A novel β-galactosidase gene family and DNA sequences derived from the cloning of cDNAs encoding products of these genes are provided, as exemplified by a β-galactosidase II protein which is encoded by a cDNA clone, pZBG2-1-4. A method for modifying cell wall metabolism which involves modifying the activity of at least one β-galactosidase, and thus modifying the quality of the fruit is also provided. Also provided by the present invention is a DNA construct including some or all of a β-galactosidase DNA sequence under control of a transcriptional initiation region operative in plants, so that the construct can generate RNA and, optionally, β-galactosidase polypeptide in plant cells. The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of β-galactosidase polypeptides or peptides by recombinant techniques. The present invention also provides plant cells containing DNA constructs of the present invention; plants derived therefrom having modified β-galactosidase gene expression; and seeds produced from such plants.
Fig. 1
Fig. 2

sheet 1 of 12
Fig. 2

sheet 2 of 12
Fig. 2
Fig. 2

sheet 4 of 12
Fig. 2

sheet 9 of 12
Fig. 2
<table>
<thead>
<tr>
<th>Location</th>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>306</td>
<td>ATG GGT TTT TGG ATG GCA ATG TTG CGT ATG TTG TTA TTG TGT TTA GGG ATT CCT GCC ATT GCT TCT</td>
<td>Met Gly Phe Trp Met Ala Met Leu Leu Leu Leu Cys Leu Trp Val Ser Cys Gly Ile Ala Ser</td>
</tr>
<tr>
<td>375</td>
<td>GTT TCA TAT GAC CAT AAA GCT ATC ATT GTA AAT GGA AAA AGA AAA ATT CTC ATT TCT GGA TCC ATT CAC</td>
<td>Val Ser Tyr Asp His Lys Ala Ile Ile Val Asn Gly Gln Arg Lys Ile Leu Ile Ser Gly Ser Ile His</td>
</tr>
<tr>
<td>444</td>
<td>TAC CCT AGA AGC ACC CCT GAG ATG TGG CCA GAT CCT ATT CAG AAG GCA AAA GAA GGG GGA GGT GAT GTT</td>
<td>Tyr Pro Arg Ser Thr Pro Glu Met Trp Pro Asp Leu Ile Gln Lys Ala Lys Glu Gly Gly Val Asp Val</td>
</tr>
<tr>
<td>513</td>
<td>ATA CAG ACT TAT GTT TTC TGG AAT GGG CAT GAG CCT GAA GAA GGG AAA TAT TAT TTT GAA GAG AGG TAT</td>
<td>Ile Gln Thr Tyr Val Phe Trp Asn Gly His Glu Pro Glu Glu Gly Lys Tyr Phe Glu Glu Arg Tyr</td>
</tr>
<tr>
<td>582</td>
<td>GAT TTA GTG AAG TTC ATT AAA GTG GTG CAA GAA GCA GGA CTT TAT GTG CAT CTT AGG ATT GGA CCT TAT</td>
<td>Asp Leu Val Lys Phe Ile Lys Val Val Gln Glu Ala Gly Leu Tyr Val His Leu Arg Ile Gly Pro Tyr</td>
</tr>
<tr>
<td>651</td>
<td>GCA TGT GCT GAA GAT TTT GGG GTT CCT GTT TGG CTG AAG TAT GGT CCA GGT ATT AGT TTC AGA</td>
<td>Ala Cys Ala Glu Trp Asn Phe Gly Pro Val Trp Leu Lys Tyr Val Pro Glu Gly Ile Ser Phe Arg</td>
</tr>
<tr>
<td>720</td>
<td>ACA AAC AAT GAG CCA TTC AAG GCT GCA ATG CAA AAG TTC ACT ACT AGG ATT GAT ATG ATG AAA GCA</td>
<td>Thr Asn Asn Glu Pro Phe Lys Ala Ala Met Gln Lys Phe Thr Thr Lys Ile Val Asp Met Met Lys Ala</td>
</tr>
<tr>
<td>789</td>
<td>GAA AAG CTC TAT GAA ACT CAG GGT GGT CCA ATT ATT CTA TCT CAG ATA GAA AAT GAA TAT GGA CCT ATG</td>
<td>Glu Lys Leu Tyr Glu Thr Glu Gly Gly Pro Ile Ile Leu Ser Glu Ile Glu Asn Gly Tyr Gly Pro Met</td>
</tr>
</tbody>
</table>

**FIG. 2A-1**
FIG. 2A-2
FIG. 2A-3
FIG. 2A-4
FIG. 2A-5
FIG. 2B-1
FIG. 2B-2
FIG. 2B-3
FIG. 2B-4
FIG. 2C-1

AGAGTTCTATTATTTTTGCTATTTTGAAA

AAGAGAAAAATAAAGTTAAGGGGGGAAAAAGTTTTCATTAGCTTTGCTTTAAGGCTAAAGAGGATAATTTTAC

ATG GGT TGT ACG CTT ATA CTA ATG TTG AAT GTG TGG GTG TTG GTG TCA TGG GTT TTT TCT GGA
1 Met Gly Cys Thr Leu Ile Leu Met Leu Asn Val Leu Leu Val Leu Gly Ser Trp Val Phe Ser Gly

ACA GCT TCT GTT TCA TAT GAC CAT AGG GCT ATT ATT GTA AAT GGA CAA AGA AGA ATA CTT ATT TCT GGT
191 Thr Ala Ser Val Ser Tyr Asp His Arg Ala Ile Ile Val Asn Gly Gln Arg Arg Ile Leu Ile Ser Gly

TCT GTT CAT TAT CCA AGA AGC ACT CCT GAG ATG TGG CCA GGT ATT ATTCAA AAG GCT AAA GAA GGA GGT
260 Ser Val His Tyr Pro Arg Ser Thr Pro Glu Met Trp Pro Gly Ile Ile Gln Lys Ala Lys Glu Gly Gly

GTA GAT GTG ATT CAG ACT TAT GTT TTC TGG AAT GGA CAT GAG CCT CAA CAA GGG AAA TAT TAT TTT GAA
329 Val Asp Val Ile Gln Thr Tyr Val Phe Trp Asn Gly His Glu Pro Gln Gln Gly Lys Tyr Tyr Phe Glu

GGG AGA TAT GAT TTA GTG AAG TTT ATT AAG CTG GTG CAC CAA GCA GGA CTT TAT GTC CAT CTT AGA GTT
398 Gly Arg Tyr Leu Val Tyr Phe Ile Leu Val His Gln Ala Gly Leu Tyr Val His Leu Arg Val

GGA CCT TAT GCT TGT GCT GAA TGG AAT TTT GGG GGC TTC CCT GCT TGG CTG AAA TAT GGT CCA GGT ATC
467 Gly Pro Tyr Ala Cys Ala Trp Asn Phe Gly Gly Gly Phe Gly Tyr Val Tyr Pro Glu Ile

AGT TTC AGA ACA GAT AAT GGA CCT TTC AAG GCT GCA ATG CAA AAA TTT ACT GCC AAG ATT GTC AAT ATG
536 Ser Phe Arg Thr Asp Asn Gly Pro Phe Lys Ala Ala Met Gln Lys Phe Thr Ala Lys Ile Val Asn Met

ATG AAA GCG GAA GTG TAT GAA ACT CAA GGG GGC CCA ATA ATT TTA TCT CAG ATT GAG AAT GAA TAT
605 Met Lys Ala Glu Arg Leu Tyr Glu Thr Glu Gly Gly Pro Ile Ile Leu Ser Gln Ile Glu Asn Gly Tyr

162
FIG. 2C-3
FIG. 2C-4
FIG. 2C-5
FIG. 2D-1
FIG. 2D-2
FIG. 2D-3
FIG. 2E-1
FIG. 2E-2
FIG. 2F-1
FIG. 2F-2
FIG. 2G-1
FIG. 2G-3
FIG. 2G-4
FIG. 2G-5
<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-24</td>
<td>TSG1-ORF</td>
</tr>
<tr>
<td>1-14</td>
<td>TSG2-ORF</td>
</tr>
<tr>
<td>1-20</td>
<td>TSG3-ORF</td>
</tr>
<tr>
<td>1-22</td>
<td>TSG4-ORF</td>
</tr>
<tr>
<td>1</td>
<td>TSG5-ORF</td>
</tr>
<tr>
<td>1</td>
<td>TSG6-ORF</td>
</tr>
<tr>
<td>1</td>
<td>TSG7-ORF</td>
</tr>
<tr>
<td>1-16</td>
<td>apple</td>
</tr>
<tr>
<td>1-20</td>
<td>asparagus</td>
</tr>
<tr>
<td>1-20</td>
<td>broccoli</td>
</tr>
<tr>
<td>1-12</td>
<td>Lupin</td>
</tr>
<tr>
<td>27</td>
<td>TSG1-ORF</td>
</tr>
<tr>
<td>37</td>
<td>TSG2-ORF</td>
</tr>
<tr>
<td>31</td>
<td>TSG3-ORF</td>
</tr>
<tr>
<td>29</td>
<td>TSG4-ORF</td>
</tr>
<tr>
<td>51</td>
<td>TSG5-ORF</td>
</tr>
<tr>
<td>51</td>
<td>TSG6-ORF</td>
</tr>
<tr>
<td>56</td>
<td>TSG7-ORF</td>
</tr>
<tr>
<td>77</td>
<td>TSG1-ORF</td>
</tr>
<tr>
<td>87</td>
<td>TSG2-ORF</td>
</tr>
<tr>
<td>81</td>
<td>TSG3-ORF</td>
</tr>
<tr>
<td>79</td>
<td>TSG4-ORF</td>
</tr>
<tr>
<td>101</td>
<td>TSG5-ORF</td>
</tr>
<tr>
<td>100</td>
<td>TSG6-ORF</td>
</tr>
<tr>
<td>76</td>
<td>TSG7-ORF</td>
</tr>
<tr>
<td>110</td>
<td>apple</td>
</tr>
<tr>
<td>80</td>
<td>carnation</td>
</tr>
<tr>
<td>85</td>
<td>asparagus</td>
</tr>
<tr>
<td>81</td>
<td>broccoli</td>
</tr>
<tr>
<td>89</td>
<td>Lupin</td>
</tr>
<tr>
<td>127</td>
<td>TSG1-ORF</td>
</tr>
<tr>
<td>137</td>
<td>TSG2-ORF</td>
</tr>
<tr>
<td>131</td>
<td>TSG3-ORF</td>
</tr>
<tr>
<td>129</td>
<td>TSG4-ORF</td>
</tr>
<tr>
<td>151</td>
<td>TSG5-ORF</td>
</tr>
<tr>
<td>151</td>
<td>TSG6-ORF</td>
</tr>
<tr>
<td>150</td>
<td>TSG7-ORF</td>
</tr>
<tr>
<td>149</td>
<td>apple</td>
</tr>
<tr>
<td>134</td>
<td>carnation</td>
</tr>
<tr>
<td>130</td>
<td>asparagus</td>
</tr>
<tr>
<td>130</td>
<td>broccoli</td>
</tr>
<tr>
<td>138</td>
<td>Lupin</td>
</tr>
<tr>
<td>176</td>
<td>TSG1-ORF</td>
</tr>
<tr>
<td>186</td>
<td>TSG2-ORF</td>
</tr>
<tr>
<td>180</td>
<td>TSG3-ORF</td>
</tr>
<tr>
<td>178</td>
<td>TSG4-ORF</td>
</tr>
<tr>
<td>200</td>
<td>TSG5-ORF</td>
</tr>
<tr>
<td>200</td>
<td>TSG6-ORF</td>
</tr>
<tr>
<td>199</td>
<td>TSG7-ORF</td>
</tr>
<tr>
<td>129</td>
<td>apple</td>
</tr>
<tr>
<td>184</td>
<td>carnation</td>
</tr>
<tr>
<td>180</td>
<td>asparagus</td>
</tr>
<tr>
<td>180</td>
<td>broccoli</td>
</tr>
<tr>
<td>188</td>
<td>Lupin</td>
</tr>
<tr>
<td>226</td>
<td>TSG1-ORF</td>
</tr>
<tr>
<td>236</td>
<td>TSG2-ORF</td>
</tr>
<tr>
<td>230</td>
<td>TSG3-ORF</td>
</tr>
<tr>
<td>228</td>
<td>TSG4-ORF</td>
</tr>
<tr>
<td>250</td>
<td>TSG5-ORF</td>
</tr>
<tr>
<td>250</td>
<td>TSG6-ORF</td>
</tr>
<tr>
<td>249</td>
<td>TSG7-ORF</td>
</tr>
<tr>
<td>229</td>
<td>apple</td>
</tr>
<tr>
<td>234</td>
<td>carnation</td>
</tr>
</tbody>
</table>

Fig. 3
<table>
<thead>
<tr>
<th>asparagus</th>
<th>broccoli</th>
<th>lupin</th>
</tr>
</thead>
<tbody>
<tr>
<td>181</td>
<td>181</td>
<td>189</td>
</tr>
<tr>
<td>182</td>
<td>183</td>
<td>190</td>
</tr>
<tr>
<td>183</td>
<td>184</td>
<td>191</td>
</tr>
<tr>
<td>184</td>
<td>185</td>
<td>192</td>
</tr>
<tr>
<td>185</td>
<td>186</td>
<td>193</td>
</tr>
<tr>
<td>186</td>
<td>187</td>
<td>194</td>
</tr>
<tr>
<td>187</td>
<td>188</td>
<td>195</td>
</tr>
<tr>
<td>188</td>
<td>189</td>
<td>196</td>
</tr>
<tr>
<td>189</td>
<td>190</td>
<td>197</td>
</tr>
<tr>
<td>190</td>
<td>191</td>
<td>198</td>
</tr>
<tr>
<td>191</td>
<td>192</td>
<td>199</td>
</tr>
<tr>
<td>192</td>
<td>193</td>
<td>200</td>
</tr>
<tr>
<td>193</td>
<td>194</td>
<td>201</td>
</tr>
<tr>
<td>194</td>
<td>195</td>
<td>202</td>
</tr>
<tr>
<td>195</td>
<td>196</td>
<td>203</td>
</tr>
<tr>
<td>196</td>
<td>197</td>
<td>204</td>
</tr>
<tr>
<td>197</td>
<td>198</td>
<td>205</td>
</tr>
<tr>
<td>198</td>
<td>199</td>
<td>206</td>
</tr>
<tr>
<td>199</td>
<td>200</td>
<td>207</td>
</tr>
<tr>
<td>200</td>
<td>201</td>
<td>208</td>
</tr>
<tr>
<td>201</td>
<td>202</td>
<td>209</td>
</tr>
<tr>
<td>202</td>
<td>203</td>
<td>210</td>
</tr>
<tr>
<td>203</td>
<td>204</td>
<td>211</td>
</tr>
<tr>
<td>204</td>
<td>205</td>
<td>212</td>
</tr>
<tr>
<td>205</td>
<td>206</td>
<td>213</td>
</tr>
<tr>
<td>206</td>
<td>207</td>
<td>214</td>
</tr>
<tr>
<td>207</td>
<td>208</td>
<td>215</td>
</tr>
<tr>
<td>208</td>
<td>209</td>
<td>216</td>
</tr>
<tr>
<td>209</td>
<td>210</td>
<td>217</td>
</tr>
<tr>
<td>210</td>
<td>211</td>
<td>218</td>
</tr>
<tr>
<td>211</td>
<td>212</td>
<td>219</td>
</tr>
<tr>
<td>212</td>
<td>213</td>
<td>220</td>
</tr>
<tr>
<td>213</td>
<td>214</td>
<td>221</td>
</tr>
<tr>
<td>214</td>
<td>215</td>
<td>222</td>
</tr>
<tr>
<td>215</td>
<td>216</td>
<td>223</td>
</tr>
<tr>
<td>216</td>
<td>217</td>
<td>224</td>
</tr>
<tr>
<td>217</td>
<td>218</td>
<td>225</td>
</tr>
<tr>
<td>218</td>
<td>219</td>
<td>226</td>
</tr>
<tr>
<td>219</td>
<td>220</td>
<td>227</td>
</tr>
<tr>
<td>220</td>
<td>221</td>
<td>228</td>
</tr>
<tr>
<td>221</td>
<td>222</td>
<td>229</td>
</tr>
<tr>
<td>222</td>
<td>223</td>
<td>230</td>
</tr>
<tr>
<td>223</td>
<td>224</td>
<td>231</td>
</tr>
<tr>
<td>224</td>
<td>225</td>
<td>232</td>
</tr>
<tr>
<td>225</td>
<td>226</td>
<td>233</td>
</tr>
<tr>
<td>226</td>
<td>227</td>
<td>234</td>
</tr>
<tr>
<td>227</td>
<td>228</td>
<td>235</td>
</tr>
<tr>
<td>228</td>
<td>229</td>
<td>236</td>
</tr>
<tr>
<td>229</td>
<td>230</td>
<td>237</td>
</tr>
<tr>
<td>230</td>
<td>231</td>
<td>238</td>
</tr>
</tbody>
</table>

