccag cccc 14
<210> SEQ ID NO 36
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 36
gagcgccag ctcg 14
<210> SEQ ID NO 37
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 37
gcgcgttgg cggc 14
<210> SEQ ID NO 38
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 38
gcgcgttgg gtgg 14
<210> SEQ ID NO 39
<211> LENGTH: 13
-continued

<210> SEQ ID NO 40
<211> LENGTH: 13
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 40

ggcgcgcgt ttc

<210> SEQ ID NO 41
<211> LENGTH: 13
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 41

cgttctcctg ccc

<210> SEQ ID NO 42
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 42

cgttctcctg cgct

<210> SEQ ID NO 43
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 43

gasgsggcc aasggccg

<210> SEQ ID NO 44
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44

gasgsggcc aasgacgc

<210> SEQ ID NO 45
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45

tcagcagcgt gcct

<210> SEQ ID NO 46
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 46

cagcagcctg ccgcag
<210> SEQ ID NO 47
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 47
agttcttgag gcactgga

<210> SEQ ID NO 48
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 48
caagttcttg aggcaagtctgc ata

<210> SEQ ID NO 49
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 49
togcgggggg g

<210> SEQ ID NO 50
<211> LENGTH: 11
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 50
togcgggggg c

<210> SEQ ID NO 51
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 51
gsattggatc aggctgtag

<210> SEQ ID NO 52
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 52
agaastggat caggtcttgt

<210> SEQ ID NO 53
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (5)...(5)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 53
tggtaccca tacaaggcag at
<210> SEQ ID NO: 54
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 54
gttraccat acaaggcaqg cg 22

<210> SEQ ID NO: 55
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 55
cgtctgccct gtatggttaa 20

<210> SEQ ID NO: 56
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 56
cgtctgccct gtatggtga 20

<210> SEQ ID NO: 57
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 57
ggasgcccag gacatagg 18

<210> SEQ ID NO: 58
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 58
aggaagccca ggacataga 19

<210> SEQ ID NO: 59
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 59
tatgtcctgy cttccttta ta 22

<210> SEQ ID NO: 60
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 60
gtcttgccct tctttaetg 19

<210> SEQ ID NO: 61
<211> LENGTH: 21
gtctgtaat catgaccaccat

gtctgtaat catgaccaccagat

gtcttgrtg atsggcccgt

gtcttgrtg atsggccat

tgcgtgcag gtccctgt

tgcgtgcag tccctgc

tcgagtggcc aasaggc
<210> SEQ ID NO 68
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 68

ttttcagtgg gcasaaggt 19

<210> SEQ ID NO 69
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 69
tcaattacgc tttanaggtga gc 22

<210> SEQ ID NO 70
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 70
tcaattacgc tttanaggtga gg 22

<210> SEQ ID NO 71
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 71
cccasaacg tcaacoacac a 21

<210> SEQ ID NO 72
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 72
cccasaacg ttcacoacaa ct 22

<210> SEQ ID NO 73
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 73
gaqaatagtg ggcttcagag aqt 23

<210> SEQ ID NO 74
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 74
gaqaatagtg ggcttcagag att 23

<210> SEQ ID NO 75
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 75
-continued

gaaatgccc tctccgagaa actcg 25

<210> SEQ ID NO 76
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 76

gaaatgccc tctccgagaa actcc 26

<210> SEQ ID NO 77
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 77

cctcttctcc ccctcccaaa attct 25

<210> SEQ ID NO 78
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 78

cctcttcccc ccctcccaaa atccc 23

<210> SEQ ID NO 79
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 79

gcgggcttcc tctgtgacac g 21

<210> SEQ ID NO 80
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 80

gcgggcttcc tctgtgacac ca 22

<210> SEQ ID NO 81
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 81

gattgtaagc acccctgta 19

<210> SEQ ID NO 82
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 82
	tgtaagcac cccctgg 17

<210> SEQ ID NO 83
<211> LENGTH: 23

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 83

ccactatcat tgatttttc cca

<210> SEQ ID NO 84
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 84

ccactatcat tgatttttc ccg

<210> SEQ ID NO 85
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (7)...(7)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 85

agggcagrtat gggcttagaa gcacctga

<210> SEQ ID NO 86
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 86

gggcagrtatg gggcttagaag cactgg

<210> SEQ ID NO 87
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 87

agggcagga aqstgygccac tatcat

<210> SEQ ID NO 88
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 88

agggcagga aqstgygccact atccc

<210> SEQ ID NO 89
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 89

ggacctgtag cacgcgca
<210> SEQ ID NO 90
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 90

ggacctgtag cacgccgc

<210> SEQ ID NO 91
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)..(4)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 91
tggttagtgt gcagcccac q

