The NHGRI Sample Repository for Human Genetic Research

Coriell Institute for Medical Research, 403 Haddon Avenue, Camden, NJ 08103

Background

The NHGRI Sample Repository for Human Genetic Research (NHGRI Repository), established by NHGRI in 2006 at the Coriell Institute for Medical Research, facilitates studies of human genetic and genomic variation by characterizing and distributing biospecimens from thousands of participants living around the world (>3,700), including the 1000 Genomes Project samples, to the research community. NHGRI Repository DNA and lymphoblastoid cell lines (LCLs) are publicly available, renewable, and are characterized with publicly available whole genome sequencing data and other large-scale genomic data. NHGRI Repository cell lines are a unique resource that can be used for in vitro modeling, and as a reliable and renewable source of DNA and RNA. NHGRI Repository cell lines and DNAs provide a gold standard reference set for testing new assays and validating new platforms, pipelines and workflows, and as sequencing technologies change over time, NHGRI Repository samples can be used to confirm reproducible results, and enable comparison between different technologies and data analysis methods.

Hundreds of thousands of NHGRI Repository DNA and cell culture samples (>200,000) have been distributed to researchers (Figure 1) in 56 countries around the world supporting a wide range of research studies. The most commonly cited research topics include use as reference materials in genetic/genomic assay development and clinical genetic testing, mutation discovery and characterization, functional studies including a wide range of functional genomic studies, and evolutionary studies. A more detailed summary of the types of research using NHGRI Repository samples is shown in Figure 2. Finally, thousands of resulting scientific publications have used HapMap and 1000 Genomes Project data and samples. The NHGRI Repository website can be accessed at https://catalog.coriell.org/NHGRI.

Distribution by Research Category

The NHGRI Sample Repository for Human Genetic Research is supported by the National Human Genome Research Institute (U24HG008736).

Table 1: NHGRI Repository Population Samples

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>African Ancestry in SW, USA</td>
<td>62</td>
</tr>
<tr>
<td>2</td>
<td>African Caribbean in Barbados</td>
<td>120</td>
</tr>
<tr>
<td>3</td>
<td>Bengali in Bangladesh</td>
<td>144</td>
</tr>
<tr>
<td>4</td>
<td>British from England and Scotland</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>Chinese Dai in Xishuangbanna, China</td>
<td>102</td>
</tr>
<tr>
<td>6</td>
<td>Chinese in Metropolitan Denver, CO, USA</td>
<td>129</td>
</tr>
<tr>
<td>7</td>
<td>Colombian in Medellin, Colombia</td>
<td>136</td>
</tr>
<tr>
<td>8</td>
<td>Eesir in Nigeria</td>
<td>173</td>
</tr>
<tr>
<td>9</td>
<td>Finnish in Finland</td>
<td>103</td>
</tr>
<tr>
<td>10</td>
<td>Gambian in Western Division, Mandinka</td>
<td>179</td>
</tr>
<tr>
<td>11</td>
<td>Gujarati Indians in Houston, TX, USA</td>
<td>109</td>
</tr>
<tr>
<td>12</td>
<td>Han Chinese in Beijing, China</td>
<td>120</td>
</tr>
<tr>
<td>13</td>
<td>Han Chinese South</td>
<td>163</td>
</tr>
<tr>
<td>14</td>
<td>Iberian Populations in Spain</td>
<td>157</td>
</tr>
<tr>
<td>15</td>
<td>Indian Telugu in the UK</td>
<td>118</td>
</tr>
<tr>
<td>16</td>
<td>Japanese in Tokyo, Japan</td>
<td>120</td>
</tr>
<tr>
<td>17</td>
<td>Kinh in Ho Chi Minh City, Vietnam</td>
<td>124</td>
</tr>
<tr>
<td>18</td>
<td>Luhya in Webuye, Kenya</td>
<td>120</td>
</tr>
<tr>
<td>19</td>
<td>Maasai in Kinyawa, Kenya</td>
<td>405</td>
</tr>
<tr>
<td>20</td>
<td>Mende in Sierra Leone</td>
<td>128</td>
</tr>
<tr>
<td>21</td>
<td>Mexican Ancestry in Los Angeles, CA, USA</td>
<td>71</td>
</tr>
<tr>
<td>22</td>
<td>Peruvian in Lima, Peru</td>
<td>122</td>
</tr>
<tr>
<td>23</td>
<td>Puerto Rican in Puerto Rico</td>
<td>139</td>
</tr>
<tr>
<td>24</td>
<td>Punjabi in Lahore, Pakistan</td>
<td>158</td>
</tr>
<tr>
<td>25</td>
<td>Sri Lankan Tamil in the UK</td>
<td>128</td>
</tr>
<tr>
<td>26</td>
<td>Toscani in Italy</td>
<td>114</td>
</tr>
<tr>
<td>27</td>
<td>Yoruba in Ibadan, Nigeria</td>
<td>120</td>
</tr>
</tbody>
</table>

Figure 1. NHGRI Repository DNA and cell product distribution by population sample. Population IDs correspond to Table 1.

Figure 2. NHGRI Repository biospecimen distribution by research category.