**Fig. 3**

Sheet 2 of 4
<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>Description</th>
<th>Allele</th>
<th>Allele Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBG1-ORF</td>
<td>DQ821334</td>
<td>510-550</td>
<td>500-560</td>
<td></td>
</tr>
<tr>
<td>TBG2-ORF</td>
<td>DQ821335</td>
<td>560-600</td>
<td>576-626</td>
<td></td>
</tr>
<tr>
<td>TBG3-ORF</td>
<td>DQ821336</td>
<td>610-650</td>
<td>626-676</td>
<td></td>
</tr>
<tr>
<td>TBG4-ORF</td>
<td>DQ821337</td>
<td>660-710</td>
<td>676-726</td>
<td></td>
</tr>
<tr>
<td>TBG5-ORF</td>
<td>DQ821338</td>
<td>720-770</td>
<td>726-776</td>
<td></td>
</tr>
<tr>
<td>TBG6-ORF</td>
<td>DQ821339</td>
<td>780-830</td>
<td>776-826</td>
<td></td>
</tr>
<tr>
<td>TBG7-ORF</td>
<td>DQ821340</td>
<td>840-890</td>
<td>826-876</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 3**
Sheet 3 of 4
<table>
<thead>
<tr>
<th>Plant</th>
<th>Accession</th>
<th>Aminogram</th>
<th>Protein</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBG2-ORF</td>
<td>687</td>
<td>GSTIDPAADP</td>
<td>SKEKDGKW</td>
<td>VPAP-NKS-MC-PDGE-VAIN</td>
</tr>
<tr>
<td>TBG3-ORF</td>
<td>681</td>
<td>AGNDPLAAL</td>
<td>NTHKQYWT</td>
<td>LVSLQGRF</td>
</tr>
<tr>
<td>TBG4-ORF</td>
<td>679</td>
<td>GONDPLAAL</td>
<td>ASGKGQTV</td>
<td>AGLSGHMP</td>
</tr>
<tr>
<td>TBG5-ORF</td>
<td>701</td>
<td>CSGKLGDP</td>
<td>KREWQY</td>
<td>SASHGEK-PRLG</td>
</tr>
<tr>
<td>TBG6-ORF</td>
<td>701</td>
<td>CSGKLGDP</td>
<td>KREWQY</td>
<td>SASHGEK-PRLG</td>
</tr>
<tr>
<td>TBG7-ORF</td>
<td>700</td>
<td>PGNEFALDP</td>
<td>DHHCKAVL</td>
<td>NEKCKTVP-RKTSKMW</td>
</tr>
<tr>
<td>apple</td>
<td>680</td>
<td>PGDAFPLAAL</td>
<td>GSHNGTQTV</td>
<td>NEKCKTVP-RKTSKMW</td>
</tr>
<tr>
<td>carnation</td>
<td>685</td>
<td>GONDPLAAL</td>
<td>ASGKGQTV</td>
<td>AGLSGHMP</td>
</tr>
<tr>
<td>asparagus</td>
<td>681</td>
<td>PGNEFALDP</td>
<td>DHHCKAVL</td>
<td>NEKCKTVP-RKTSKMW</td>
</tr>
<tr>
<td>broccoli</td>
<td>681</td>
<td>PGNEFALDP</td>
<td>DHHCKAVL</td>
<td>NEKCKTVP-RKTSKMW</td>
</tr>
<tr>
<td>Lupin</td>
<td>689</td>
<td>AGNDPLAAL</td>
<td>NTHKQYWT</td>
<td>LVSLQGRF</td>
</tr>
<tr>
<td>TBG1-ORF</td>
<td>727</td>
<td>ECKHYLSE</td>
<td>GQCDWVLE</td>
<td>SAPTYGAELL-VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG2-ORF</td>
<td>737</td>
<td>ECKHYLSE</td>
<td>GQCDWVLE</td>
<td>SAPTYGAELL-VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG3-ORF</td>
<td>731</td>
<td>ECKHYLSE</td>
<td>GQCDWVLE</td>
<td>SAPTYGAELL-VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG4-ORF</td>
<td>729</td>
<td>ECKHYLSE</td>
<td>GQCDWVLE</td>
<td>SAPTYGAELL-VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG5-ORF</td>
<td>751</td>
<td>ECKHYLSE</td>
<td>GQCDWVLE</td>
<td>SAPTYGAELL-VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG6-ORF</td>
<td>751</td>
<td>ECKHYLSE</td>
<td>GQCDWVLE</td>
<td>SAPTYGAELL-VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG7-ORF</td>
<td>750</td>
<td>ECKHYLSE</td>
<td>GQCDWVLE</td>
<td>SAPTYGAELL-VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>apple</td>
<td>730</td>
<td>DCKHLYSE</td>
<td>EAPYNDTS</td>
<td>DSYDAP-VN-PEPNS-DYITVYIG</td>
</tr>
<tr>
<td>carnation</td>
<td>735</td>
<td>DCKHLYSE</td>
<td>EAPYNDTS</td>
<td>DSYDAP-VN-PEPNS-DYITVYIG</td>
</tr>
<tr>
<td>asparagus</td>
<td>731</td>
<td>DCKHLYSE</td>
<td>EAPYNDTS</td>
<td>DSYDAP-VN-PEPNS-DYITVYIG</td>
</tr>
<tr>
<td>broccoli</td>
<td>731</td>
<td>DCKHLYSE</td>
<td>EAPYNDTS</td>
<td>DSYDAP-VN-PEPNS-DYITVYIG</td>
</tr>
<tr>
<td>Lupin</td>
<td>739</td>
<td>DCKHLYSE</td>
<td>EAPYNDTS</td>
<td>DSYDAP-VN-PEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG1-ORF</td>
<td>777</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG2-ORF</td>
<td>787</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG3-ORF</td>
<td>781</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG4-ORF</td>
<td>779</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG5-ORF</td>
<td>801</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG6-ORF</td>
<td>801</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG7-ORF</td>
<td>800</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>apple</td>
<td>780</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>carnation</td>
<td>785</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>asparagus</td>
<td>781</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>broccoli</td>
<td>781</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>Lupin</td>
<td>789</td>
<td>AGCDWVLE</td>
<td>SAPTYGAELL</td>
<td>VPVPEPNS-DYITVYIG</td>
</tr>
<tr>
<td>TBG1-ORF</td>
<td>827</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>TBG2-ORF</td>
<td>837</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>TBG3-ORF</td>
<td>831</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>TBG4-ORF</td>
<td>829</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>TBG5-ORF</td>
<td>851</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>TBG6-ORF</td>
<td>851</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>TBG7-ORF</td>
<td>850</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>apple</td>
<td>830</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>carnation</td>
<td>835</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>asparagus</td>
<td>831</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>broccoli</td>
<td>831</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>Lupin</td>
<td>839</td>
<td>KFASQPQEG</td>
<td>VCNTQGGC</td>
<td>RCPQYDKF</td>
</tr>
<tr>
<td>TBG1-ORF</td>
<td>877</td>
<td>ENRQGSEP-CR</td>
<td>INVLLKELSVE</td>
<td>XGS-</td>
</tr>
<tr>
<td>TBG2-ORF</td>
<td>887</td>
<td>GVFGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>TBG3-ORF</td>
<td>881</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>TBG4-ORF</td>
<td>879</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>TBG5-ORF</td>
<td>901</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>TBG6-ORF</td>
<td>901</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>TBG7-ORF</td>
<td>900</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>apple</td>
<td>880</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>carnation</td>
<td>885</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>asparagus</td>
<td>881</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
<tr>
<td>broccoli</td>
<td>881</td>
<td>EVQGQ-Sep-CR</td>
<td>HIVSNLAQV</td>
<td>KCSPPDDSST</td>
</tr>
</tbody>
</table>

**Fig. 3**

Sheet 4 of 4
<table>
<thead>
<tr>
<th>10</th>
<th>20</th>
<th>30</th>
<th>35</th>
<th>40</th>
<th>Br</th>
<th>Tr</th>
<th>Pk</th>
<th>Rd</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBG1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBG2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBG3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBG4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBG5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBG6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBG7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 5**
Fig. 6
Fig. 7
Fig. 8
Fig. 10
Flat plate compression to 3 mm
Breaker + 7 d

* Standard Deviation

Fig. 11a
Fig. 11b
4-mm cylindrical indentor to 3 mm
Breaker + 7 d

Fig. 11c
4-mm cylindrical puncture to 10 mm Breaker + 7 d

Fig. 11d
Fig. 11e
Fig. 12
Fig. 13

NOS-P = nopaline synthase promoter
NOS-T = nopaline synthase terminator
NPTII = neomycin phosphotransferase
GUS = beta-glucuronidase
2-1-4 cDNA = TBG4 gene
CaMV35S-P = cauliflower mosaic virus 35S promoter

pBII21 cut with SacI and BamHI
to make pBII21ΔGUS (11 kb)

pZBG2-1-4 cut with SacI and partially digested with BamHI
to make 1.55 kb 2-1-4Δ2 fragment

p2-1-4Δ2BII21ΔGUS (12.5 kb)
GENES CODING FOR TOMATO
BETA-GALACTOSIDASE POLYPEPTIDES

FIELD OF THE INVENTION

[0001] The present invention relates to a family of novel plant genes encoding polypeptides characterized by their ability to hydrolyze terminal non-reducing β-D-galactosyl residues from β-D-galactosides. More specifically, a polypeptide sequence derived from a cDNA clone designated pZBG2-1-4 (referred to in U.S. Provisional Application No. 60/088,805 as pTomgal 4), which encodes a specific plant polypeptide named β-galactosidase II, is provided. Also provided are cDNA clones encoding six other homologous polypeptides, methods of using these cDNA clones for producing β-D-galactosidase polypeptides of the invention, and methods of modifying fruit quality by employment of a polynucleotide or polypeptide of the present invention.

BACKGROUND OF THE INVENTION

[0002] The most conspicuous and important processes related to post-harvest quality of climacteric fruit are the changes in texture, color, taste, and aroma which occur during ripening. Because of the critical relationship that deleterious changes in texture have to quality and post-harvest shelf-life, emphasis has been placed on studying the mechanisms involved in the loss of firmness that occurs during tomato fruit ripening. Although fruit softening may involve changes in turgor pressure, anatomical characteristics and cell wall integrity, it is generally assumed that cell wall disassembly leading to a loss of wall integrity is a critical feature. The most apparent changes, in terms of composition and size, occur in the pectic fraction of the cell wall (see references in Seymour and Gross, 1996. Postharvest News Info. 7: 45N-52N).


[0004] Among the other known pectin modifications that occur during fruit development, one of the best characterized is the significant net loss of galactosyl residues which occurs in the cell walls of many ripening fruit (Gross and Sams, 1984. Phytochem. 23: 2457-2461; Seymour and Gross, supra). Although some loss of galactosyl residues could result indirectly from the action of PG, β-galactosidase (exo-β(1→4)-D-galactopranosidase; E.C. 3.2.1.23) is the only enzyme identified in higher plants capable of directly cleaving β(1→4) galactan bonds, and probably plays a role in galactan sidechain loss (DeVau et al. 1993. Physiol. Plant 87: 279-285; Carey et al. 1995. Plant Physiol. 108: 1099-1107; Carrington and Pressey, 1996. J. Amer. Soc. Hort. Sci. 121: 132-136). No endo-acting galactanase has yet been identified in higher plants. The view that β-galactosidase is active in releasing galactosyl residues from the cell wall during ripening is supported by the dramatic increase in free galactose, a product of β-galactosidase activity (Gross, K. C. 1984. Physiol. Plant 62: 25-32) and a concomitant increase in activity of a particular enzyme, designated β-galactosidase II, in tomatoes during ripening (Carey et al., supra). β-galactosidase activity is thought to be important in cell wall metabolism (Carey et al., supra). β-galactosidases are generally assayed using artificial substrates such as p-nitrophenyl-β-D-galactopyranoside (PNP), 4-methylumbelliferyl-β-D-galactopyranoside and 5-bromo-4-chloro-3-indoxyl-β-D-galactopyranoside (X-GAL). However, it is clear that β-galactosidase II is also active against natural substrates, i.e., β(1→4) galactan (Carey et al., supra; Carrington and Pressey, supra; Pressey, R. 1983. Plant Physiol. 71: 132-135). β-galactosidase proteins have been purified and characterized in a number of other fruits including kiwifruits (Ross et al. 1993. Planta 180, 499-506; coffee (Golden et al. 1993. Phytochem. 34: 355-360); persimmon (Kang et al. 1994. Plant Physiol. 105:975-979), and apple (Ross et al. 1994. Plant Physiol. 106: 521-528).

[0005] Carey et al. (supra) were able to purify one of the three previously identified β-galactosidases from ripening tomato fruit (Pressey, supra), but only one β-galactosidase II was active against β(1→4) galactan. Even though they were able to identify putative β-galactosidase cDNA clones, none of the cDNA’s deduced amino acid sequences matched the amino terminal sequence of the β-galactosidase II protein. Although β-galactosidase II, a protein present in tomato (Lycopersicon esculentum Mill.) fruit during ripening and capable of degrading tomato fruit galactan has been purified, cloning of the corresponding gene has been elusive.

[0006] The modification of plant gene expression has been achieved by several methods. The molecular biologist can choose from a range of known methods to decrease or increase gene expression or to alter the spatial or temporal expression of a particular gene. For example, the expression of either specific antisense RNA or partial (truncated) sense RNA has been utilized to reduce the expression of various target genes in plants (as reviewed by Bird and Ray, 1991. Biotechnology and Genetic-Engineering Reviews 9: 207-227). These techniques involve the incorporation into the genome of the plant of a synthetic gene designed to express either antisense or sense RNA. They have been successfully used to down-regulate the expression of a range of individual genes involved in the development and ripening of tomato fruit (Gray et al. 1992. Plant Molecular Biol. 9: 69-87). Methods to increase the expression of a target gene have also been developed. For example, additional genes designed to express RNA containing the complete coding region of the target gene may be incorporated into the genome of the plant to “over-express” the gene product. Various other methods to modify gene expression are known; for example, the use of alternative regulatory sequences. The complete disclosure of each of the references cited above is fully incorporated herein by reference.
The need therefore exists to clone a gene for β-galactosidase II and related polypeptides, and using known methods of modification of plant gene expression, thereby to provide methods for modifying quality of fruits, particularly by modifying the cell wall, thereby directly affecting the ripening of the fruit.

SUMMARY OF THE INVENTION

The present invention is based on the discovery of novel DNA sequences derived from cDNA clones from a family of genes encoding β-galactosidases. The phylogenetic tree based on the shared amino acid sequence identities for the DNA sequences of the present invention is shown in FIG. 1A, B. Five cDNA and two RT-PCR clones, designated herein as TBG1, TBG2, TBG3, TBG4, TBG5, TBG6, and TBG7 and having the nucleic acid sequences designated SEQ ID Nos. 1-7, respectively as shown in FIG. 2, were identified which had a high degree of shared sequence identity to other known β-galactosidas. The corresponding amino acid sequences are designated herein as SEQ ID Nos. 8-16, respectively and are shown in FIGS. 2 and 3.

The nucleotide sequences for SEQ ID Nos. 1-7 are recorded in GenBank with the following respective Accession Numbers:

SEQ ID NO: 1 TGB1 AF023847 deposited Sep. 10, 1997
SEQ ID NO: 2 TGB2 AF154420 deposited May 19, 1999
SEQ ID NO: 3 TGB3 AF154421 deposited May 20, 1999
SEQ ID NO: 4 TGB4 AF020390 deposited Aug. 21, 1997
SEQ ID NO: 5 TGB5 AF154423 deposited May 20, 1999
SEQ ID NO: 6 TGB6 AF154424 deposited May 20, 1999
SEQ ID NO: 7 TGB7 AF154422 deposited May 20, 1999

Throughout the following discussion, wherever TBG4 is indicated in the description of the invention, it is to be understood that TBG1-3 and 5-7 are also to be included in that description, unless otherwise indicated.

A method of providing a DNA sequence of the invention, either by cloning a cDNA (for instance, p2BG2-1-4) that codes for a protein of the present invention, such as β-galactosidase II, or by deriving the DNA sequence from genomic DNA, or by synthesis of a DNA sequence ab initio using the cDNA sequence as a guide is also provided.

A method for modifying cell wall metabolism which involves modifying the activity of at least one galactosidase, and thus modifying the quality of the fruit is also provided.

Also provided by the present invention is a DNA construct including some or all of an exemplary β-galactosidase DNA sequence under control of a transcriptional initiation region operative in plants, so that the construct can generate RNA in plant cells.

Also discovered is a promoter gene associated with expression of the genes encoding β-galactosidase.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of β-galactosidase polypeptides or peptides by recombinant techniques.

The present invention also provides plant cells containing DNA constructs of the present invention; plants derived therefrom having modified β-galactosidase gene expression; and seeds produced from such plants.

The β-galactosidase II protein of the present invention has demonstrated enzyme activity in cell wall disassembly leading to loss of tissue integrity and fruit softening. The β-galactosidase II protein also may be involved in cell turnover, which could be involved in cell extension and/or expansion and therefore plant growth and development.

By hydrolyzing galactose from the cell wall, the enzyme may allow ripening to commence and/or progress, since galactose may be involved in stimulating ethylene production alone or in conjunction with unconjugated N-glycans.

The β-galactosidase of the invention may be involved in conversion of chloroplasts (green—chlorophyll) to chromoplasts (red—lycopene) during fruit ripening by degrading chloroplast membrane galactolipids.

The family of genes represented by the nucleotide sequences shown in FIG. 2 is expected to code for a group of similar enzymes with the same type of hydrolytic activity but with different tissue and/or substrate specificities or cellular compartmentation profiles.

The β-galactosidase II protein of the present invention as well as other proteins encoded in the nucleotide sequences shown in FIG. 2 may be used for preparation of pectin and other cell wall derived polymers with lowered galactosyl content for use in biotests and solutions (for example in clarification of fruit juices) requiring lower or higher cross-linking or viscometric properties.

The present invention also provides β-galactosidase enzymes for use as components of enzyme mixtures for protoplast isolation.

BRIEF DESCRIPTION OF THE FIGURES

FIGS. 1A and 1B shows a phylogenetic tree based on shared amino acid sequence identity among tomato β-galactosidase clones TGB1-7 and other known plant β-galactosidase polypeptides.

FIG. 2 shows cDNA sequences [SEQ ID Nos: 1-7, respectively] for the seven β-galactosidase genes of the invention: TGB1, TGB2, TGB3, TGB4, TGB5, TGB6, TGB7.

FIG. 3 shows multiple sequence alignment of the N-terminal amino acid sequences of β-galactosidase II protein from tomato fruit for TGB1, TGB2, TGB3, TGB4, TGB5, TGB6 and TGB7 [SEQ ID Nos: 8-16, respectively] and various plant β-galactosidase cDNA clones.
FIG. 4 shows autoradiograph of northern blot analysis of TBG expression in various plant tissues (flowers, leaves, roots and stems).