<210> SEQ ID NO 92
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (4)..<(6)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 92
tggttagtgt gcagcccac ca

<210> SEQ ID NO 93
<211> LENGTH: 13
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 93

gggggcctgg tgg

<210> SEQ ID NO 94
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 94

agggggcctg gtga

<210> SEQ ID NO 95
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 95
ccocctgccca ctgccca

<210> SEQ ID NO 96
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 96
cocctgcacc tcgccg
<210> SEQ ID NO 97
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (15)...(15)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 97

cc tgtg cact ggcc ggcggcttg ggcacacttg

<210> SEQ ID NO 98
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (15)...(15)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 98

cc tgtg cact ggcc ggcggcttg ggcacacttt

<210> SEQ ID NO 99
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (14)...(14)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 99

cc tgtg cact ggc ggcggcttg ggcacactgt

<210> SEQ ID NO 100
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (14)...(14)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 100

cc tgtg cact ggc ggcggcttg ggcacactttct

<210> SEQ ID NO 101
<211> LENGTH: 13
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 101

cggccg a g t

<210> SEQ ID NO 102
<211> LENGTH: 13
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 102

cggccg cc a c t
<210> SEQ ID NO 103
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 103
tgacotcccc ctgcacct

<210> SEQ ID NO 104
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 104
tgacotcccc ctgcagct

<210> SEQ ID NO 105
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 105
gtctcactgg ctgagc

<210> SEQ ID NO 106
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 106
gtgtctcagt ggctggt

<210> SEQ ID NO 107
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 107
cgcqrggtt cacatc

<210> SEQ ID NO 108
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 108
cgcqrggtt cacacag

<210> SEQ ID NO 109
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (7)..<(7)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 109

gaggcgtcg tgtaccccg

<210> SEQ ID NO 110
<211> LENGTH: 10
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (8)..<(8)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 110
cagagcgtcg tgtaccccg

<210> SEQ ID NO 111
<211> LENGTH: 12
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 111
gcgtcggcg gt

<210> SEQ ID NO 112
<211> LENGTH: 13
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 112
gggtgtctgg cgt

<210> SEQ ID NO 113
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 113
gctgtgcacc atcacc

<210> SEQ ID NO 114
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 114
cgctgtgcacc ctcacatc

<210> SEQ ID NO 115
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)..<(3)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 115
ccrasacca ggtctggt g
<210> SEQ ID NO 116
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 116
ccrasacca ggtgcggtgat g

<210> SEQ ID NO 117
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 117
tggsgccgc gcgcpg

<210> SEQ ID NO 118
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 118
tggsgccgc gcgcpg

<210> SEQ ID NO 119
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 119
cgagtgcgt tctcctc

<210> SEQ ID NO 120
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 120
cgagtgcgt tctcctc
cagagcgcgt tctcctc

<210> SEQ ID NO 121
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (16)...(16)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 121
cagagcgcgt tctcctcctc

<210> SEQ ID NO 122
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (16)...(16)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 122

cagagcgct tcttcttg

<210> SEQ ID NO 123
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 123
tggcaagaa tgcgttggag gc

<210> SEQ ID NO 124
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 124
tggcaagaa tgcgttggag ga

<210> SEQ ID NO 125
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 125
gcrgcctct cggcacccc

<210> SEQ ID NO 126
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 126
gcrgcctct cggcaccac

<210> SEQ ID NO 127
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 127
gcrgcctct cggcaccacc t

<210> SEQ ID NO 128
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (3)...(3)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 128

gcgcctcct cgcgcacccca c 21

<210> SEQ ID NO 129
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 129
cgcacaaag gcgcgcagc 17

<210> SEQ ID NO 130
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 130
gcgcacaaag ggcgcgcagc 18

<210> SEQ ID NO 131
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 131
tggtgcgcct tgccgca 16

<210> SEQ ID NO 132
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 132
tgcggccctt gcgcgc 15

<210> SEQ ID NO 133
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (11)..<(11)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 133
ggcctgcgc racacactgcg 20

<210> SEQ ID NO 134
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (12)..<(12)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 134
ggcctgcgc cccacactccc t 21

<210> SEQ ID NO 135
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 135

catatccac ccgcag

<210> SEQ ID NO: 136
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 136

catatccac ccgcag

<210> SEQ ID NO: 137
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (5)...(5)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 137
	agagccctt gggcgc

<210> SEQ ID NO: 138
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (6)...(6)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 138

agagccctt tgggcga

<210> SEQ ID NO: 139
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 139

cggtttgtc cagagg

<210> SEQ ID NO: 140
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 140

acggtttgt cccagag

<210> SEQ ID NO: 141
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 141

ccacagcct gaggsagtt
-continued

<210> SEQ ID NO 142
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 142

cagcagctg aggaagc

<210> SEQ ID NO 143
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 143

tcaggtgct gacccagct cagga

<210> SEQ ID NO 144
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 144

caggtgctg gacccagctc agggs

<210> SEQ ID NO 145
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 145

tgcacctagc tcagggga

<210> SEQ ID NO 146
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 146

gcgcctagc caggggag

<210> SEQ ID NO 147
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (24)...(24)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 147

gctgcgtgac ctagctcagggagrac

<210> SEQ ID NO 148
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (24)...(24)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 148

gctgcgtgac ctagctcagggagrat
<210> SEQ ID NO 149
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 149
cocgaactctctttc
c
16

<210> SEQ ID NO 150
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 150
gccagctcctctttc
17

<210> SEQ ID NO 151
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (11)..(11)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 151
tocctctgca ratococagc c
t
21

<210> SEQ ID NO 152
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (12)..(12)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 152
gtcccctcgac scatocccac cgt
t
22

<210> SEQ ID NO 153
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 153
cggatagcc tcgttaactgac gacagg
ctggatgac tcggtaactgac gacagg
27

<210> SEQ ID NO 154
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 154
cggatagcc tcggtaactgagacagg
cggatagcc tcggtaactgagacagg
27

<210> SEQ ID NO 155
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 155
gtggtgtag ctgtaactg agcaca

<210> SEQ ID NO 156
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 156

gtggtgtag ctgtaactg agcagcg

<210> SEQ ID NO 157
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 157

gggacccagc ccagcaca

<210> SEQ ID NO 158
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 158

gggacccagc ccagcacc

<210> SEQ ID NO 159
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (9)..(9)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 159

gccagccg gccagagcc gagg

<210> SEQ ID NO 160
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (9)..(9)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 160

gccagccg ccgagagcct gact

<210> SEQ ID NO 161
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 161

tggcaagcct tgtctacctt t

<210> SEQ ID NO 162
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 162
tggcgccacctttcaccctc 19

<210> SEQ ID NO: 163
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 163
gttcaatga tgaagacctgt c 21

<210> SEQ ID NO: 164
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 164
ggtttcattgt atgaagacctgt gt 22

<210> SEQ ID NO: 165
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 165
gcagaagaca qtagcccac cactat 26

<210> SEQ ID NO: 166
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 166
gcagaagaca qtagcccac cactag 26

<210> SEQ ID NO: 167
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (20)..(21)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 167
tgggctcacc ctgcaacactcrgag 25

<210> SEQ ID NO: 168
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (17)..(18)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 168
gtcacgcctgcagacactgtca t 21

<210> SEQ ID NO: 169
<211> LENGTH: 22
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens  
<220> FEATURE:  
<221> NAME/KEY: modified_base  
<222> LOCATION: (17)..(18)  
<223> OTHER INFORMATION: I  
<400> SEQUENCE: 169  
gctacggtg cacatrggag 22  
<210> SEQ ID NO 170  
<211> LENGTH: 21  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens  
<220> FEATURE:  
<221> NAME/KEY: modified_base  
<222> LOCATION: (17)..(18)  
<223> OTHER INFORMATION: I  
<400> SEQUENCE: 170  
gctacggtg cacatrgga t 21  
<210> SEQ ID NO 171  
<211> LENGTH: 26  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens  
<220> FEATURE:  
<221> NAME/KEY: modified_base  
<222> LOCATION: (23)..(23)  
<223> OTHER INFORMATION: I  
<400> SEQUENCE: 171  
ggcccttcctg tctagatct acrtc 26  
<210> SEQ ID NO 172  
<211> LENGTH: 27  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens  
<220> FEATURE:  
<221> NAME/KEY: modified_base  
<222> LOCATION: (24)..(24)  
<223> OTHER INFORMATION: I  
<400> SEQUENCE: 172  
ggcgcccttcctg tctagatct tacrtct 27  
<210> SEQ ID NO 173  
<211> LENGTH: 20  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens  
<400> SEQUENCE: 173  
tgggtcaacg ctgaacatct 20  
<210> SEQ ID NO 174  
<211> LENGTH: 19  
<212> TYPE: DNA  
<213> ORGANISM: Homo sapiens  
<400> SEQUENCE: 174  
ggcgccctgc tgcacatcc 19  
<210> SEQ ID NO 175  
<211> LENGTH: 24
-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 175