FIG. 5 shows autoradiograph of northern blot analysis of TBG expression in fruit tissues at different stages of development.

FIG. 6 shows autoradiograph of northern blot analysis of TBG expression in fruit tissues (mature green or turning stage fruit peel, outer pericarp, inner pericarp and locular).

FIG. 7 shows autoradiograph of northern blot analysis of TBG expression in normal and mutant fruit tissues.

FIG. 8 shows autoradiograph of northern blot analysis of TBG expression in response to ethylene treatment of mature green fruit tissues.

FIG. 9 shows Western blot analysis of TBG expression by yeast.

FIG. 10 shows detection of β-galactosidase activity from pZBG2-1-4 expression in E. coli.

FIGS. 11A-E shows the comparative results of texture measurements for fruit from tomato plants containing antisense constructs to suppress TBG4 mRNA and fruit from the parental line.

FIGS. 12A-B show Northern blot analysis of TBG4 expression in transgenic fruit containing TBG4 antisense construct.

FIG. 13 shows a Binary construct used to transform plants and express TBG4 (pZBG2-1-4) in the antisense orientation.

DETAILED DESCRIPTION

The following detailed description is directed to a preferred embodiment of the present invention and is intended as illustrative of each of other DNA sequences of the present invention.

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding β-galactosidase polypeptides, particularly a β-galactosidase II polypeptide having the amino acid sequence shown in FIG. 2. The DNA sequence of the exemplary β-galactosidase II cDNA clone of the invention, which was determined from a cDNA clone, pZBG2-1-4, encoding β-galactosidase II, is recorded in GenBank as Accession Number AF020390. Not all β-galactosidases possess in vitro activity against extracted cell wall material via the release of galactose from wall polymers containing β (1→4)-D-galactan. The polypeptide expressed from the exemplary β-galactosidase II clone, pZBG2-1-4, has been shown to exhibit β-galactosidase activity and exogalactanase activity.

The exemplary β-galactosidase II protein of the present invention, as shown in FIG. 2, shares sequence homology with the amino acid sequence deduced from β-galactosidase cDNA clones of TBG2-7 and cDNA clones of the fruits of asparagus (accession number P45582), apple (accession number P48981), and cation (accession number Q00662), as well as with β-galactosidase cDNA clones of a previously published sequence of a tomato β-galactosidase cDNA clone designated pIombgal 61 (accession number P48986) isolated from ripe ‘Ailsa Craig’ fruit (Carey et al., supra). The ORF of the clone TBG1 disclosed herein by the inventors (accession number AF023847) is nearly identical to the cDNA previously described by Carey et al. As shown in FIG. 2, the shared deduced sequence identity is high among all the published plant β-galactosidases of the seven clones (TBG1-7) and the other plant β-galactosidases. BLAST searches of the database also indicated significant shared sequence identity between domains of the plant β-galactosidases and mammalian and fungal β-galactosidases, however little share sequence identity was detected with bacterial β-galactosidases.

As shown in FIG. 1, the shared amino acid identity of TBG1 and TBG3 was high. TBG4 was also very similar to both TBG1 and 3. The amino acid sequences of TBG2 and 7 were unique because several regions of amino acid insertions appear throughout their sequence (FIG. 2).

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using a PCR-based deoxyribonucleotide terminator protocol and an ABI automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc., Foster City, Calif.), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

By “nucleotide sequence” of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U).

Using the information provided herein, such as the exemplary nucleotide sequence shown in FIG. 2[SEQ ID NO: 4], a nucleic acid molecule of the present invention encoding a β-galactosidase II polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in FIG. 2[SEQ ID NO: 4] was discovered in a cDNA library derived from breeder, turning and pink fruit pericarp from ‘Rutgers’ tomato plants. A second library comprised poly(A)RNA isolated from all fruit tissues (except seeds) from immature green, mature green, breaker, turning, pink, red-ripe and over-ripe fruit of ‘Rutgers’ plants.

The complete sequence of the cDNA insert of pZBG2-1-4 is accessible in the GenBank (no. AF020390)
and is provided in FIG. 2[SEQ ID NO: 4]. The cDNA insert is 2532 nucleotides (nt) long and contains a single, long open reading frame (ORF) predicted to start with the first in-frame ATG at nt 64 and end with a TAA at nt 2238. This ORF codes for a 79 kD protein 724 amino acids long. The deduced amino acid sequence of pZBG2-1-4 shared significant amino acid identity to all published plant \(-\)-galactosidase sequences in the database (FIGS. 1A.B). When the entire ORF of each \(-\)-galactosidase gene was compared to pZBG2-1-4, the shared sequence identity was about 64% for tomato pTombgal 1 (P48980), about 67.6% for apple (P48981), about 63.5% for asparagus (P45582) and about 55% for carnation (Q00662). As one of ordinary skill would appreciate, due to the possibilities of sequencing errors discussed above, the actual complete \(-\)-galactosidase II polypeptide encoded by the deposited cDNA, which comprises about 724 amino acids, may be somewhat longer or shorter. More generally, the actual open reading frame may be anywhere in the range of \(\pm\)20 amino acids, more likely in the range of \(\pm\)10 amino acids, of that predicted from either the first methionine codon from the N-terminus shown in FIG. 2[SEQ ID NO: 4]. In any event, as discussed further below, the invention further provides polypeptides having various residues deleted from the N-terminus of the complete polypeptide, including polypeptides lacking one or more amino acids from the N-terminus of the \(-\)-galactosidase II polypeptide described herein.

[0053] Leader and Mature Sequences

[0054] Analysis of the deduced amino acid sequence of pZBG2-1-4 suggested a high probability for secretion based on the presence of a hydrophobic leader sequence, a leader sequence cleavage site and three possible N-glycosylation sites. The programs PSORT V6.4 (Nakai and Kano, 1992. Genomics 14, 897-911, incorporated herein by reference) and SignalP V1.1 (Nielsen et al. 1997. Protein Engineering 10: 1-6, incorporated herein by reference), were used to predict that the ORF contains a hydrophobic leader sequence that would be cleaved between the alanine and serine residues at positions 23 and 24 respectively, and that the mature polypeptide has an extracellular location. The mature polypeptide contains three possible N-glycosylation sites at asparagine numbers 282, 459 and 713, however the asparagine position 713 is unlikely to be glycosylated due to the proline at position 714. The predicted molecular mass of the unglycosylated mature polypeptide was 75 kD with a pi of 8.9.

[0055] Accordingly, the amino acid sequence of the complete \(-\)-galactosidase II protein of the invention includes a leader sequence and a mature protein, as shown in FIG. 3[SEQ ID NO: 4]. More in particular, the present invention provides nucleic acid molecules encoding a mature form of the \(-\)-galactosidase II protein. Thus, according to the signal hypothesis, secreted proteins have a signal or secretary leader sequence which is cleaved from the complete polypeptide to produce a secreted "mature" form of the protein. In some cases, cleavage of a secreted protein is not entirely uniform, which results in two or more mature species of the protein. Further, it has long been known that the cleavage specificity of a secreted protein is ultimately determined by the primary structure of the complete protein, that is, it is inherent in the amino acid sequence of the polypeptide. Therefore, the present invention provides a nucleotide sequence encoding the mature \(-\)-galactosidase II polypeptide having the amino acid sequence encoded by the cDNA shown in FIG. 2[SEQ ID NO: 4] and provided in GenBank (Accession No. AF20390). By the "mature \(-\)-galactosidase II polypeptide having the amino acid sequence encoded by the cDNA clone shown in FIG. 2[SEQ ID NO: 4] is meant the mature form(s) of the \(-\)-galactosidase II protein produced by expression in a plant cell of the complete open reading frame encoded by the cDNA sequence of the clone shown in FIG. 2[SEQ ID NO: 4] and provided in GenBank (Accession No. AF20390).

[0056] The exemplary \(-\)-galactosidase II cDNA of the present invention (TBG4) has been expressed in E. coli strain XL1 blue MR (lacZ) (Stratagene, La Jolla, Calif.), as described hereinbelow (see Example).

[0057] Analysis of the deduced amino acid sequence of cDNA clones representing the other genes of the \(-\)-galactosidase of the invention also revealed open reading frames and, in some cases, suggested a high probability for secretion of the encoded proteins. All the full-length cDNA clones were predicted to have a signal sequence (FIG. 2). Using the two prediction programs SignalP and PSORT, TBG4 was predicted to be secreted by both programs. TBG1, 2 and 3 were predicted to have cleavable signal sequences by SignalP, but uncleavable signal sequences by PSORT. TBG7 was suggested to be targeted to the chloroplast by PSORT. Particular observations for each of the seven clones are as follows, based on the presence of a hydrophobic leader predicted by the programs PSORT V6. and SignalP V1.1:

TBG1: initiation codon at 306 [SEQ ID NO: 1], ORF=835 amino acids [SEQ ID NO: 8], signal sequence at 1-24;
TBG2: initiation codon not determined [SEQ ID NO: 2], ORF=888 amino acids [SEQ ID NO: 9], signal sequence at 1-25;
TBG3: initiation codon at 32 [SEQ ID NO: 3], ORF=838 amino acids [SEQ ID NO: 10], signal sequence at 1-22;
TBG5: initiation codon not determined [SEQ ID NO:5], ORF=251 amino acids [SEQ ID NO: 12], signal sequence not determined;
TBG6: initiation codon not determined [SEQ ID NO:6], ORF=248 amino acids [SEQ ID NO:13], signal sequence not determined;
TBG7: initiation codon at 104 [SEQ ID NO: 7], ORF=870 amino acids [SEQ ID NO:14], signal sequence at 1-35.

[0058] The deduced amino acid sequences of the seven clones was also subjected to analysis using the program DNAseq and the predictions for molecular mass, cellular targeting, pl and potential N-linked glycosylation sites are summarized in Table 1.

<table>
<thead>
<tr>
<th>CLONE</th>
<th>mRNA (kb)</th>
<th>kD</th>
<th>pl</th>
<th>N-LINK*</th>
<th>TARGET</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBG1</td>
<td>3.2</td>
<td>90.8</td>
<td>6.2</td>
<td>2</td>
<td>ER/OUT*</td>
</tr>
<tr>
<td>TBG2</td>
<td>3.0</td>
<td>97.0</td>
<td>6.2</td>
<td>6</td>
<td>PM*</td>
</tr>
<tr>
<td>TBG3</td>
<td>2.8</td>
<td>90.5</td>
<td>8.2</td>
<td>1</td>
<td>ER/OUT</td>
</tr>
<tr>
<td>TBG4</td>
<td>2.6</td>
<td>77.9</td>
<td>8.9</td>
<td>3</td>
<td>OUT</td>
</tr>
<tr>
<td>TBG5</td>
<td>-3</td>
<td>-3</td>
<td>-3</td>
<td>-3</td>
<td>-3</td>
</tr>
<tr>
<td>TBG6</td>
<td>-3</td>
<td>-3</td>
<td>-3</td>
<td>-3</td>
<td>-3</td>
</tr>
<tr>
<td>TBG7</td>
<td>3.0</td>
<td>93.3</td>
<td>8.0</td>
<td>6</td>
<td>CHLOR*</td>
</tr>
</tbody>
</table>

*Possible N-linked glycosylation sites
*Endoplasmic Reticulum; OUT = Secreted
*Tethered to Plasma Membrane
*Chloroplast

[0059] As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and
genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

[0060] By “isolated” nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include in vivo or in vitro RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

[0061] Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) with an initiation codon at position 64 of the nucleotide sequence shown in FIG. 2 [SEQ ID NO: 4]. Also included are DNA molecules comprising the coding sequence for the mature β-galactosidase II protein shown at positions 135-2532 of FIG. 2 [SEQ ID NO: 4].

[0062] In addition, isolated nucleic acid molecules of the invention include DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode the β-galactosidase II protein. Of course, the genetic code and species-specific codon preferences are well known in the art. Thus, it would be routine for one skilled in the art to generate the degenerate variants described above, for instance, to optimize codon expression for a particular host (e.g., change codons in the plant mRNA to those preferred by a bacterial host such as E. coli). Preferably, this nucleic acid molecule will encode the mature polypeptide encoded by the above-described deposited cDNA clone.

[0063] The invention further provides an isolated nucleic acid molecule having the nucleotide sequence shown in FIG. 2 [SEQ ID NO: 4] or a nucleic acid molecule having a sequence complementary to the above sequence. Such isolated molecules, particularly DNA molecules, are useful as probes for gene mapping, by in situ hybridization with chromosomes, and for detecting expression of the β-galactosidase II gene in plant tissue, for instance, by Northern blot analysis.

[0064] The present invention is further directed to nucleic acid molecules encoding portions of the nucleotide sequences described herein as well as to fragments of the isolated nucleic acid molecules described herein. In particular, the invention provides a polynucleotide having a nucleotide sequence representing the portion of FIG. 2 [SEQ ID NO: 4] which consists of positions 1-2538 of FIG. 2 [SEQ ID NO: 4].

[0065] In addition, the invention provides additional nucleic acid molecules having nucleotide sequences related to extensive portions of FIG. 2 [SEQ ID NO: 4] which have been determined from the following related cDNA clones: TBG1-3 and TBG5-7 as shown in FIG. 2, SEQ NO's 1-3 and 5-7.

[0066] In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above, for instance, the cDNA clone shown in FIG. 2 [SEQ ID NO: 4]. By “stringent hybridization conditions” is intended overnight incubation at 42° C. in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt’s solution, 10% dextran sulfate, and 20 µg/ml denatured, sheared salmon sperm DNA, followed by washing with the filters in 0.1x SSC at about 65° C.

[0067] As indicated, nucleic acid molecules of the present invention which encode β-galactosidase II polypeptide may include, but are not limited to those encoding the amino acid sequence of the mature polypeptide, by itself; and the coding sequence for the mature polypeptide and additional sequences, such as those encoding the about 1-23 amino acid leader sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences.

[0068] Also discovered is an enhancer/promoter associated with expression of the genes encoding β-galactosidase. The inventors have characterized the expression profile of TBG2 mRNA and have cloned a lambda genomic cDNA. TBG2 is expressed before the onset of fruit ripening and continues at uniform level throughout all the ripening stages. TBG2 has been found to be expressed in all fruit tissues and has also been found to be fruit specific. Experiments have shown TBG2 to be unaffected by ethylene. TBG2 is expressed in the ripening mutants rin, nor and Nrat the normal chronological time after anthesis. The promoter discovered would be useful to express any gene in the sense or antisense orientation, specifically in tomato fruit, in all tomato fruit tissues, starting before and continuing throughout the entire ripening process. The promoter could also be used to express any gene in the ripening mutants rin, nor and Nrat the need to gas the fruit with exogenous ethylene.

[0069] Variant and Mutant Polynucleotides

[0070] The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the β-galactosidase II protein. Variants may occur naturally, such as a natural allelic variant. By an “allelic variant” is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, 1985, Lewin, B., ed., John Wiley & Sons, New York). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

[0071] Such variants include those produced by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the β-galactosidase II protein or portions thereof. Also especially preferred in this regard are conservative substitutions.

[0072] Most highly preferred are nucleic acid molecules encoding the mature protein having the amino acid sequence shown in FIG. 2 as pZBG2-1-4 or the mature β-galactosidase II amino acid sequence encoded by the deposited cDNA clone.

[0073] Further embodiments include an isolated nucleic acid molecule comprising a polynucleotide having a nucle-
otide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to a poly-
nucleotide selected from the group consisting of: (a) a nucleotide sequence encoding the β-galactosidase II polypeptide having the complete amino acid sequence in FIG. 2[SEQ ID NO: 4] (b) a nucleotide sequence encoding the mature β-galactosidase II polypeptide shown in FIG. 2[SEQ ID NO: 4]. © a nucleotide sequence complementary to any of the nucleotide sequences in (a) or (b) above.

[0074] Vectors and Host Cells

[0075] The present invention also relates to vectors which include the isolated DNA molecules of the present inven-
tion, host cells which are genetically engineered with the recombinant vectors, and the production of β-galactosidase II polypeptides or fragments thereof by recombinant tech-
niques. The vector may be, for example, a phage, plasmid, viral or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complement-
ing host cells.

[0076] The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

[0077] The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL pro-
moter, the E. coli lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAAG or UAG) appropriately positioned at the end of the polypeptide to be translated.

[0078] As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampi-
cillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293 and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

agene; and ptrp9a, PKK223-3, PKK233-3, pDRS450, pRT15 available from Pharmacia. Among preferred eukaryotic vec-
tors are pWLINEO, pSV2CAT, pOG44, pXT1 and pBG available from Stratagene; and pSVK3, pBPV, PM3G and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

[0080] Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dex-
tran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other meth-
ods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986).

EXAMPLES

Example 1

RNA Extraction

[0081] Tomato (Lycopersicon esculentum Mill., cv. 'rut-
gers') plants were grown in a greenhouse using standard cultural practices. The ripening mutants, ripening inhibitor (rin), non-ripening (nor) and never ripe (Nr) (Tigeljaar et al. 1978. Hort. Science 13: 508-513), were all cultivated under 'rut-
gers' background. Flowers were tagged at anthesis and fruit were harvested according to the number of days post-
anthesis (dpa) or based on their surface color using ripeness stages as previously described (Mitcham et al. 1989. Plant Physiol. 89: 477-481), the complete disclosure of which is hereby fully incorporated herein by reference. For gene expression studies, a variety of leaf, flower, and stem tissues were harvested from greenhouse-grown plants and roots were harvested from seedlings grown in basal tissue culture medium for 4 weeks after seed germination.