gtgcacaca ggcgtacag gace

<210> SEQ ID NO: 176
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 176
tgctcaccag ggcgtacag acg

<210> SEQ ID NO: 177
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (9)..<(9)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 177
tcagcagctc gactagggcag gtcgag

<210> SEQ ID NO: 178
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (10)..<(10)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 178
tcagcagctc gactagggcag gtcgag

<210> SEQ ID NO: 179
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 179
tcagcgcctt tgtggacac

<210> SEQ ID NO: 180
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 180
gtgcagcgcct tgtggacac

<210> SEQ ID NO: 181
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (6)..<(6)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 181
-continued

ggacaraagt cccctgca 17

<210> SEQ ID NO 182
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> HNM/KEY: modified_base
<222> LOCATION: (6),(6)
<223> OTHER INFORMATION: I

<400> SEQUENCE: 182

ggacaraagt cccctggg 17

<210> SEQ ID NO 183
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 183

agacgcttc acttctcacc act 23

<210> SEQ ID NO 184
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 184

acggcctca acttctcacc c 21

<210> SEQ ID NO 185
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 185

cggacggatg ggcgctctcg ga 22

<210> SEQ ID NO 186
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 186

tgacgagatga ggcctctcgg g 21

<210> SEQ ID NO 187
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 187

ccttctgctt c cacccc 18

<210> SEQ ID NO 188
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 188

ccttctgctt caccoccg 18
<210> SEQ ID NO 189
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (13)..(13)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 189

cgtttcacc cccracactt cccg

<210> SEQ ID NO 190
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (14)..(14)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 190

ccgcttcac cccracactt tcctg

<210> SEQ ID NO 191
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 191

ccccccccca cagccccac

<210> SEQ ID NO 192
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 192

ccccccccac aggccccac

<210> SEQ ID NO 193
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 193

cagttggccac cgagaagctg

<210> SEQ ID NO 194
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 194

tccagttggc acogyagaagc ta

<210> SEQ ID NO 195
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 195
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
<th>Length</th>
<th>Type</th>
<th>Organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>ctgacctaa ctgctgcc</td>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 196</td>
<td>19</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORGANISM: Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQUENCE: 196</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ctgacctaa ctgctgcc</td>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 197</td>
<td>33</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORGANISM: Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQUENCE: 197</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cattctatac tgtatttat acaaaaatga gag</td>
<td>33</td>
<td>DNA</td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 198</td>
<td>33</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORGANISM: Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQUENCE: 198</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cattctatac tgtatttat acaaaaatga gac</td>
<td>33</td>
<td>DNA</td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 199</td>
<td>27</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORGANISM: Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQUENCE: 199</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ttctatsca tagttgcaaa tttgta</td>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 200</td>
<td>27</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORGANISM: Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQUENCE: 200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ttctatsca atagttgca attttata</td>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 201</td>
<td>27</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORGANISM: Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQUENCE: 201</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ttgcaacct taatattaaga acttota</td>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 202</td>
<td>27</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ORGANISM: Homo sapiens</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQUENCE: 202</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>attgcaacct atgatattaga aacctc</td>
<td></td>
<td></td>
<td>Homo sapiens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEQ ID NO 203</td>
<td>28</td>
<td>DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LENGTH: 28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 203

gtttgyttt acattaggg taaatitag  28

<210> SEQ ID NO: 204
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 204

gtttgyttt tacattaggg taaatitag  28

<210> SEQ ID NO: 205
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 205

gtgggttgag gtacogt  17

<210> SEQ ID NO: 206
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 206

gtgggttgag gtacoga  17

<210> SEQ ID NO: 207
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 207

tgccasagg csgga  15

<210> SEQ ID NO: 208
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 208

tgccasagg gcsagtt  16

<210> SEQ ID NO: 209
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 209

gocctttggc actgygt  16

<210> SEQ ID NO: 210
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 210

cocctttggc ctggc  15
<210> SEQ ID NO 211
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 211