[0082] Fruits were processed immediately after harvest in the greenhouse by chilling on ice, excising the various tissues and freezing them in liquid nitrogen. Tissue samples were ground using a mortar and pestle and stored at −80°C. RNA was extracted using the method described in Verwoerd et al. (1989. Nuc. Acids. Res. 17: 2362). Poly(A) RNA was purified from total RNA using oligo(dT) columns (Phar-
macia, Piscataway, N.J.). RNA was quantified by measuring A260 using a dual beam spectrophotometer.

Example 2

RTPCR

[0083] Degenerate primers were designed based on the highest shared deduced amino acid sequence identity we found between an apple (accession number P48980), asparagus (P45582) and carnation (P00662) β-galactosidase cDNA clones. The two primers used for the first reaction were BGS5E1 (WSNNGWSNHAYCATAYCC) and BGS3E (CRRAYTCTCNADNGGNGG). A second reaction was done on the products of the first reaction using BGS5H1 (AITHACARACNTAYGNTTTYGG) and BGS3E. The degeneracy code for the primer sequences is N=aa+1 c+c; H=at+1+c; B=tc+c+g; D=at+1+g; V=ac+c+g; R=ag+g; Y=cc+c+t; M=ac+c+g; S=cc+c+g and W=at+1. The 5' and 3' primers corresponded to amino acids 72-78 and 321-315 of the apple clone, respectively. Amplification was done using AmpliTaq DNA polymerase (Perkin Elmer, Norwalk, Conn.) and standard PCR conditions using the cDNA made for the first cDNA library described below as a template (Ausubel et al. 1987. In: Current Protocols in Molecular Biology. John Wiley and Sons, New York, N.Y.). PCR products were separated in an agarose gel and fragments of the expected size (approximately 750 bp) were purified, cloned into pCRScript (Stratagene, La Jolla, Calif.), and sequenced.

Example 3

cDNA Library Construction

[0084] Two cDNA libraries were constructed. The first comprised poly(A) RNA isolated from breaker, turning and
pink fruit pericarp from ‘Rutgers’ plants. The cDNA synthesis and library construction was done exactly according to the manufacturers instructions for the ZAP-cDNA Giga-pack II Gold Cloning Kit (Stratagene), the complete disclosure of which is fully incorporated herein by reference. First-strand cDNA synthesis was primed using a poly(dT) primer and inserts were directionally cloned into the Uni-Zap XR vector using EcoRI and XhoI restriction sites. The second library comprised poly(A) RNA isolated from all fruit tissues (except seeds) from immature green, mature green, breaker, turning, pink, red-ripe and over-ripe fruit of ‘Rutgers’ plants. The cDNA synthesis and library construction was done exactly according to the manufacturers instructions for the SuperScript Lambda System for cDNA synthesis and Cloning (Gibco BRL, Gaithersburg, Md.). First-strand cDNA synthesis was primed using a cDNA oligo(dT) primer and cDNA inserts were directionally cloned into the ZipLox cloning vector using Sall and NotI restriction sites. Both libraries were amplified and maintained using the host strains provided by the manufacturer, according to their instructions.

One of the clones (RT-PCR-1) was used to screen 106 plaques from the tomato fruit cDNA libraries at low stringency (hybridization at 45°C, no formamide and final wash with 0.2xSSC at 42°C). Thirty positive cDNA clones were identified and partially sequenced. Complete sequencing and characterization of the RT-PCR and cDNA clones revealed the possibility of seven unique β-galactosidase genes.

Example 4
D NA and RNA Gel Blot Analysis

Southern analysis was done using the 3’ UTR of each full length clone and the RT-PCR clones as probes against restriction enzyme digested genomic DNA. The genes corresponding to the clones appeared to be present as single copies (data not shown). The same probes were used to map 6 of the 7 genes using RFLPs of recombinant inbred lines and the loci names and map positions are shown in Table II (James Giowannone, Texas A&M University, personal communication).

Table II

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chromosome</th>
<th>Map Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBG1</td>
<td>12^*</td>
<td>Overlap of IL 12-2, IL 12-3</td>
</tr>
<tr>
<td>TBG2</td>
<td>9</td>
<td>IL 9-3</td>
</tr>
<tr>
<td>TBG3</td>
<td>3</td>
<td>IL 3-5</td>
</tr>
<tr>
<td>TBG4</td>
<td>12^*</td>
<td>Overlap of IL 12-2, IL 12-3</td>
</tr>
<tr>
<td>TBG5</td>
<td>11</td>
<td>IL 11-3</td>
</tr>
<tr>
<td>TBG6</td>
<td>2</td>
<td>Overlap of IL 2-4, IL 2-5</td>
</tr>
<tr>
<td>TBG7</td>
<td>no RFLP</td>
<td></td>
</tr>
</tbody>
</table>

^TBG1 and TBG4 are loosely linked

Total RNA (20 mg/lane) was separated in a formamide/Mops agarose gel, transferred to Hybond-N+ membrane (Amersham, Arlington Heights, Ill.), fixed by incubating for 2 h at 80°C, hybridized overnight in a hybridization incubator (Robbins Scientific, Sunnyvale, Calif.) using a buffer described by Church and Gilbert (1984. Proc. Natl. Acad. Sci. USA 81: 1991-1995) washed to a final stringency of 0.1xSSC with 0.2% SDS at 65°C, and autoradiographed essentially as described by Ausubel et al. (supra). An RNA ladder standard (GibcoBRL) was used to estimate the length of the RNAs. Probes were synthesized using a random priming kit with [32P]-dATP as the label (Boehringer Mannheim, Indianapolis, Ind.). Northern analysis was done using the 3’ UTR of each full length clone and the RT-PCR clones as templates for probe synthesis. As a loading control, RNA blots were stripped and re-probed at a reduced hybridization and washing stringency using a soybean 26S rDNA fragment (Turan et al. 1997. Plant Physiol. 113: 1329-1341). For all hybridizations, [32P(dATP]-labeled probe was diluted to 1-2x10^6 dpm/mL. DNA gel blot analysis was done essentially as described in Smith and Federoff (1995. Plant Cell 7: 735-745) except that 3 mg of genomic DNA was used for each digest. The complete disclosures of the above references are fully incorporated herein by reference.

Example 5
Sequence Analysis

Northern blot Analysis

Northern blot analysis was done to reveal which, if any, of the β-galactosidase genes had a fruit-specific expression pattern. With the exception of TBG2, transcripts of all clones were detected in non-fruit tissues (FIG. 4). Transcripts of TBG1, 4, 5 and 6 were detected in all the tissues tested. TBG3 transcript was detected at low levels in root and stem tissues, while TBG7 transcript was detected in flower and stem tissues.

Example 7
Temporal Expression Pattern in Fruit

The temporal expression pattern of the seven genes in fruit tissue was examined using RNA extracted from all fruit tissues except seeds. Transcripts for all seven genes were detected during some stage of fruit development (FIG. 5). TBG1 and 3 had similar expression patterns and their transcripts were detected throughout the breaker to over-ripe stages. TBG2 had a unique expression pattern and its transcript was detected at a constant level from 30 dpp to the over ripe stage. TBG4 expression pattern was similar to TBG1 and 3, but differed in that the transcript level was significantly higher at the turning stage. TBG5 had a similar expression pattern to TBG4 during the ripening stages of development, however TBG5 transcript was also detected throughout all the earlier stages of fruit development. TBG6 had an interesting expression pattern and its transcript was only detected at high levels in all pre-ripening stages tested.
TBG7 also had a unique expression pattern and its transcript was detected at very low levels throughout all the stages tested, and at moderate levels at 10 dpp and the over-ripe stage.

Example 8

Spatial Expression Pattern in Fruit

Northern blot analysis was also done to determine transcript accumulation in various fruit tissues. Since there were temporal differences in the expression patterns of the TBG genes both the mature green and turning fruit stages were used for RNA extractions (FIG. 6). Both TBG2 and TBG6 transcripts were detected in all mature green fruit tissues tested. TBG7 transcript was present in all fruit tissues tested except for locules. Both TBG1 and TBG4 transcripts were detected in RNA samples extracted from all turning stage fruit tissues. TBG4 transcript was notably more abundant in the peel. TBG3 and TBG5 expression patterns were unique and their transcripts were detected in all tissues except the outer pericarp and locular respectively.

Example 9

Expression in Normal Versus Mutant Fruit

In order to better understand the potential roles of the TBG products and transcriptional regulatory mechanisms, northern analysis was performed using fruit tissue from the ripening mutants rin, nor and N'. This analysis was important because it might give clues for preliminary determination of any potential ripening and/or softening role any of the TBGs might possess.

The results of mutant fruit Northern analysis suggested that the transcriptional regulation of TBG1, 2, 3, 5 and 7 was unaffected in mutant fruit tissue and that their transcripts were present in a normal chronological (dpp) pattern (FIG. 7). The abundance of TBG4 and 6 transcripts were however different in the mutant fruit. TBG4 transcript was not detected in fruit tissue of N' and was detected at much lower levels in rin and nor than wild type fruit tissues. Normally TBG6 transcripts are detectable at high levels throughout the early stages of fruit development but are not detectable after the mature green stage (40-42 dpp). TBG6 transcripts persisted even to 50 dpp in fruit of all three mutants.

Example 10

Transcriptional Regulation by Ethylene

The northern analysis done using mutant and wild type fruit suggested that TBG4 expression might be up-regulated by ethylene and that TBG6 expression might be down-regulated by ethylene. In order to evaluate this hypothesis mature green fruit were harvested and subjected to a continuous flow of 10 ppm ethylene mixed in air. Control and ethylene-treated fruit were used for RNA extractions at 1, 2, 12 and 24 hours. The results of this experiment confirmed the findings from the mutant fruit northern analysis. As expected, the presence and abundance of TBG1, 2, 3, 5 and 7 transcripts was essentially unaffected in mature green tissues subjected to exogenous ethylene treatment (FIG. 8). However, TBG4 transcript abundance was increased in mature green tissues in the presence of ethylene. From the data presented it was unclear whether TBG6 transcript abundance was reduced by exogenous ethylene treatment since its transcript level was normally reduced at this stage of fruit development.

Example 11

Enzyme Activity

In order to determine the role of the TBG encoded products we initiated experiments to express the cDNA encoded enzymes using heterologous expression systems. Several E. coli expression systems were tested, but the yield of product was very low due to toxicity (See the example). Therefore we used a yeast expression system which secretes a mature amino-terminal-FLAG fusion protein into the culture medium. The TBG4 cDNA was tested first and resulted in the production of approximately 1 mg TBG4 active protein per 50 mls culture. TBG4 was used first because the cDNA codes for the enzyme β-galactosidase II which was purified from tomato fruit and has been characterized in some detail (Carey et al., supra; Smith et al. 1998. Plant Physiol. 117: 417-423). Therefore we could compare the activity of the heterologous system-expressed protein to the native enzyme purified from tomato. The TBG4 protein was successfully affinity purified using an anti-FLAG affinity resin (FIG. 9).

The affinity-purified TBG4 enzyme was shown to have β (1→4)-D-galactosidase activity by virtue of its ability to hydrolyze the synthetic substrate 5-nitrophenyl-β-D-galactopyranoside (Smith et al., supra). The enzyme can cleave galactosyl residues from a variety of cell wall substrates and therefore has exo-galactanase activity (Table III). The remaining full-length cDNA clones are currently being tested for successful expression of active enzyme. Preliminary results have shown that TBG1 codes for an enzyme which also has both β-D-galactosidase and exo-galactanase activity (Table III).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate&lt;br&gt;</th>
<th>Boiled</th>
<th>Live</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBG4a</td>
<td>CSP</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>ASP</td>
<td>0</td>
<td>14.5</td>
</tr>
<tr>
<td>TBG1b</td>
<td>HCP</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>ASP</td>
<td>0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

*0.005 units enzyme/ρx
*b0.005 units enzyme/ρx
**2 mg substrate; 4 hr 37°C.

Example 12

pZBG2-1-4 Codes for a β-Galactosidase

The TBG4 ORF was cloned in-frame into the repressible/inducible bacterial expression vector pFLAG-CTC. The host strain XL1-Blue MR is a mutant strain containing no endogenous β-galactosidase activity nor...
a-complementation. Induction of gene transcription by IPTG caused the immediate cessation of E. coli growth at 30 to 37°C. However, induction at 20°C did allow for some limited E. coli growth. When clones containing the pZBG2-1-4 ORF were grown at 20°C and induced with IPTG, the cells slowly turned blue after 36 hrs growth in medium containing the β-galactosidase substrate X-GAL, (FIG. 10). If not induced with IPTG, no blue color was seen, even after extended growth in media containing X-GAL. As an additional negative control, clones consisting of XL1-Blue MR transformed with the FLAG vector alone never showed any β-galactosidase activity with or without IPTG-induction, even after 7-days growth (FIG. 10). As a positive control for maximal β-galactosidase (derived from E. coli β-galactosidase) activity the cloning vector pGEM was transformed into the host strain DH5α and the results are also shown in FIG. 10. FIG. 10 shows the detection of β-galactosidase activity from pZBG2-1-4 expression in E. coli. Cells were harvested and extracts were prepared every 12 hours and the A₅₄₀ measured. Cultures were grown with the addition of the chromogenic substrate X-GAL (open symbols) or X-GAL and the transcriptional inducer IPTG (closed symbols) in the medium. The vector used as a positive control for E. coli β-galactosidase activity was pGEM (black) and the vector used as a negative control and for expression was pFLAG-C1C either without (○), containing the pZBG2-1-4 ORF (●).

Example 13

Effects of TBG4 encoded β-galactosidase II on Plant Tissue Texture

[0098] To further demonstrate the function of TBG4 encoded β-galactosidase II the following experiments were carried out.

[0099] Fruit from tomato plants containing antisense constructs to suppress TBG4 mRNA were up to 40% firmer [compare means of parental line #1 with antisense line #2 in FIGS. 11A-11E (1-4)] than fruit from the parental line. Among the transformants the line with the firmer fruit also had the lowest overall levels of TBG4 mRNA (FIGS. 12A, B). This correlation suggests that a reduction in TBG4 mRNA is associated with increased fruit firmness. Firmer fruit might result in (1) less shipping damage (a) less loss due to damage and (b) ability to harvest at later stage resulting in better flavor at market (2) longer shelf life for both market and consumer. (3) better quality fruit for fresh slice market; fruit cut better at the pink/red stage when firmer.

[0100] To determine the function of TBG4 encoded β-galactosidase II, antisense constructs were made using the constitutively expressed 3SS CaMV promoter to express TBG4 antisense RNA (FIG. 13). Constructs were moved into tomato using Agrobacterium-mediated transformation. Four tomato cultivars have been transformed in order to evaluate the effect of TBG4 suppression on processing tomato (cv ‘UC82B’) fruit paste quality and three fresh pick cultivars. Of the fresh pick cultivars one is a soft fruit large cherry tomato (cv ‘Alisa Craig’), the second is a soft fruit old breeding line (cv ‘ruters’g’) and the third is a recently developed somewhat firm cultivar ‘New Rutgers’. Among the lines where TBG4 mRNA is suppressed we expect to observe an increase in firmness and paste viscosity.

[0101] Although this project is nearly finished the complete biochemical and molecular analysis is not finished. The preliminary results on the analysis of the ‘New Rutgers’ cultivar is presented in FIGS. 11A-E (1-4) and 12A, B. In this example a fresh pick cultivar called ‘New Rutgers’ was used. Plants of the purchased seed were grown and allowed to self and the resulting seed was used as the parental control (line 1). Seven independent transformed plants (lines 2-8) containing TBG4 antisense constructs were grown and allowed to self. Transformation (T-DNA insertion) was confirmed by southern analysis (data not shown). From each transformed line, five plants were grown along with 10 parental line plants. Fruit were tagged at the breaker stage (1° onset of color change) and were harvested at breaker plus 7 days. Data were taken using 15-20 fruit from each line. Each type of texture measurement was done twice for each fruit and fruit were subjected to 4 types of texture measurements using a Stable Micro System’s TA-XT2i texture analyzer. The 4 measurements were: 1, 2-inch flat plate compression to 3 mm (FIG. 1A), 2, 4 mm spherical indenter compression to 3 mm (FIG. 1B), 3, 4 mm cylindrical indenter compression to 3 mm (FIG. 1C) and 4, 4 mm cylindrical indenture puncture to 10 mm (FIG. 1D). The summary of this data is shown in FIGS. 1E (1-4). In FIGS. 11A-E (1-4) line 1 was the parental line and lines 2-8 each represent an independent transformant containing one T-DNA copy of the TBG4 antisense construct. Statistical analysis (Duncans and Scheffe) of the data revealed that fruit from the transformed lines 3, 7 and 8 were not significantly different from the parental line but that transformed lines 2, 4, 5 and 6 were significantly firmer than the parental fruit. Most noteworthy is that fruit from transformed line 2 had fruit with a mean firmness that was 40% firmer than that of the parental line (FIGS. 11A-D).

Example 14

Northern Blot Analysis

[0102] We are currently investigating any changes in the biochemical composition of fruit where TBG4 mRNA levels have been suppressed. These experiments are designed to show a link between increased fruit firmness and TBG4 mRNA suppression, TBG4 encoded enzyme activity suppression, possible cell wall modification (e.g. increased galactosyl residue content) and a decrease in free galactose levels during fruit ripening.

[0103] These experiments are not complete, however some preliminary Northern blot experiments were done and the data is shown in FIGS. 12A, B. There is no parental or azygous control fruit RNA shown in FIGS. 12A, B because these plants were the last to grow and RNA extractions are just being done now. As a comparison of normal fruit TBG4 mRNA levels refer to FIG. 5 above. The data from FIG. 5 showed that TBG4 mRNA levels are low at the mature green stage, peak at the eating stage and are reduced at the red stage. All the lines except for 2 and 3 expressed antisense TBG4 mRNA (FIGS. 12A, B). The antisense transcripts appear as two bands, smaller in length than the endogenous mRNA. The two bands probably resulted from 1, the expected transcriptional stop signal provided by the NOS- terminator and 2, a cryptic transcriptional stop signal in the antisense TBG4 cDNA. The most notable result was in line 2 where no TBG4 mRNA was detected at the turning stage.
Line 2 also had the firmer red fruit (see FIGS. II A-D). The absence of detectable TBG4 mRNA probably was the result of cosuppression of both the endogenous and antisense mRNAs. When compared to earlier blots (e.g. FIG. 4), all of the lines appeared to have an overall reduced level of TBG4 mRNA, but it is impossible to assign numbers to this statement without the parental and azygous control RNA on the same Northern blot.