ggagttcccc gttgtctaa

<210> SEQ ID NO 212
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 212

ggagttcccc gttgtctag

<210> SEQ ID NO 213
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 213

gggtcaccct ctttctcag

<210> SEQ ID NO 214
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 214

gggtcaccct ctttctcat

<210> SEQ ID NO 215
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 215

gttggctcgc gcgcaca

<210> SEQ ID NO 216
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 216

tgggtcgcga ggcaca

<210> SEQ ID NO 217
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 217

cctgccataa ccacacatgt

<210> SEQ ID NO 218
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 218
gtgcatacacc acaactgcg 18

<210> SEQ ID NO: 219
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 219
cacacccagc tga taaa s a tta 23

<210> SEQ ID NO: 220
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 220
cacacccagc tga taaa s a ttt 23

<210> SEQ ID NO: 221
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 221
tca aagca actcttc aas ctc 23

<210> SEQ ID NO: 222
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 222
tca aagca actcttc aas c tg 23

<210> SEQ ID NO: 223
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 223
tttctcc t a g g p c a c a g t c a 20

<210> SEQ ID NO: 224
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 224
tctct c t g g c a c a t c t g g 18

<210> SEQ ID NO: 225
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 225
ggcttgaas a g t c a c t g a c t t g 24
-continued

<210> SEQ ID NO 226
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 226

aattgcttga aataagcact gtacctta

<210> SEQ ID NO 227
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 227
gccatagaca cagggcga

<210> SEQ ID NO 228
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 228
gccatagaca cagggcag

<210> SEQ ID NO 229
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 229
ccatgacat tgaattggtt gttta

<210> SEQ ID NO 230
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 230
catgactatt gattgagttg tttg

<210> SEQ ID NO 231
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 231
goaggtaccc ttaggcttg

<210> SEQ ID NO 232
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 232
gggcagagg ccctgcta

<210> SEQ ID NO 233
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 233
-continued

tatccagctg ggacccaa 18

<210> SEQ ID NO: 234
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 234
ccagccccat ggtctt 16

<210> SEQ ID NO: 235
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 235
cagacgacc ccgcgttt 17

<210> SEQ ID NO: 236
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 236
cgggccgcc cgttt 14

<210> SEQ ID NO: 237
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 237
tgggttcgcc cctgaa 16

<210> SEQ ID NO: 238
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 238
tggtctatgc aggagtt 18

<210> SEQ ID NO: 239
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 239
gcacgccca ggtgaa 17

<210> SEQ ID NO: 240
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 240
catgcagcc caccaatg 17

<210> SEQ ID NO: 241
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 241
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 241

agccaccttg gtagtgagcg agc
23

<210> SEQ ID NO: 242
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 242

gacccaccttg gtagtgagcg cagc
24

<210> SEQ ID NO: 243
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (1)...(4)
<223> OTHER INFORMATION: I
<400> SEQUENCE: 243

aggtggtat tgacacacca cag
23

<210> SEQ ID NO: 244
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 244

attccaggtaa ttcacacag gc
22

<210> SEQ ID NO: 245
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 245

gacctgaggc gacotgttt
18

<210> SEQ ID NO: 246
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 246

gcacaagtgca gacotgott
18

<210> SEQ ID NO: 247
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 247

cagatgaaa gacacacatt
19

<210> SEQ ID NO: 248
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 248
aagcccccat tttgtaacca t\n22

<210> SEQ ID NO: 249
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 249
gcctggtggc atgaattttca a\n21

<210> SEQ ID NO: 250
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 250
cctgaasaa cccagaaccc tt\n22

<210> SEQ ID NO: 251
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 251
tacactgcca qqca\n14

<210> SEQ ID NO: 252
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 252
tctctgtctg tccttgqca\n20

<210> SEQ ID NO: 253
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 253
aaggtgcctca tggataggca g\n21

<210> SEQ ID NO: 254
<211> LENGTH: 15
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 254
cgagagggcca acgca\n15

<210> SEQ ID NO: 255
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 255
cccagaatgc tcacacagctg\n21
<210> SEQ ID NO 256
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 256

gccagcagc tttatttgcc

<210> SEQ ID NO 257
<211> LENGTH: 12
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 257

cagcgccggc ct

<210> SEQ ID NO 258
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 258

tctctgtctg tctottgcc cc

<210> SEQ ID NO 259
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 259

tctctgtctg tcctottgcc cca

<210> SEQ ID NO 260
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 260

tctctgtctg tcctottgcc caq

<210> SEQ ID NO 261
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 261

atacagcagc qtcgcgc ac a

<210> SEQ ID NO 262
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (12)...(12)
<223> OTHER INFORMATION: 2-thio-deoxythymidine
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (19)...(19)
<223> OTHER INFORMATION: 2-amino-adenosine

<400> SEQUENCE: 262

atacagcagc qtcgcgc ac a
<210> SEQ ID NO: 263
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (9)..(9)
<223> OTHER INFORMATION: 4-methylindole

<400> SEQUENCE: 263
atacaccat gtgcacacac a

<210> SEQ ID NO: 264
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (8)..(8)
<223> OTHER INFORMATION: 2-aminoadenosine

<400> SEQUENCE: 264
gcacgcacat tttttaggc

<210> SEQ ID NO: 265
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (8)..(8)
<223> OTHER INFORMATION: 2-aminoadenosine

<400> SEQUENCE: 265
gcacgcacat tttttaggc

<210> SEQ ID NO: 266
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 266
tgcttgaat tattacgtt cattta

<210> SEQ ID NO: 267
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 267
tgcttgaat tattacgtt cattta

<210> SEQ ID NO: 268
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 268

cacactgtg cacacacg

<210> SEQ ID NO: 269
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (16) (16)
<223> OTHER INFORMATION: 4-methylindole

<400> SEQUENCE: 269
cacactgtg cacacacg

<210> SEQ ID NO: 270
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 270
	ttgctgtag gtgtgacct

<210> SEQ ID NO: 271
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (8) (8)
<223> OTHER INFORMATION: 4-methylindole

<400> SEQUENCE: 271
	ttgctgtag gtgtgacct

<210> SEQ ID NO: 272
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (8) (8)
<223> OTHER INFORMATION: 5-nitroindole

<400> SEQUENCE: 272
	ttgctgtag gtgtgacct

<210> SEQ ID NO: 273
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: modified_base
<222> LOCATION: (8) (8)
<223> OTHER INFORMATION: abasic nucleotide

<400> SEQUENCE: 273
	ttgctgtag gtgtgacct

<210> SEQ ID NO: 274
<211> LENGTH: 14
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
1. A biochip comprising a solid substrate comprising an array comprising:
   a) at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY1A1;
   b) at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY1A2;
   c) at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY1B1;
   d) at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY2C19;
   e) at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY2D6;
   f) at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY2E1; and
   g) at least one capture probe substantially homologous to a portion of the sense strand of a nucleic acid encoding CPY3A4.

2. A biochip according to claim 1 wherein said first portion of said sense strand of said nucleic acid encoding CPY2D6 is adjacent to a single nucleotide polymorphism (SNP) position of interest.

3. A biochip according to claim 1 wherein said first portion of said sense strand of said nucleic acid encoding CPY2D6 includes at a terminus a single nucleotide polymorphism (SNP) position of interest.

4. A biochip according to claim 1 wherein said array further comprises at least one capture probe substantially homologous to a portion of the antisense strand of a nucleic acid encoding a protein selected from the group consisting of CYP1A1; CYP1A2; CYP1B1; CYP2C9; CYP2D6; CYP2E1 and CYP3A4.

5. A biochip according to claim 1 wherein said array further comprises at least one capture probe substantially homologous to a portion of a CYP pseudogene.

6. A biochip according to claim 1 wherein said solid support is selected from the group consisting of glass, plastic, ceramic, and PCB board.