The specification discloses that β-galactosidase II polypeptide is involved in the degradation of cell wall pectin during fruit ripening. In the present invention, the role of β-galactosidases in tomato during fruit ripening and softening and the description of the cloning of a β-galactosidase cDNA clone that codes for a β(1→4) galactan degrading enzyme, which is expressed in ripening tomato fruit tissues, has been shown.

The present work indicates that pZBG2-1-4 is a cDNA derived from the transcript of the TBG4 gene which codes for β-galactosidase II for the following reasons:

First, the deduced amino acid sequence of the highly conserved amino-terminal portion of the expected mature pZBG2-1-4 translation product matches almost exactly (28 of 30 amino acids) with the amino-terminal sequence of β-galactosidase II as purified by Carey et al. (supra) and designated TOMAA. Importantly, the two amino acids (KY) in the β-galactosidase II sequence (TOMAA), that do not match the pZBG2-1-4 deduced amino acid sequence of the present invention are believed to be incorrect since all plant β-galactosidase sequences in the database and four additional β-galactosidase-related cDNAs that were identified from tomato all match the deduced amino acid sequence of pZBG2-1-4 at these same two amino acid (ST) positions (FIG. 3).

Second, the transcript detected by pZBG2-1-4 is present in normal ripening fruit at the same time that β-galactosidase II activity was detected (FIG. 5; Carey et al., supra). Moreover, little or no transcript was detected in fruit at 45 and 50 dpa from the mutants nor, rin and Nr (FIG. 7). This observation also coincides with the data presented by Carey et al. (supra) that β-galactosidase II activity remained at levels equal to mature green fruit and did not rise in fruit 45-65 dpa from nor or rin plants. Interestingly, Carrington and Pressey (supra) have reported that β-galactosidase II activity was only detected in ‘Rutgers’ fruit after the turning stage of ripeness. The Northern data in the present invention indicates that maximum β-galactosidase II activity occurs only after the turning stage, assuming mRNA levels predict extractable enzyme activity (FIG. 5).

Third, the apparent molecular weight of 77.9 kD and pl of 8.9 for the mature protein predicted from the pZBG2-1-4 sequence is similar to that determined for β-galactosidase II, Pressey (supra), estimated a molecular weight of 62 kD by gel-filtration column chromatography and a pl of 7.8 by isoelectric focusing, while Carey et al. (supra) estimated a molecular weight of 75 kD by SDS-PAGE and a pl of 9.8 by isoelectric focusing.

Fourth, enzyme produced from pZBG2-1-4 ORF using a heterologous yeast expression system has both β-galactosidase activity and exogalctinase activity.

---

**SEQUENCE LISTING**

- **SEQ ID NO: 1**
- **LENGTH: 3224**
- **ORGANISM: Lycopersicon esculentum**

**SEQUENCE: 1**

```
ttttttttttt gtttcttttg tcagctctag agcctagag aagggaaaaa aqagntstgg 60
actaagagaa taacctaaca aagagagaa aaaaaaata aagaaat sttcgacaa 120
caaaaaagtc ttttttctct tctactatct atctatata aatgtgcagga 180
atatgatgat gttgacact tcgttaca aagagagaa taagcttgagg ggtgagggag 240
gaagttagtt cattattgca ttgctttctg tgaagctgtcg aagacatct tgaagctgtcg 300
catataggg ttcttttgag gtgaatgctgc tgaagctgtcg atgctttttt tgaagctgtcg 360
gttgatagtc tctgtgttct tgaagctgtcg tgaagctgtcg caagagaa 420
ttcatacttc ggatcactct cactcactc gagcagccgg tcgaattctgc gcctattgttc 480
ttcagagctg aaaaaggaag ggagttgtag ttaagagagac ttgatgggc tggatgggc 540
agagcagcgc aagagagcg ttaattaag aagaggtgct gccagctcttc gcagcgcgc 600
agatgctgca aagagcagca ctataagctgc ccagctcttc gcagcgcgc 660
agatgctgca aagagcagca ctataagctgc ccagctcttc gcagcgcgc 720
```
caasaataga gcctcctcaag gctgcaatgc aaaaatctca tcaataagtt gttgatagta 780
tgaagcaaga aaagctcat gaaactcagg gttgcaaat tattctactt cagataagaaa 840
atgaatagtg acaactgagc tgggaaactggt gtaaacctgg taaagtttac tcgaatgagg 900
cagccaaat ggtcttggtg ctggtgcactg gttgcatctg gatctagtgc aagaaagattg 960
atgcttctga tctctataatt aataactgca aatgcttctta ctttgtcaac ttcacaccaa 1020
ataacgcat ctaaccccaag atgtgcaact aagcctcgac tgcctggtttt accgaaatg 1080
gagtttcagt tctctaccgt cctgcagagc atatggcatt gtcgctcga acagttatatac 1140
aaccctgggga cctcctcnac caaactataa tgtatcttgt eggacacaaac ttggggaaggt 1200
cctctggtgg cccctatttt gctatctyttt attgattatga ttgacaccctat agatgaatttgg 1260
ggtcattcgc gcgccctaa tgggctcctc tgggaggct acataagagca ataacaacct 1320
gtgacgccgc ttatgtatct gtagacccaa ctgtgcacat cttgagaaac tataaagagg 1380
cactactttc cagctcgacg tctgggcccct cgctgcctct ttcctgcaact tcaacaccagc 1440
actctttttgc taaatgggaca tttggggaaca ttctattataa ttcacaccocc tgtgtotota 1500
gcacttcctc cgaactcagc aaaaactctct atataactgtg aaggggtttgg gctaacacg 1560
cctcagatgaa gtagcctoaa gtcacgtgac gttttctatg gatgctattg aatgaaacgc 1620
cacccgctgg aagagacgc aacgtacttc cttggtggatt atggctattt gtagtaacg 1680
cagctcctgg aagagacgc aacgtacttc cttggtggatt atggctattt gtagtaacg
caagagattt atcggacttt tgtgtggtata tgtgagctct gcagagctttt acacaaagc
1740
gatttttcga tagtccttcta tgttgtcttct tgttgtcttc ctgatccttg 1800
atgtatctct gcagagctttt atgtgcttca aatgctacgtt aatgttagaga aacccaaacc
1860
taatctctgc ccaaccttata aatctgagacctctgtgctg cagatcctct cgtgatc 1920
tgctgttgct tcttcgcacgt cttggtctccct atttctgacgc aaggttagttgt ctgtcctct 1980
gacccgttto atctataagga attaagagga cttctactgg cgaattgtgt 2040
ttctcactgg tggccctaaa ggagagcgcct gcgtcttactc ttcactcagc ggtgagcc 2100
cctgtggtagt gttgtgacgct tctttgtgct ctcagagtcc gcaactcag tggcttaaga 2160
cctagctccac atgctgctagct cgcagacgtc gagatgtatat ggtatggtgctt gcgggocct 2220
aaggttcagtt atggtatgac gttgagcgcct cttggtgacgct cttgggctca cattatatct 2280
cctggaatct tgtgtgctt atctataacct gctgtgtgctt tyggatctaa aggctaaact 2340
actggtgtgc ccgggcaacaa agatgtcatc agctaccccg ctggctggctg tattctactg 2400
gaaatgtggt atgtattcac ggaagatggg gggagactcct ttattgtcaacttctttatt 2460
aaagagataa aggggggttg tgtgctgtaa tatcttgctg gcaaccaccag ttattgatat 2520
gggagagctt atgtatggatt aagttgctca aacactctcag ctgctctgctct gcttttaccc 2580
gtgctctctgg tcaagattg tcttactata aatttgcaagc cttggtgacgct cttgagggag 2640
cttgcggtgga ctctctccac gcgagctgctc atgcctccag cctctctatt cgttttaccc 2700
agatgtgtgg tggagagcat tcttgctagc accusagaaac ttgggggttg 2760
atctccagcg aacgtctttca aagaaactctt ctggtgtacaatttgctt gtagttgttt 2820
gagtattcac gttggaaaagtc ttccttccaa ctttcttctac atggagctact 2880
agacatctca taaccacacac tataaatatttt tgggttgcgt gggggtgcga ttttctctat 2940
aacacacca cctttcttggat caaactcgaca tctgtatgta aagatgattg caaagactc 3000
-continued

tgtacctgta ayytctcgc taaatcac ccatactgt aacatctggt ggtagctgt 3060
caaaatttt ttgttctagg gtagtagaa cttcataaa actgtgttct catgattgct 3120
tacattata caccataa ctagtagc aagccaat attgagtttt cactgaaa 3180
aattgcattt actatgtgtgc aattcactaa aacaaaaa aaaaaa 3224

<210> SEQ ID NO 2
<211> LENGTH: 2984
<212> TYPE: DNA
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 1

ggagcagag aaaaaactgt aatcttctgt taataacaa ggtgtttact atcaactttg 60
tgtcgctgc cggcagatgt tcaagcagtt tcaagctcct ctacatattc ccagcttaac 120
tcagctggtg caacggcctct atcgctagtt agcagagtttt gttgtcagat gtcctcgaag 180
cgtgatagtc gccaccagct attcttcttg cgcgaagagttt gttgctaat ggtctcttctct 240
tctatctta tggagatgct ccgcgctcgc caagggcagg gcgataatt ccaaggaagat 300
atgatttgg ccagctgtgtg gatcctcaag gatccttatg gatcctcaag 360
tagctctta tggatgtgca gaaaggtact tcggggatt coccatattg cttcgtgata 420
taatcaagatt caagattgga aacatattac agaatcactaa acaatcactaa gggagatgtg 480
taaaaaaagt agttcatctt atgatatctgt agttcatctt ctgctgtoaca ggtgttcctta 540
tacccttctg cgccttctag aatgtatatt gaaatattgaa aagctttctttcg ttcctctccag 600
ggaagttata atgaagttctt ggtcctgtgaa tggctgttctg gttgctcttctctg 660
gggtcatgtg cagcacaact gattgctccag aactatctag atgatctctgt atcgcatact 720
attgtagagt gttccagcgg aatcagcaca aagagacagaa aatttctgtt gagaatggga 780
atggtatgtg tctggtattg ggtggagact tctacatctag accttccgag atattargcat 840
tgtaagtctg tgcagttctgt ccaaggtggg gcagattctc caagactatct tatattttttg 900
gttggcagaata ttgctgctgc agttctgttg gccacactta aatcactcag tattgattatg 960
atgctctct ggtgaaatag gaaactattc tcaagctctcg ctaccggctc tggggaggattcg 1020
tggtgagaatag tataaagttg tgggacccag ctctgtagtgc tgggctattac cctaggtata 1080
ttacactgag accacacacg gaggccactg ttcattctgtg aacatcaca aaccattggcc 1140
aataatatgt gtttaacgag gctagtattc cagcaatatttg gattgactctg 1200
aactacacg aatgttatatt tattcagtttt acattctggt gcattgttcat 1260
tctgctcactgt tgcagcattaa cctgctcactgt gtaatctgatg 1320
ccaaaacttg ggttagaatg ctgcctttctg tgggaatatt cccctctttc tgaagatctg 1380
cactaagaa aacagcttagc agatatttctc acacagttact gacactgga aagcactctg 1440
gttgtggg gcagcagatg ttcactcttctc aagggactcag gcactgtcct atgtgcagc 1500
aagccacagc agattctcttg ttttctttcag cccatattct gattagttcct gttgatattct 1560
catttctgg gagaatattg ccatcctctg ctaattcttg tgggctctg 1620
ttcgcttttt gcagttttgc gtagtctgaa aaggcactgg actctctggctg 1680
tcactcctat cagcagttgt aagggcactgt gctttcttttt tgggctctg 1740
gattgcggaa ttactgtgtgc tttctttttt ggggtttttaa gggtctgataa 1800
-continued

```
aggttatcag atgcaaaagc ggggatatca atctcaaacg atcttatggg acctaccagg 1860
tgggtgatat ggggaatttc ttggaagatg atagttgcacga tggattcgtta agtgcggat 1920
ggactgtgtt cttctaactgtg aacaactcgtg cagcttttttc gggttaacag aacaagtgg 1980
atgctcaac gaggatggat cagttgcttc ggtatattga tagcagtgagga aagagtcaag 2040
cagtggataa ttgcaacacc atatggaagt atgtgccttt ggttggacac aataatggat 2100
gtggagaaac tattttgatt tctagtctgat acacacgtgacus aatagcattg aacaactgtg 2160
gagaggtac gccgctctgg tacattcatac cttgacatcg gctaaaagaca ttaatattg 2220
tatcgattat cttggagaaa cacgatataaa ctctgattga tattttcatt ttcacgcgtt 2280
tcgttcgac cccttgggtg aagttgtctc tggatgtataa accacccagaa aagctctgcg 2340
tctgctgga gttgacagca aagttgtctc tggatgtataa accacccagaa aagctctgcg 2400
agtgctgac gacacatatac atctctctt aatgaaatgc aagttgtgcg aagctctgcg 2460
gcagttctgca aacggctctca cagaaagaaat gcagttctgca aacggctctca cagaaagaaat 2520
tcgggtctgtg tttgaagaa acagttgccc ttcctgtcgg tttttagaagc 2580
atcgttcg aacggctctca cagaaagaaat gcagttctgca aacggctctca cagaaagaaat 2640
acccagcgc ccctgcttcgac cttgagagg actcctctga tcagagatca ctgtagacac 2700
aacactggcc ctcgcttcgac cttgagagg actcctctga tcagagatca ctgtagacac 2760
atgcttcgac ctcgcttcgac cttgagagg actcctctga tcagagatca ctgtagacac 2820
ggttactgtc tttggataa cgagcgcagc gatgttcggtg ataattttgt tttttagaagc 2880
tgtagtctggt ttttggataa cgagcgcagc gatgttcggtg ataattttgt tttttagaagc 2940
tttttagaagc tttttagaagc tttttagaagc tttttagaagc tttttagaagc tttttagaagc 2990
```

<210> SEQ ID NO: 3
<211> LENGTH: 3069
<212> TYPE: DNA
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 3

```
aaggttatcag attttttctg cttttgaaaa aaggggaaaa aatassagtt aaaggggggg 60
gaaaagtgtt tttaaattgc ctaaaaaagc aacactctga tagaaagagcta gtaatatta 120
cacggtgtct gacctttctac tattgggtcg cttattttgg tglgggtttgg gttttgtggg 180
	ttttatgc cttctttctatg caatagttcg atattttgata ctgatggaga atggacaag 240

ggacacagtt attttttctg cttctttctatg caatagttcg atattttgata ctgatggaga 300

tttatattgc aaggggataag cggagttggt ggtatggtat cagacttttg ttcctttggaa 360
tggggcagc ccccaacagc ggaatatta ttttgaaggg agatattgatt tagtgaaggttt 420

ttatgaactcg gttgcacacc cagacacgata ttgcttctgg atagttggag cctattctttg 480
ttgctactgtt aatattttgt cttcttaggg gttttaggg gttttaggg gttttaggg 540

cgggctaagc gctacacagtt ctaagagctcg aactgcaaaaa tttactgtcga aatgggtcga 600
tattgtgaa ccgaaaagtt ctttttattt cttgggtaaat ggtaattttgctatataggat 660
tgctagtcgca tggggaaaaa cctgtgggtct gctttttgct gttttaggg gttttaggg 720

tcgggtctgtg ttttggataa cgagcgcagc gatgttcggtg ataattttgt tttttagaagc 780

ggcagagctg ccctgcttcgac cttttagaagc tttttagaagc tttttagaagc tttttagaagc 840
```
-continued

tccaaacaag gtttataaac caaagtatacg gcgtgagcc tggactgcat ggtttaactgg 900
ttttgaaat ccaagcctc aocgtcttgg cgagactgc aacatcttctc tggaaacatt 960
ttacagaaag gcaggttctct tttctcaatct caacagcttat cttctgggaa caaacccttgg 1020
cacgtcttgc gttggctacoat tttactgtaag gagctgacac actagtctaga 1080
atatgctaattctgcacca caaataggtgc tcagctgata caagagctataa 1140
gtggtaata ccaagccttgc ttcatctgaga ttcagtttctgc ccagacattctgg 1200

gagagccggc tgggtttaggt cgaagcttgg ctcttgctct gcatttccttg ccacttcaga 1260
ccaaacacttc tttctctcaac tgctattccg acacaaacat caaacttccag catggtctgc 1320
aatcagcactt ctcccagcagtc gcacacacac ctgtaattaac acacagccag tcagtgctca 1380

gacgtcttcag atgaagattgc ctccagctag cagaggtgtcc occttgccag ctacatccga 1440
agagacactca ttttctagag acaagaggttt tccagttgtct gggctatttg aacagataaa 1500

tccacacagc gagcgtcctgc atttattttgg attccaccaag gatgtcacaag tggattcag 1560

gaaaaagtttt ggagagggcg gaaatgggcc tgggtttcag acatgctcag cttgggctagc 1620
aatgcgtttt ttttgagatag gcctattgca aagacctgca ttggaagattg tgaanacac 1680

gaanactact ttcgctaaag ctgtagaactg gggacaggtt gtttcacagag tttttctact 1740
gagcttcgttc gagacgctttg caaatgctgg ccaccattctt gacagctggc atgcccgtgttg 1800
tctggggcata gtccctcaacct ctggtcctgc caagggamaa aagagatattt acatggcaag 1860
atgctcttcag aagctGGtccg ttttaaagagac accttcagaag cttccctcagc ctcgcttgag 1920
cgtctcagttt ggaggtctttt atggccctca aagacagocag agoatggygta 1980

cagagcagact tttttgctcc cagctggaaa tggcctttgct gcattacagtg tgaatcact 2040

gggcaacagc caaacaatga caaagctggg atatgctttg cgtctttttggct cttgctatca 2100
agacatctgg aacgcttggc ctcgtaactg tcaggtctgg tttaggtaa gccaatatac 2160
aatctagaat ggcagggcgtt cacaagatgt gatacatggtt ccacgctttt cgcgttatac 2220

tactgtgaatt gtctgtagat ctttggagga atgggagattt gacgctctag gaatcttttt 2280

ggtttaggtc gaagttcaga gctgtttggtc agataaattg gacgctccac caacagtttgct 2340

gaatgtgcaat atgcaagtct cttgtaaatg tcgcacacat cttgagacatct cttcaact 2400

cgtctcagttt gttgcttctca gaggatttcac attcaaatcc tccagctttt ggccacacaag 2460

tagagggcttgg gcaacccgcct cggagtttgtg cttggcggttc atggtgcttt 2520

tgggaaggttt tagcggtggc aacactctgtt ctcaagcttgg gcacacacag 2580

ggttctgcct gtctggcacaat cttggctttgg actggtagattt gacgctaatgtct 2640

aacagtgtgaa gaccacaaa aagttgtttt ctagtttaggt tttgaaacata ctgcaaaatct 2700

ggttctgcct gtctggcacaat cttggctttgg actggtagattt gacgctaatgtct 2760

tgaatgctga cggcccaagc tttgctatcat atatgttggcc tagcttcagc tttgcttagg 2820

tttgcaaatg ccaacagata gttatgcttc ctcggctaatgc aagctgactgc atgggtgctat 2880

tgggctttag tagttacagct tcagctcagtt gtttttctgc caagctgtgc cttaggatcct 2940

catatatat ccaacctagat gctccatagg actgcttccc gaggcttcag tcagttttgtag 3000

tgggctttag tagttacagct tcagctcagtt gtttttctgc caagctgtgc cttaggatcct 3060

aaaaaaaas

3069
<210> SEQ ID NO 4
<211> LENGTH: 2554
<212> TYPE: DNA
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 4

```
adadadaagtt tctctttttt ttcctaaaata aaaaaaataa cttttttttttt gggtggaat 60
aadadgaatt ggctaatgt gttgtgtta ttgttttttt ttttttttttgc 120
agacytgaag ctagtcttcc tattgatgac agagcatataa tctaaatgg gaaagaaaaa 180
attcttttt tctggccaat tcttatctaa agagaaccc caaacaatttc gctgtcctt 240
atacaaaagg caaagaggtt aggcatggtt gttattgtaa ctttattttt cttgtaagaa 300
catcagcct tctcctgaaac atataatattt gaggagagat atgtccttgt tagattctc 360
aadagttcag aagagcagg actttatgtc aattnaagtta tggccttaa ctgtcgtgct 420
gatgaactc tgtgggtggt tgtggttggag cacctgttgg ctaaatatatg tgtccttgtaa 480
acacacactca agcctttaaa ggtggcgttg cacgatatgtg ttgagagataa agatcataat 540
atgaactcag aaaaaattgt tgtaactctaa ggcggacacaa taattaacaac cacactaca 600
aatgagtctag gactagtaaat atgggaattt ggtgctctctg ttaaagctta taacaatagg 660
gacagctttaa tgtctgtgat tttgaaactt ttgtctcctc ggtcattgta taagcgagatg 720
gacgcttctg atctctgtat tgtactatgtt aaggtctttt ttcgctcagag tctggttcct 780
attactttttt cactacacag cagatagtaa gcaagttgaa cttgctgtga ttgagagataa 840
gagcgcaca agagctcaca gactgagcctt tttgaggttt caggttggct 900
cagagactataa atgactacta aatgagatag gabgacacaa tttgagcgg 960
acactaatcc ggtctgtcct ttgactatggc aagctcttctc cagatgtactcagatg 1020
ggtgtgatca atgacacaa gttctgacgt tattgacagt taacatagctg atcaaatgtct 1080
tctgacaccc gttctgttctt ctatagatgt gcctctggtta gctcttggtat gactgtgatac 1140
gtccttttttta atccgtgttat atggctacgcc ttgctgtgcct cttatcatctctcct 1200
agatcattcttaaagttgaactctgcaagctcgtcagcaagttgctaaagctttcaagcttct 1260
agatcttttttt ttaagctttttaa gtaatggcact ttttcttctgct tggctgtgtgcct 1320
tggagtcataa agatgctcgcct tgcagttgcttg ggtctttctc cggagcgact caagaagtat 1380
acgctctactg tgcagctccg ccctatatgcg cagaatggg gagaatggcagcttctatgctt 1440
gttttacgttat cttctctgtgg cttattgtctgtt atgctgtattt caggttttattt 1500
aaaccttt gcctaatcag ctcagcgtgta aatgctttga ttttcttcttct cgtttttcct 1560
ttgacttttttt ctaagacgctt gagatgtcctt attttctttcgtttatcagctcttttt 1620
aactattcaca ctagatgcaag ctgtagagct gttttctttt ttttgatgtgta gttttttct 1680
aggttttttggg gttttccttc ggtctgttattt tatttttttt cttactagatg gclctgttct 1740
cagcgtttctc tattttttctg cagcgtttattt tattttttttttt ctttctttcctcct 1800
ttttttttt tggctgctcat cttcttttttt ttgctgtttctct cggagcgtttttt 1860
tttttttttt cggagcgtttttt gagaatggg gagaatggg gaaacatct 1920
agagaacta ttatgatagt cttttttttttt ctttctttcctcct cttttttttttt 1980
gagatagtta tcttttttcccc gttttttttttt gttttttttttt tttttttttttttt 2040
```
<table>
<thead>
<tr>
<th>-continued</th>
</tr>
</thead>
<tbody>
<tr>
<td>gcacaaggcg aagtacgacaa atgcagttat gctggacgct tcacggaga gaaagtgcgg</td>
</tr>
<tr>
<td>actacaagcg gacaacctct atcgtagatg taccattttg caacagctgt gctggaacca</td>
</tr>
<tr>
<td>agtgggaact tggtagtttg attcggagaa tggggagcta atccacacag aatattttca</td>
</tr>
<tr>
<td>gtaagcagat caggtatsag aacctgaaa gtaaactttt ttcgaatatt atgggtctgg</td>
</tr>
<tr>
<td>aactggagcc gaaatacaca taceacgaag tcacaatgg ggctacagtt tgcaaaattgc</td>
</tr>
<tr>
<td>aagtgaatat aaccattasa gataagaaaa ttcgtgaatt aaaggttat ataaattttc</td>
</tr>
<tr>
<td>agagaatttt cttaattttt tgeaaaaacct tggttttaaa aytttataca gaattttttctg</td>
</tr>
<tr>
<td>ttatttgat tataagattg aagagagttg tacaatctcc atatatttatt agaatacaaa</td>
</tr>
<tr>
<td>taattttctct gtaaaaaaaa aaaaaaaaaa aaaa</td>
</tr>
</tbody>
</table>

**<210> SEQ ID NO 5**
**<211> LENGTH: 785**
**<212> TYPE: DNA**
**<213> ORGANISM: Lycopersicon esculentum**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;400&gt; SEQUENCE: 5</td>
<td></td>
</tr>
<tr>
<td>atcagacact acgtttttctgc ggacttctcat gaaccttttc acgatcgtta tgaatigas</td>
<td>60</td>
</tr>
<tr>
<td>ggaaggaag atttatttta tttttagaag tgggtgagga gacgtggcct attggtcttt</td>
<td>120</td>
</tr>
<tr>
<td>ataaagttg gcggctattg tttgtcagaa tgacattag tgggttttcc tctttgggct</td>
<td>180</td>
</tr>
<tr>
<td>catttccttc cttttttttttt aacatgagc gtcagagtgg acaaattgaa</td>
<td>240</td>
</tr>
<tr>
<td>cgtttttttttttt aacatgagc gtcagagtgg acaaattgaa</td>
<td>300</td>
</tr>
<tr>
<td>gggcggcttt tcttctacta gataaaaaat gatgttgctaa atggtgtata ttgctccctg</td>
<td>360</td>
</tr>
<tr>
<td>tatgttttac gttcgaacac ttaatgtaac ttggtctatg caatgcttctac gtcttttaaat</td>
<td>420</td>
</tr>
<tr>
<td>aagggagtgc aactggcttt tggctcagca acaagactccc ctcttctcct atatgaaaact</td>
<td>480</td>
</tr>
<tr>
<td>ttgaactatg ttaatagttg ccaatttctc caaatctctg ataaacactt cgaatgattg</td>
<td>540</td>
</tr>
<tr>
<td>actagaaact gcagcggttg ctgttctgct tttgtgtgtc ctgttccctca cagaaacttg</td>
<td>600</td>
</tr>
<tr>
<td>gaagactctg ctttctttttct ttcacggag gggaacctt gccgaatctt ccaaaact</td>
<td>660</td>
</tr>
<tr>
<td>tatactagcc acggggagtt aacttttttt ggaaccctgt ttggacggct tattcagact</td>
<td>720</td>
</tr>
<tr>
<td>agaatgact atagtgccct ttctacgagaa taagg</td>
<td>755</td>
</tr>
</tbody>
</table>

**<210> SEQ ID NO 6**
**<211> LENGTH: 749**
**<212> TYPE: DNA**
**<213> ORGANISM: Lycopersicon esculentum**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;400&gt; SEQUENCE: 6</td>
<td></td>
</tr>
<tr>
<td>atcagagact atgtttttttt gaaatgttctc gcgttttctc ccgttcaattaa caaatttgaa</td>
<td>60</td>
</tr>
<tr>
<td>ggaagaatag acgttggtag ttgtttaaa aagattcaata aacgaggcct gtattgcctat</td>
<td>120</td>
</tr>
<tr>
<td>ccctgctct agcctttaag tttgtgcagag tgaatattttt gagggtttttt cgatagtggct</td>
<td>180</td>
</tr>
<tr>
<td>aagatagtgc cttggcattg cttcaacgct caatgacta cccaaatgacaa</td>
<td>240</td>
</tr>
<tr>
<td>gggtgtatgc aagacagtttt taactttttta ctattttttg tctttgctgc</td>
<td>300</td>
</tr>
<tr>
<td>gtcactcat atctctcaag atgagggcct tcaaccaag gttaatttgc</td>
<td>360</td>
</tr>
<tr>
<td>caagggaga caaatgatta acaaggtctgt caaatgtgcc gttggttattt ggaacagcgg</td>
<td>420</td>
</tr>
<tr>
<td>gtcagatgg ggtagtggcag gaaagagcat ctcacagact cttgtgcaaa ccatgcaat</td>
<td>480</td>
</tr>
</tbody>
</table>
-continued

gttctact gtgataattt cttccccacc aacccataaa aaacgccat tggagctgat  
540

gttctgatg gtttgctggt gaaattggc ggtcccttc acaagacacc aagtatgctg  
600
tggctgttgg ctgtgccaca aatctatcgt aagggaagggt ctcttgtttta ccatctatg  
660
taacgctgg ggcaagaact tggagccact gcgggttgac catctatcct taccagctat  
720
gattatcag cacccctctga cggatattg  
749

<210> SEQ ID NO 7
<211> LENGTH: 2972
<212> TYPE: DNA
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 7

gcaacccctcc cgtgaatata cacgggttaaa ccggcagatgg caaactctgt ccgaatctga  
60
atagtattt aagcagcctta gctagctaac ttttgtcctc gcacggaaca caagtatgtg  
120
 tgttgcctct aatcatcaag tgtttttctc tcgtgctgat ggtgatgta gagctgaact  
180
gcggctcgag ttcgcaagcc agtagttacat gctatctgtca ctgataatgt  
240
ggcttattcg gcagctgcgt tatttataa ccggcaagag aagctctcgc ttcgctcctc  
300
cattccat ctacgcaagtgt tccctgctat gttggtctggt ctgtgtctgt tgggaaagga  
360
agagagctgt gcgtgtaattt aacagctgtg tttciggaac gcctgagaac cttctcgcgg  
420
caatttacat tttgaggggt gttttgcot ctgtcaasat ttagacctga ttcgagccgc  
480
tgagagttgt atgctttgct ggtgacgac attgtgtcgt gcgaatgga aaccttggtgg  
540
actttctcag tcggctctgt atgtgcaggg taaatacgcttt ccgctgtgata ttggaacatt  
600
taagctatca atgcagaact tcatacatac tacaatgaaac ttaaagaaa gaggagacgtc  
660
 ttgctgctc ccaggggtct ccaattgctct gcacagccta gaaaatgagt aagctacta  
720
tsagagagctg ctagggttaa gagggaaaggt cttgtccaaga tggggtggta aatytgcctt  
780
 ttcacacat cttctgcag tttggaataa gcgtcagctg ctaggctgtc tggctcctg  
840
gctgacaca tcgatctctc tttactgca ccaatattaa ccaatctcctt ccaacaggcc  
900
cacatttgg ccagagaact ggctgggtatg gcacacacga tttggggcct gcgacactca  
960
cggctcttt ccagagcctgt cttatctgtg gtttgcttttt ttcacaggaa gaaggaagct  
1020
 gcgacatat tcatgtcact atgtctggagc gacccctggt gcggcaagcc gcggcctttt  
1080
cattcccaca aggtatgact aagtgccccc aattggcaga tatttgttcac caaggttccc  
1140
 aaatggggt caccttaaag aacctctataa agctcataaa tgcgttgcac atgttctgtg  
1200
gaacaggtg ccacattctt tttctcattt tcctcttcaaa ggcttgctgt tgttctgcag  
1260
tgcgccgag gtttgggctg cttctttcgc gaaatggat gacaaaaatg acaaggttgg  
1320
aagcagctgt gtcattgact aaccctgcgc aacgtggctc gttggctttt tgcggccactg  
1380
caaaaagta gcggcctcag cagccaggtt gcgtgatctc cctttctcttgt cgaatctggc  
1440
aacccatgat cggtgcctga ccgagagctc aaccagaca gacatcaagct cttcctactg  
1500
 ggaagctcctt cagggacac gcggagtttag gctggtgctg gattctacta aacaaatgct  
1560
tggtagaatt tttgacccca ccaagagttgc ttcagactac ctccgtctac caaacaagtct  
1620
 tttgtgtctg gcaagggagg aattctctaa aaccagacgg ctgagcaagtc ttctgctgca  
1680
 atccaaaggt ccagcagtctg atgcttcctct caataaaaagg ctttccccaca gttggatctgg  
1740
aaatggcaca gtgcacac gtccagtttgg aacatctatg aaccttaaagg caggygas
1800

ccatatttcc tggctatcga ggtgaataat cggctcagtg ttcgcaagag ctggagctg
1860

tctgctatc gcataaggg aatagttaaat tggagattat ccttcttgctta ggtggtctggc
1920
ggcttggtgc cagtcctgctg gcccagcagc ctgctcgctg cacatcagc
1980

tatgctttt cggctcagcttt cagacatcagc attgtagatg tggagcttgctg cccagtcttc
2040
ttcataaa ctaaagagcttt aatgttttggc ccttttccag ctcgactcttg ggtgttggtg
2100

ttcagcactttc tggagcgttg ggtgtggtt ccttttcttt ctttctttcttt ctttcttttt
2160

ttcagtcttg ggtgtggtgg ccttttcttt ctttcttttt ctttcttttttt ctttcttttttt
2220

ttcagttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued

Lys Ala Ala Met Gln Lys Phe Thr Thr Lys Ile Val Asp Met Met Lys
145 150 155 160

Ala Glu Lys Leu Tyr Glu Thr Gln Gly Pro Ile Ile Leu Ser Gln
165 170 175

Ile Glu Asn Glu Tyr Gly Pro Met Glu Trp Glu Leu Gly Glu Pro Gly
180 185 190

Lys Val Tyr Ser Glu Trp Ala Ala Lys Met Ala Val Asp Leu Gly Thr
195 200 205

Gly Val Pro Trp Ile Met Cys Lys Glu Asp Asp Val Pro Asp Pro Ile
210 215 220

Ile Asn Thr Cys Asn Gly Phe Tyr Cys Asp Tyr Phe Thr Pro Asn Lys
225 230 235 240

Ala Asn Lys Pro Lys Met Trp Thr Glu Ala Trp Thr Ala Trp Phe Thr
245 250 255

Glu Phe Gly Gly Pro Val Pro Tyr Arg Pro Ala Glu Asp Met Ala Phe
260 265 270

Ala Val Ala Arg Phe Ile Gln Thr Gly Gly Ser Phe Ile Asn Tyr Tyr
275 280 285

Met Tyr His Gly Gly Thr Asn Phe Gly Arg Thr Ser Gly Gly Gly Pro Phe
290 295 300

Ile Ala Thr Ser Tyr Asp Tyr Asp Ala Pro Leu Asp Glu Phe Gly Ser
305 310 315 320

Leu Arg Gln Pro Lys Trp Gly His Leu Lys Asp Leu His Arg Ala Ile
325 330 335

Lys Leu Cys Glu Pro Ala Leu Val Ser Val Asp Pro Thr Val Thr Ser
340 345 350

Leu Gly Asn Tyr Gln Glu Ala Arg Val Phe Lys Ser Glu Ser Gly Ala
355 360 365

Cys Ala Ala Phe Leu Ala Asn Tyr Asn Gln His Ser Phe Ala Lys Val
370 375 380

Ala Phe Gly Asn Met His Tyr Asn Leu Pro Pro Trp Ser Ile Ser Ile
385 390 395 400

Leu Pro Asp Cys Lys Asn Thr Val Tyr Asn Thr Ala Arg Val Gly Ala
405 410 415

Gln Ser Ala Gln Met Lys Met Thr Pro Val Ser Arg Gly Phe Ser Trp
420 425 430

Glu Ser Phe Asn Glu Ala Ala Ser His Glu Asp Thr Phe Thr
435 440 445

Val Val Gly Leu Leu Glu Gln Ile Asn Ile Thr Arg Asp Val Ser Asp
450 455 460

Tyr Leu Trp Tyr Met Thr Asp Ile Glu Ile Asp Pro Thr Glu Gly Phe
465 470 475 480

Leu Asn Ser Gly Asn Trp Pro Trp Leu Thr Val Phe Ser Ala Gly His
485 490 495

Ala Leu His Val Phe Val Asn Gly Gln Leu Ala Gly Thr Val Tyr Gly
500 505 510

Ser Leu Glu Asn Pro Lys Leu Thr Phe Ser Asn Gly Ile Asn Leu Arg
515 520 525

Ala Gly Val Asn Lys Ile Ser Leu Leu Ser Ile Ala Val Gly Leu Pro
530 535 540
<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
<td>Arg</td>
<td>Lys</td>
<td>Thr</td>
<td>Leu</td>
<td>Asn</td>
<td>Phe</td>
<td>Pro</td>
<td>Leu</td>
<td>Ile</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td>Lys</td>
<td>Thr</td>
<td>Leu</td>
<td>Asn</td>
<td>Phe</td>
<td>Pro</td>
<td>Leu</td>
<td>Ile</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>His</td>
<td>Phe</td>
<td>Val</td>
<td>Ile</td>
<td>Val</td>
<td>Ala</td>
<td>Gly</td>
<td>Thr</td>
<td>Tyr</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Tyr</td>
<td>Asp</td>
<td>Arg</td>
<td>Ala</td>
<td>Leu</td>
<td>Ile</td>
<td>Gly</td>
<td>Gly</td>
<td>Lys</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Ser</td>
<td>Ala</td>
<td>Gly</td>
<td>Ile</td>
<td>His</td>
<td>Tyr</td>
<td>Pro</td>
<td>Arg</td>
<td>Ala</td>
<td>Thr</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 9
<211> LEXENTRY: 887
<212> TYPE: PRT
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 9