7. A biochip according to claim 1 wherein said array comprises an array of electrodes.

8. A biochip according to claim 1 wherein said array comprises an array of polymer gel pads.

9. A method of determining the identification of a nucleotide at a detection position in at least one target sequence.
selected from the group consisting of CYP1A1, CYP1A2, CYP1B1, CYP2C19, CYP2D6, CYP2E1 and CYP3A4, said method comprising:

a) providing an array comprising:

i) at least one first capture probe substantially homologous to a first portion of a nucleic acid encoding CPY1A1, wherein said first capture probe is directly adjacent to or includes at its terminus a detection position;

ii) at least one second capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY1A2, wherein said second capture probe is directly adjacent to or includes at its terminus a detection position;

iii) at least one third capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY1B1, wherein said third capture probe is directly adjacent to or includes at its terminus a detection position;

iv) at least one fourth capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2C19, wherein said fourth capture probe is directly adjacent to or includes at its terminus a detection position;

v) at least one fifth capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2D6, wherein said fifth capture probe is directly adjacent to or includes at its terminus a detection position;

vi) at least one sixth capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY2E1, wherein said sixth capture probe is directly adjacent to or includes at its terminus a detection position; and

vii) at least one seventh capture probe substantially homologous to a first portion of the sense strand of a nucleic acid encoding CPY3A4, wherein said seventh capture probe is directly adjacent to or includes at its terminus a detection position;

b) hybridizing said target sequence to at least one of said capture probes to form a hybridization complex; and

c) determining the nucleotide at said detection position.

11. A method of determining the identification of a nucleotide at a detection position in a target sequence comprising:

a) providing a solid support with a first surface comprising at least one non self-extension probe wherein self-extension of said non self-extension probe does not occur in the absence of said target and wherein, said non self-extension probe includes an interrogation nucleotide;

b) hybridizing said target sequence to said non self-extension probe to form a hybridization complex;

c) contacting said surface with:

i) an extension enzyme; and

ii) at least one chain terminating nucleotide comprising a hapten;

under conditions whereby if said chain terminating nucleotide is perfectly complementary to the base of the target sequence immediately adjacent to the 3' end of said non self-extension probe in the hybridization complex, said chain terminating nucleotide is added to said non self-extension probe to form a modified extension probe;

d) contacting said modified extension probe with the binding partner of said hapten, wherein said hapten is labeled; and

e) detecting the presence of said label to determine the nucleotide at said detection position.

12. A method according to claim 11 wherein said interrogation nucleotide in said non self-extension probe is within two bases of its 3' terminal end and wherein, said 3' terminal end nucleotide is non-complementary to a corresponding base when a self-hybridizing structure of said non self-extension probe is formed.

13. A method according to claim 12 wherein said interrogation nucleotide is the 3' terminal nucleotide.

14. A method according to claim 12 wherein said interrogation nucleotide is the penultimate 3' terminal nucleotide.

15. A method according to claim 11 wherein said non self-extension probe comprises at least two modified nucleotides.

16. A method according to claim 15 wherein said modified nucleotides are exo-cyclic amine modified bases.

17. A method according to claim 15 wherein said modified nucleotides are terminator bases.

18. A method according to claim 16 wherein said exo-cyclic amine modified bases are selected from a group consisting of 2-thio thymine, 2-amino adenine, amine modified cytosine and amine modified guanine.

19. A method according to claim 17 wherein said terminator base is 4-methyl indole.

20. A method according to claim 15 wherein said modified nucleotides alter protein binding and are present in the stem region of said non self-extension probe.

21. A method according to claim 20 wherein said modified nucleotide comprises a sugar modification.

22. A method according to claim 20 wherein said modified nucleotide comprises a phosphate modification.
23. A method according to claim 22 wherein said phosphate modifications are selected from a group consisting of phosphorothioates, phosphoramidates, methyl phosphates, methyl phosphonates, H-phosphonates.

24. A method according to claim 11 wherein self-extension of said non self-extension probe is inhibited by short complementary oligonucleotides.

25. A method of determining the identification of a nucleotide at a detection position in a target sequence comprising:
   a) amplifying the target DNA using random primers to generate DNA amplicons;
   b) transcribing said DNA amplicons to generate RNA target sequences (in vitro transcription);
   c) providing a solid support with a first surface comprising at least one extension probe wherein said extension probe includes an interrogation nucleotide;
   b) hybridizing said RNA target sequence to said extension probe to form a hybridization complex;
   c) contacting said surface with:
      i) a modified reverse transcriptase; and
      ii) at least one chain terminating nucleotide comprising a hapten;
   under conditions whereby if said chain terminating nucleotide is perfectly complementary to the base of the target sequence immediately adjacent to the 3’ end of said non self-extension probe in the hybridization complex, said chain terminating nucleotide is added to said non self-extension probe to form a modified extension probe;
   d) contacting said modified extension probe with the binding partner of said hapten, wherein said binding partner is labeled; and
   e) detecting the presence of said label to determine the nucleotide at said detection position.