| Ser | Arg | Arg | Lys | Thr | Leu | Asn | Phe | Pro | Leu | Ile | Leu | Thr | Val | Leu | Thr | Leu |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ile | His | Phe | Val | Ile | Val | Ala | Gly | Thr | Tyr | Phe | Lys | Pro | Phe | Asn | Val |
| 20  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Thr | Tyr | Asp | Arg | Ala | Leu | Ile | Gly | Gly | Lys | Arg | Arg | Met | Leu |
| 35  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ile | Ser | Ala | Gly | Ile | His | Tyr | Pro | Arg | Ala | Thr | Pro | Glu | Met | Trp | Pro |
| 50  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Asn | Val | Gly | Pro | His | Phe | Glu | Thr | Asn | Ala | Gly | Val | Leu | Gly | Pro |
| 545 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Val | Ser | Leu | Asn | Gly | Leu | Asn | Gly | Thr | Arg | Asp | Leu | Thr | Trp | Gln |
| 565 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Lys | Trp | Phe | Tyr | Lys | Val | Gly | Leu | Lys | Gly | Glu | Ala | Ser | Leu | His |
| 580 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ser | Leu | Ser | Gly | Ser | Pro | Ser | Val | Glu | Trp | Val | Glu | Gly | Ser | Leu | Val |
| 595 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ala | Gln | Lys | Glu | Pro | Leu | Ser | Trp | Tyr | Lys | Thr | Thr | Phe | Asn | Ala | Pro |
| 610 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Asp | Gly | Asn | Glu | Pro | Leu | Ala | Leu | Asp | Met | Asn | Thr | Met | Gly | Lys | Gly |
| 625 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Gln | Val | Trp | Ile | Asn | Gly | Glu | Ser | Leu | Gly | Arg | His | Trp | Pro | Ala | Tyr |
| 645 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Lys | Ser | Ser | Gly | Ser | Cys | Ser | Val | Cys | Asn | Tyr | Thr | Gly | Trp | Phe | Asp |
| 660 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Glu | Lys | Lys | Cys | Leu | Thr | Asn | Cys | Gly | Glu | Gly | Ser | Glu | Gly | Ser | Gln |
| 675 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| His | Val | Pro | Arg | Ser | Trp | Leu | Tyr | Pro | Thr | Gly | Asn | Leu | Val | Leu | Val |
| 690 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Phe | Glu | Glu | Trp | Gly | Asp | Pro | Tyr | Glu | Ile | Thr | Leu | Val | Lys | Arg |
| 705 |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Glu | Ile | Gly | Ser | Val | Cys | Ala | Asp | Ile | Tyr | Glu | Trp | Gln | Pro | Gln | Leu |
| 725 |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Leu | Aan | Trp | Gln | Arg | Leu | Val | Ser | Gly | Lys | Phe | Asp | Arg | Pro | Leu | Arg |
| 740 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Pro | Lys | Ala | His | Leu | Lys | Cys | Ala | Pro | Gly | Gln | Lys | Ile | Ser | Ser | Ile |
| 755 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Lys | Phe | Ala | Ser | Phe | Gly | Thr | Pro | Glu | Gly | Val | Cys | Gly | Asn | Phe | Gln |
| 770 |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Gln | Gly | Ser | Cys | His | Ala | Pro | Arg | Ser | Tyr | Asp | Ala | Phe | Lys | Lys | Asn |
| 785 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Cys | Val | Gly | Lys | Ser | Cys | Ser | Val | Gln | Val | Thr | Pro | Glu | Aen | Phe |
| 805 |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Gly | Gly | Asp | Pro | Cys | Arg | Asn | Val | Leu | Lys | Leu | Ser | Val | Glu | Ala |
| 820 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ile | Cys | Ser |     |     |     |     |     |     |     |     |     |     |     |     |
| 835 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

<210> SEQ ID NO 9
<211> LEXENTRY: 887
<212> TYPE: PRT
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 9

| Ser | Arg | Arg | Lys | Thr | Leu | Asn | Phe | Pro | Leu | Ile | Leu | Thr | Val | Leu | Thr | Leu |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1   |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ile | His | Phe | Val | Ile | Val | Ala | Gly | Thr | Tyr | Phe | Lys | Pro | Phe | Asn | Val |
| 20  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Thr | Tyr | Asp | Arg | Ala | Leu | Ile | Gly | Gly | Lys | Arg | Arg | Met | Leu |
| 35  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Ile | Ser | Ala | Gly | Ile | His | Tyr | Pro | Arg | Ala | Thr | Pro | Glu | Met | Trp | Pro |
| 50  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
Thr Leu Ile Ala Arg Ser Lys Glu Gly Ala Asp Val Ile Glu Thr 65 70 75 80
Tyr Thr Phe Trp Asn Gly His Glu Pro Thr Arg Gly Gln Tyr Asn Phe 85 90 95
Glu Gly Arg Tyr Asp Ile Val Lys Phe Ala Lys Val Val Gly Ser His 100 105 110
Gly Leu Phe Leu Phe Ile Arg Ile Gly Pro Tyr Ala Cys Ala Glu Trp 115 120 125
Asn Phe Gly Gly Phe Pro Ile Trp Leu Arg Asp Ile Pro Gly Ile Glu 130 135 140
Phe Arg Thr Asp Ala Pro Phe Lys Glu Glu Met Glu Arg Tyr Val 145 150 155 160
Lys Lys Ile Val Asp Leu Met Ile Ser Glu Ser Leu Phe Ser Trp Gln 165 170 175
Gly Gly Pro Ile Ile Leu Leu Gln Ile Glu Asn Glu Tyr Gly Asn Val 180 185 190
Glu Ser Ser Phe Gly Pro Lys Gly Lys Leu Tyr Met Lys Trp Ala Ala 195 200 205
Glu Met Ala Val Gly Leu Gly Ala Gly Val Pro Trp Val Met Cys Arg 210 215 220
Gln Thr Asp Ala Pro Glu Tyr Ile Ile Asp Thr Cys Asn Ala Tyr Tyr 225 230 235 240
Cys Asp Gly Phe Thr Pro Asn Ser Glu Lys Lys Pro Lys Ile Trp Thr 245 250 255
Glu Asn Trp Asn Gly Trp Phe Ala Asp Trp Gly Glu Arg Leu Pro Tyr 260 265 270
Arg Pro Ser Glu Asp Ile Phe Ala Ile Ala Arg Phe Phe Gln Arg 275 280 285
Gly Gly Ser Leu Gln Asn Tyr Tyr Met Tyr Phe Gly Gly Thr Asn Phe 290 295 300
Gly Arg Thr Ala Gly Pro Thr Gln Ile Thr Ser Tyr Asp Tyr Asp 305 310 315 320
Ala Pro Leu Asp Glu Tyr Gly Leu Leu Arg Gln Pro Lys Trp Gly His 325 330 335
Leu Lys Asp Leu His Ala Ala Ile Lys Leu Cys Glu Pro Ala Ala Val 340 345 350
Ala Ala Asp Ser Pro Gln Tyr Ile Lys Leu Gly Pro Lys Gln Glu Ala 355 360 365
His Val Tyr Arg Gly Thr Ser Asn Ile Gly Gln Tyr Met Ser Leu 370 375 380
Asn Gly Ile Cys Ala Ala Phe Ile Ala Asn Ile Asp Glu His Glu 385 390 395 400
Ser Ala Thr Val Lys Phe Tyr Gly Glu Phe Thr Leu Pro Pro Trp 405 410 415
Ser Val Val Phe Cys Gln Ile Ala Glu Ile Gln Leu Ser Thr Gln Leu 420 425 430
Arg Trp Gly His Lys Leu Gln Ser Lys Gln Trp Ala Gln Ile Leu Phe 435 440 445
Gln Leu Gly Ile Ile Leu Cys Phe Tyr Lys Leu Ser Leu Lys Ala Ser 450 455 460
-continued

Ser Glu Ser Phe Ser Gln Ser Trp Met Thr Leu Lys Glu Pro Leu Gly
 465 470 475 480
Val Trp Gly Asp Lys Asn Phe Thr Ser Lys Gly Ile Leu Glu His Leu
 485 490 495
Asn Val Thr Lys Asp Gln Ser Asp Tyr Leu Trp Tyr Leu Thr Arg Ile
 500 505 510
Tyr Ile Ser Asp Asp Ile Ser Phe Trp Glu Asn Asp Val Ser
 515 520 525
Pro Thr Ile Asp Ile Asp Ser Met Arg Asp Phe Val Arg Ile Phe Val
 530 535 540
Asn Gly Gln Leu Ala Gly Ser Val Lys Gly Lys Trp Ile Lys Val Val
 545 550 555 560
Gln Pro Val Lys Leu Val Gln Gly Tyr Asn Asp Ile Leu Leu Leu Ser
 565 570 575
Glu Thr Val Gly Leu Gln Asp Tyr Gly Ala Phe Leu Glu Lys Aep Gly
 580 585 590
Ala Gly Phe Lys Gly Gln Ile Lys Leu Thr Gly Cys Lys Ser Gly Asp
 595 600 605
Ile Asn Leu Thr Thr Ser Leu Thr Thr Tyr Gin Val Gin Gly Leu Arg Gly
 610 615 620
Glu Phe Leu Glu Val Tyr Asp Val Asn Ser Thr Glu Ser Ala Gly Trp
 625 630 635 640
Thr Glu Phe Pro Thr Gly Thr Thr Pro Ser Val Phe Ser Trp Tyr Lys
 645 650 655
Thr Lys Phe Asp Ala Pro Gly Gly Thr Asp Pro Val Ala Leu Asp Phe
 660 665 670
Ser Ser Met Gly Lys Gly Gln Ala Trp Val Asn Gly His His Val Gly
 675 680 685
Arg Tyr Trp Thr Leu Val Ala Pro Asn Asn Gly Cys Gly Arg Thr Cys
 690 695 700
Asp Tyr Arg Gly Ala Tyr His Ser Asp Lys Cys Arg Thr Asn Cys Gly
 705 710 715 720
Glu Ile Thr Gln Ala Trp Tyr His Ile Pro Arg Ser Trp Leu Lys Thr
 725 730 735
Leu Asn Asn Val Leu Val Ile Phe Glu Glu Thr Asp Lys Thr Pro Phe
 740 745 750
Asp Ile Ser Ile Ser Thr Arg Ser Thr Glu Thr Ile Cys Ala Gln Val
 755 760 765
Ser Glu Lys His Tyr Pro Leu His Lys Trp Ser Ser Glu Phe
 770 775 780
Asp Arg Lys Leu Ser Leu Met Asp Lys Thr Pro Glu Met His Leu Gin
 785 790 795 800
Cys Asp Glu Gly His Thr Ile Ser Ile Glu Phe Ala Ser Tyr Gly
 805 810 815
Ser Pro Asn Gly Ser Cys Gin Lys Phe Ser Gin Gly Lys Cys His Ala
 820 825 830
Ala Asn Ser Leu Ser Val Val Ser Gln Ala Cys Ile Gly Arg Thr Ser
 835 840 845
Cys Ser Ile Gly Ile Ser Asn Gly Val Phe Gly Asp Pro Cys Arg His
 850 855 860
Val Val Lys Ser Leu Ala Val Gln Ala Lys Cys Ser Pro Pro Asp
Leu Ser Thr Ser Ala Ser Ser

<210> SEQ ID NO 10
<br 211> LENGTH: 838
<br 212> TYPE: PRT
<br 213> ORGANISM: Lycopersicon esculentum
<br 240> SEQUENCE: 10
<br Met Gly Cys Thr Leu Ile Leu Met Leu Asn Val Leu Val Leu Leu Leu
<br 1  5  10  15
<br Gly Ser Trp Val Phe Ser Gly Thr Ala Ser Val Ser Tyr Asp His Arg
<br 20  25  30
<br Ala Ile Ile Val Asn Gly Gln Arg Arg Ile Leu Ile Ser Gly Ser Val
<br 35  40  45
<br His Tyr Pro Arg Ser Thr Pro Glu Met Trp Pro Gly Ile Ile Gln Lys
<br 50  55  60
<br Ala Lys Glu Gly Gly Val Asp Val Ile Gln Thr Tyr Val Phe Trp Asn
<br 65  70  75  80
<br Gly His Glu Pro Gln Gln Gly Lys Tyr Tyr Phe Glu Gly Arg Tyr Asp
<br 85  90  95
<br Leu Val Lys Phe Ile Lys Leu Val His Gln Ala Gly Leu Tyr Val His
<br 100 105 110
<br Leu Arg Val Gly Pro Tyr Ala Cys Ala Glu Trp Asn Phe Gly Gly Phe
<br 115 120 125
<br Pro Val Trp Leu Lys Tyr Val Pro Gly Ile Ser Phe Arg Thr Asp Asn
<br 130 135 140
<br Gly Pro Phe Lys Ala Ala Met Gln Lys Phe Thr Ala Lys Ile Val Asn
<br 145 150 155 160
<br Met Met Lys Ala Glu Arg Leu Tyr Glu Thr Gln Gly Gly Pro Ile Ile
<br 165 170 175
<br Leu Ser Gln Ile Glu Asn Glu Tyr Gly Pro Met Glu Trp Glu Leu Gly
<br 180 185 190
<br Ala Pro Gly Lys Ser Tyr Ala Glu Trp Ala Ala Lys Met Ala Val Gly
<br 195 200 205
<br Leu Asp Thr Gly Val Pro Trp Val Met Cys Lys Gln Asp Ala Pro
<br 210 215 220
<br Asp Pro Ile Ile Asn Ala Cys Asn Gly Phe Tyr Cys Asp Tyr Phe Ser
<br 225 230 235 240
<br Pro Asn Lys Ala Tyr Lys Pro Lys Ile Trp Thr Glu Ala Trp Thr Ala
<br 245 250 255
<br Trp Phe Thr Gly Phe Gly Asn Pro Val Pro Tyr Arg Pro Ala Glu Asp
<br 260 265 270
<br Leu Ala Phe Ser Val Ala Lys Phe Ile Gln Lys Gly Gly Ser Phe Ile
<br 275 280 285
<br Asn Tyr Tyr Met Tyr His Gly Gly Thr Asn Phe Gly Arg Thr Ala Gly
<br 290 295 300
<br Gly Pro Phe Ile Ala Thr Ser Tyr Asp Tyr Asp Ala Pro Leu Asp Glu
<br 305 310 315 320
<br Tyr Gly Leu Leu Arg Gln Pro Lys Trp Gly His Leu Lys Asp Leu His
<br 325 330 335
Arg Ala Ile Lys Leu Cys Glu Pro Ala Leu Val Ser Gly Asp Pro Ala
340 345 350
Val Thr Ala Leu Gly His Gln Gln Ala His Val Phe Arg Ser Lys
355 360 365
ala Gly Ser Cys Ala Ala Phe Leu Ala Asn Tyr Asp Gln His Ser Phe
370 375 380
ala Thr Val Ser Phe Ala Asn Arg His Tyr Asn Leu Pro Pro Trp Ser
385 390 395 400
Ile Ser Ile Leu Pro Asp Cys Lys Asn Thr Val Phe Asn Thr Ala Arg
405 410 415
Ile Gly Ala Glu Ser Ala Glu Met Lys Met Thr Pro Val Ser Arg Gly
420 425 430
Leu Pro Trp Gln Ser Phe Asn Glu Thr Ser Ser Tyr Glu Asp Ser
435 440 445
Ser Phe Thr Val Val Gly Leu Leu Glu Gln Ile Asn Thr Thr Arg Asp
450 455 460
Val Ser Asp Tyr Leu Trp Tyr Ser Thr Asp Val Lys Ile Asp Ser Arg
465 470 475 480
Glu Lys Phe Leu Arg Gly Gly Lys Trp Pro Trp Leu Thr Ile Met Ser
485 490 495
ala Gly His Ala Leu His Val Phe Val Asn Gly Gln Leu Ala Gly Thr
500 505 510
ala Tyr Gly Ser Leu Glu Lys Pro Lys Leu Thr Phe Ser Lys Ala Val
515 520 525
Asn Leu Arg Ala Gly Val Asn Lys Ile Ser Leu Ser Ile Ala Val
530 535 540
Gly Leu Pro Asn Ile Gly Pro His Phe Glu Thr Trp Asn Ala Gly Val
545 550 555 560
Leu Gly Pro Val Ser Leu Thr Gly Leu Asp Glu Gly Lys Arg Asp Leu
565 570 575
Thr Trp Gln Tyr Ser Tyr Lys Val Gly Leu Lys Gly Glu Ala Leu
580 585 590
Ser Leu His Ser Leu Ser Gly Ser Ser Ser Ser Val Gln Tyr Val Glu Gly
595 600 605
Ser Leu Val Ala Gln Arg Gln Pro Leu Thr Trp Tyr Lys Ser Thr Phe
610 615 620
Asn Ala Pro Ala Gly Asn Asp Pro Leu Ala Leu Asp Leu Asn Thr Met
625 630 635 640
Gly Lys Gly Glu Val Trp Ile Asn Gly Gln Ser Leu Gly Arg Tyr Trp
645 650 655
Pro Gly Tyr Lys Ala Ser Gly Asn Cys Gly Ala Cys Asn Tyr Ala Gly
660 665 670
Trp Phe Asn Glu Lys Cys Leu Ser Asn Cys Gly Glu Ala Ser Gln
675 680 685
Arg Trp Tyr His Val Pro Arg Ser Trp Leu Tyr Pro Thr Gly Asn Leu
690 695 700
Leu Val Leu Phe Glu Glu Trp Gly Glu Pro His Gly Ile Ser Leu
705 710 715 720
Val Lys Arg Glu Val Ala Ser Val Cys Ala Asp Ile Asn Glu Trp Gln
725 730 735
Pro Gln Leu Val Asn Trp Gln Met Gln Ala Ser Gly Lys Val Asp Lys
Pro Leu Arg Pro Lys Ala His Leu Ser Cys Ala Ser Gly Gln Lys Ile 755 760 765
Thr Ser Ile Lys Phe Ala Ser Phe Gly Thr Pro Gln Gly Val Cys Gly 770 775 780
Ser Phe Arg Glu Gly Ser Cys His Ala Phe His Ser Tyr Asp Ala Phe 785 790 795 800
Glu Arg Tyr Cys Ile Gly Gln Asn Ser Cys Ser Val Pro Val Thr Pro 805 810 815
Glu Ile Phe Gly Gly Gly Asp Pro Cys Pro His Val Met Lys Leu Ser 820 825 830
Val Glu Val Ile Cys Ser 835