26. A method according to claim 25 wherein said modified reverse transcriptase only extends extension probes bound to RNA.

27. A method according to claim 11 or 25 wherein said hapten is biotin.

28. A method according to claim 11 or 25 wherein said binding partner is streptavidin.

29. A method according to claim 11 or 25 wherein said binding partner is Alcya dye labeled streptavidin.

30. A method of determining the identification of a nucleotide at a detection position in a target sequence comprising:
   a) providing a solid support with a first surface comprising a hydrogel layer comprising at least one non self-extension probe, wherein self-extension said non self-extension probe does not occur in the absence of said target and wherein, said non self-extension probe includes an interrogation nucleotide;
   b) hybridizing said target sequence to said non self-extension probe to form a hybridization complex;
   c) contacting said surface with:
      i) an extension enzyme; and
      ii) at least one chain terminating nucleotide comprising a hapten; under conditions whereby if said chain terminating nucleotide is perfectly complementary to the base of the target sequence immediately adjacent to the 3’ end of said non self-extension probe in the hybridization complex, said chain terminating nucleotide is added to said non self-extension probe to form a modified extension probe;
   d) contacting said modified extension probe with the binding partner of said hapten, wherein said binding partner is labeled; and
   e) detecting the presence of said label to determine the nucleotide at said detection position.

31. A method according to claim 30 wherein said interrogation nucleotide in said non self-extension probe is within two bases of its 3’ terminal end and wherein, said 3’ terminal end nucleotide is non-complementary to a corresponding base when a self-hybridizing structure of said non self-extension probe is formed.

32. A method according to claim 31 wherein said interrogation nucleotide is the 3’ terminal nucleotide.

33. A method according to claim 31 wherein said interrogation nucleotide is the penultimate 3’ terminal nucleotide.

34. A method according to claim 30 wherein said non self-extension probe comprises at least two modified nucleotides.

35. A method according to claim 34 wherein said modified nucleotides are exo-cyclic amine modified bases.

36. A method according to claim 34 wherein said modified nucleotides are terminator bases.

37. A method according to claim 35 wherein said exo-cyclic amine modified bases are selected from a group consisting of 2-thio thymine, 2-amino adenine, amine modified cytosine and amine modified guanine.

38. A method according to claim 36 wherein said terminator base is 4-methylcytosine.

39. A method according to claim 34 wherein said modified nucleotides alter protein binding and are present in the stem region of said non self-extension probe.

40. A method according to claim 39 wherein said modified nucleotide comprises a sugar modification.

41. A method according to claim 39 wherein said modified nucleotide comprises a phosphate modification.

42. A method according to claim 41 wherein said phosphate modifications are selected from a group consisting of phosphorothioates, phosphoramidates, methyl phosphonates, methyl phosphates, H-phosphonates.

43. A method according to claim 39 wherein self-extension of said non self-extension probe is inhibited by short complementary oligonucleotides.

44. A method of determining the identification of a nucleotide at a detection position in a target sequence comprising:
   a) amplifying the target DNA using random primers to generate DNA amplicons;
   b) transcribing said DNA amplicons to generate RNA target sequences (in vitro transcription);
   c) providing a solid support with a first surface comprising at least one extension probe wherein said extension probe includes an interrogation nucleotide;
b) hybridizing said RNA target sequence to said extension probe to form a hybridization complex;

c) contacting said surface with:

i) a modified reverse transcriptase; and

ii) at least one chain terminating nucleotide comprising a hapten;

under conditions whereby if said chain terminating nucleotide is perfectly complementary to the base of the target sequence immediately adjacent to the 3' end of said non self-extension probe in the hybridization complex, said chain terminating nucleotide is added to said non self-extension probe to form a modified extension probe;

d) contacting said modified extension probe with the binding partner of said hapten, wherein said binding partner is labeled; and

e) detecting the presence of said label to determine the nucleotide at said detection position.

45. A method according to claim 44 wherein said modified reverse transcriptase only extends extension probes bound to RNA.

46. A method according to claim 30 or 44 wherein said hapten is biotin.

47. A method according to claim 30 or 44 wherein said binding partner is streptavidin.

48. A method according to claim 30 or 44 wherein said binding partner is Alexa dye labeled streptavidin.