<210> SEQ ID NO 11
<211> LENGTH: 724
<212> ORIGIN: Lycopersicon esculentum
<400> SEQUENCE: 11

Met Leu Arg Thr Asn Val Leu Leu Leu Val Ile Cys Leu Leu Asp 1 5 10 15
Phe Phe Ser Ser Val Lys Ala Ser Val Ser Tyr Asp Asp Arg Ala Ile 20 25 30
Ile Ile Asn Gly Lys Arg Lys Ile Leu Ile Ser Gly Ser Ile His Tyr Ile 35 40 45
Pro Arg Ser Thr Pro Gln Met Trp Pro Asp Leu Ile Gln Lys Ala Lys 50 55 60
Asp Gly Gly Leu Asp Val Ile Glu Thr Val Phe Thr Asp Gly His 65 70 75 80
Glu Pro Ser Pro Gly Lys Tyr Asn Phe Glu Gly Arg Tyr Asp Leu Val 85 90 95
Arg Phe Ile Lys Met Val Gln Arg Ala Gly Leu Tyr Val Asn Leu Arg 100 105 110
Ile Gly Pro Tyr Val Cys Ala Glu Trp Asn Phe Gly Gly Phe Pro Val 115 120 125
Trp Leu Lys Tyr Val Pro Gly Met Glu Phe Arg Thr Asn Asn Gln Pro 130 135 140
Phe Lys Val Ala Met Gln Gly Phe Val Glu Gly Lys Ile Val Asn Met Val 145 150 155 160
Lys Ser Glu Asn Leu Phe Glu Ser Gln Gly Gly Pro Ile Ile Met Ala 165 170 175
Gln Ile Glu Asp Gly Tyr Gly Pro Val Glu Trp Glu Ile Gly Ala Pro 180 185 190
Gly Lys Ala Tyr Thr Lys Trp Ala Ala Glu Met Ala Val Gly Leu Lys 195 200 205
Thr Gly Val Pro Trp Ile Met Cys Lys Glu Asp Ala Pro Asp Pro 210 215 220
Val Ile Asp Thr Cys Asn Gly Phe Tyr Cys Gly Gly Phe Arg Pro Asn 225 230 235 240
Lys Pro Tyr Lys Pro Lys Met Thr Glu Val Thr Gly Trp Tyr 245 250 255
Thr Lys Phe Gly Gly Pro Ile Pro Gln Arg Pro Ala Glu Asp Ile Ala
   260 265 270

Phe Ser Val Ala Arg Phe Val Gln Asn Asn Gly Ser Phe Phe Asn Tyr
   275 280 285

Tyr Met Tyr His Gly Gly Thr Asn Phe Gly Arg Thr Ser Ser Gly Leu
   290 295 300

Phe Ile Ala Thr Ser Tyr Asp Tyr Asp Ala Pro Leu Asp Glu Tyr Gly
   305 310 315 320

Leu Leu Asn Glu Pro Lys Tyr Gly His Leu Arg Asp Leu His Lys Ala
   325 330 335

Ile Lys Leu Ser Glu Pro Ala Leu Val Ser Ser Tyr Ala Ala Val Thr
   340 345 350

Ser Leu Gly Ser Asn Gln Glu Ala His Val Tyr Arg Ser Lys Ser Gly
   355 360 365

Ala Cys Ala Ala Phe Leu Ser Asn Tyr Asp Ser Arg Tyr Ser Val Lys
   370 375 380

Val Thr Phe Gln Asn Arg Pro Tyr Asn Leu Pro Pro Trp Ser Ile Ser
   385 390 395 400

Ile Leu Pro Asp Cys Lys Thr Ala Val Tyr Asn Thr Ala Gln Val Asn
   405 410 415

Ser Gln Ser Ser Ser Ile Lys Met Thr Pro Ala Gly Gln Gly Leu Ser
   420 425 430

Trp Gln Ser Tyr Asn Glu Glu Thr Pro Thr Ala Asp Asp Ser Asp Thr
   435 440 445

Leu Thr Ala Asn Gly Leu Trp Glu Gln Asn Val Thr Arg Asp Ser
   450 455 460

Ser Asp Tyr Leu Trp Tyr Met Thr Asn Val Asn Ile Ala Ser Asn Glu
   465 470 475 480

Gly Phe Leu Lys Asn Gly Lys Pro Tyr Leu Thr Val Met Ser Ala
   485 490 495

Gly His Val Leu His Val Phe Val Asn Gly Lys Leu Ser Gly Thr Val
   500 505 510

Tyr Gln Thr Leu Asp Asn Pro Lys Leu Thr Tyr Ser Gly Asn Val Lys
   515 520 525

Leu Arg Ala Gly Ile Asn Lys Ile Ser Leu Leu Ser Val Ser Val Gly
   530 535 540

Leu Pro Asn Val Gly Val His Tyr Asp Thr Trp Asn Ala Gly Val Leu
   545 550 555 560

Gly Pro Val Thr Leu Ser Gly Leu Asn Gly Asp Arg Asn Leu Ala
   565 570 575

Lys Gln Lys Trp Ser Tyr Lys Val Gly Leu Lys Gly Glu Ser Leu Ser
   580 585 590

Leu His Ser Leu Ser Gly Ser Ser Val Glu Trp Val Arg Gly Ser
   595 600 605

Leu Met Ala Gln Lys Gln Pro Leu Thr Trp Tyr Lys Ala Thr Phe Asn
   610 615 620

Ala Pro Gly Gln Asp Pro Leu Ala Leu Asp Met Ala Ser Met Gly
   625 630 635 640

Gly Lys Ile Trp Ile Asn Gly Glu Val Gly Arg His Trp Pro
   645 650 655

Gly Tyr Ile Ala Gln Gly Asp Cys Ser Lys Cys Ser Tyr Ala Gly Thr
Phe Aan Glu Lys Lys Cys Gln Thr Aan Cys Gly Gln Pro Ser Gln Arg
675 680 685
Trp Tyr His Val Pro Arg Ser Trp Lys Pro Ser Gly Asn Leu Leu
690 695 700
Val Val Phe Glu Glu Trp Gly Gly Asn Pro Thr Gly Ile Ser Leu Val
705 710 715 720
Arg Arg Ser Arg

<210> SEQ ID NO 12
<211> LENGTH: 251
<212> TYPE: PRT
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 12

Ile Gln Thr Tyr Val Phe Trp Aan Leu His Glu Pro Val Arg Aen Gln
1 5 10 15
Tyr Ase Phe Glu Gly Arg Lys Aes Leu Ile Aen Phe Val Lys Leu Val
20 25 30
Glu Aan Ala Gly Leu Phe Val His Ile Arg Ile Gly Pro Tyr Val Cys
35 40 45
Aen Glu Trp Aen Tyr Gly Gly Phe Pro Leu Trp Leu His Phe Ile Pro
50 55 60
Gly Ile Glu Phe Arg Thr Ase Aen Glu Pro Phe Lys Ala Glu Met Lys
65 70 75 80
Arg Phe Thr Ala Lys Ile Val Aen Met Ile Lys Gln Glu Aen Leu Tyr
85 90 95
Aen Ser Glu Gly Gly Pro Val Ile Leu Ser Glu Glu Asn Glu Tyr
105 110
Gly Aen Gly Ase Ile Glu Ser Arg Tyr Gly Pro Arg Aen Lys Pro Tyr
115 120 125
Val Aan Trp Ala Ala Ser Met Ala Thr Ser Leu Aen Thr Gly Val Pro
130 135 140
Trp Val Met Cys Glu Gln Pro Asp Ala Pro Ser Val Ile Aen Thr
145 150 155 160
Cys Aen Gly Phe Tyr Cys Asp Glu Gln Phe Lys Glu Aen Ser Asp Lys Thr
165 170 175
Pro Lys Met Trp Thr Glu Aen Trp Thr Gly Trp Phe Leu Ser Phe Gly
180 185 190
Gly Phe Val Pro Tyr Arg Pro Val Glu Ase Ala Phe Ala Val Ala
195 200 205
Arg Phe Phe Glu Arg Gly Gly Thr Phe Glu Aen Tyr Tyr Met Tyr His
210 215 220
Gly Gly Thr Aen Phe Gly Arg Thr Ser Gly Gly Pro Phe Ile Ala Thr
225 230 235 240
Ser Tyr Asp Tyr Ase Ala Pro Leu Asp Glu Tyr
245 250

<210> SEQ ID NO 13
<211> LENGTH: 249
<212> TYPE: PRT
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 13
Ile Glu Thr Tyr Val Phe Trp Asn Val His Glu Pro Ser Pro Gly Asn
1  5  10  15
Tyr Asn Phe Glu Gly Arg Tyr Asp Val Arg Phe Val Lys Thr Ile
20 25  30
Gln Lys Ala Glu Leu Tyr Ala His Leu Arg Ile Gly Pro Tyr Val Cys
35 40  45
Ala Glu Trp Asn Phe Gly Phe Pro Val Trp Leu Lys Tyr Val Pro
50 55  60
Gly Ile Ser Phe Arg Ala Asp Asn Glu Pro Phe Lys Asn Ala Met Lys
65 70  75  80
Gly Tyr Ala Glu Lys Ile Val Asn Leu Met Lys Ile Ile Phe Ser
85 90  95
Ser Leu Arg Val Val Gln Ser Tyr Ser His Arg Leu Arg Met Ser Met
100 105  110
Gly Leu Lys Pro Arg Tyr Leu Glu His Arg Asp Ile Ser Ile Gln His
115 120  125
Gly Leu Gln Ile Trp Gln Leu Asp Pro Val Thr Gly Val Pro Thr Val
130 135  140
Met Cys Lys Glu Glu Phe Arg Ala Phe Asp Val Pro Ile Asn Thr Cys Asn
145 150  155  160
Gly Phe Tyr Cys Asp Asn Phe Pro Asp Pro Val Lys Pro Tyr Lys Pro Ala
165 170  175
Ile Trp Thr Glu Ala Trp Gly Thr Phe Ser Glu Phe Gly Gly Pro
180 185  190
Leu His Gln Arg Pro Val Gln Asp Leu Ala Phe Ala Val Ala Gln Phe
195 200  205
Ile Gln Arg Gly Gly Ser Phe Val Asn Tyr Tyr Met Tyr His Gly Gly
210 215  220
Thr Asn Phe Gly Arg Thr Ala Gly Pro Phe Ile Thr Thr Ser Tyr
225 230  235  240
Asp Tyr Asp Ala Pro Leu Asp Glu Tyr
245

<210> SEQ ID NO 14
<211> LENGTH: 770
<212> TYPE: PRT
<213> ORGANISM: Lycopersicon esculentum

<400> SEQUENCE: 14
Met Asn Thr Met Ser Cys Leu Ser Ser Asn Phe Lys Phe Val Phe Leu
1  5  10  15
Ala Ser Thr Val Ile Trp Met Thr Val Met Ser Ser Leu Ala Ala
20 25  30
Val Asp Ala Ser Asn Val Thr Ile Gly Thr Asp Ser Val Thr Tyr
35 40  45
Asp Arg Arg Ser Leu Ile Asn Gly Gin Arg Lys Leu Ile Ser
50 55  60
Ala Ser Ile His Tyr Pro Arg Ser Val Pro Ala Met Trp Pro Gly Leu
65 70  75  80
Val Arg Leu Ala Lys Glu Gly Val Asp Val Ile Glu Thr Tyr Val
85 90  95
Phe Trp Asn Gly His Glu Pro Ser Pro Gly Asn Tyr Phe Gly Gly
Arg Phe Asp Leu Val Lys Phe Cys Lys Ile Ile Gln Gln Ala Gly Met
Tyr Met Ile Leu Arg Ile Gly Pro Phe Val Ala Ala Glu Trp Asn Phe
Gly Gly Leu Pro Val Trp Leu His Tyr Val Pro Gly Thr Thr Phe Arg
Thr Asp Ser Glu Pro Phe Lys Tyr His Met Glu Lys Phe Met Thr Tyr
Thr Val Asn Leu Met Lys Arg Glu Arg Leu Phe Ala Ser Gln Gly Gly
Pro Ile Ile Leu Ser Gln Val Glu Asn Glu Tyr Gly Tyr Glu Ala
Ala Tyr Gly Glu Gly Gly Lys Arg Tyr Ala Leu Trp Ala Ala Lys Met
Ala Leu Ser Gln Asn Thr Gly Val Pro Trp Ile Met Cys Gln Gly Tyr
Asp Ala Pro Asp Pro Val Ile Asp Thr Cys Asn Ser Phe Tyr Cys Asp
Gln Phe Lys Pro Ile Ser Pro Asn Lys Pro Lys Ile Trp Thr Glu Asn
Trp Pro Gly Trp Phe Lys Thr Phe Gly Ala Arg Asp Pro His Arg Pro
Ala Glu Asp Val Ala Tyr Ser Val Ala Arg Phe Gln Gly Gly
Ser Val Gln Asn Tyr Met Tyr His Gly Glu Thr Asn Phe Gly Arg
Thr Ala Gly Gly Pro Phe Ile Thr Thr Ser Tyr Asp Tyr Asp Ala Pro
Ile Asp Glu Tyr Gly Leu Pro Arg Phe Pro Lys Trp Gly His Leu Lys
Glu Leu His Lys Val Ile Lys Ser Cys Glu His Ala Leu Leu Asn Asn
Asp Pro Thr Leu Leu Ser Leu Gly Pro Leu Gin Glu Ala Asp Val Tyr
Glu Asp Ala Ser Gly Ala Cys Ala Ala Phe Leu Ala Asn Met Asp Asp
Lys Asn Asp Lys Val Val Gln Phe Arg His Val Ser Tyr His Leu Pro
Ala Trp Ser Val Ser Ile Leu Pro Asp Cys Lys Asn Val Ala Phe Asn
Thr Ala Lys Val Gly Cys Gin Thr Ser Ile Val Asn Met Ala Pro Ile
Asp Leu His Pro Thr Ala Ser Ser Pro Lys Arg Asp Ile Lys Ser Leu
Gln Trp Glu Val Phe Lys Glu Thr Ala Gly Val Trp Gly Val Ala Asp
Phe Thr Lys Asn Gly Phe Val Asp His Ile Asn Thr Thr Lys Asp Ala
Thr Asp Tyr Leu Trp Tyr Thr Ser Ile Phe Val His Ala Glu Glu
Asp Phe Leu Arg Asn Arg Gly Thr Ala Met Leu Phe Val Glu Ser Lys

515 520 525

Gly His Ala Met His Val Phe Ile Asn Lys Leu Gln Ala Ser Ala

530 535 540

Ser Gly Asn Gly Thr Val Pro Gln Phe Lys Gly Thr Pro Ile Ala

545 550 555 560

Leu Lys Ala Gly Lys Asn Glu Ile Ser Leu Leu Ser Met Thr Val Gly

565 570 575

Leu Gln Thr Ala Gly Ala Phe Tyr Glu Trp Ile Gly Ala Gly Pro Thr

580 585 590

Ser Val Lys Val Ala Gly Phe Thr Gly Thr Met Asp Leu Thr Ala

595 600 605

Ser Ala Trp Thr Tyr Lys Ile Gly Leu Gin Gly Gly Lys His Leu Arg Ile

610 615 620

Gln Lys Ser Tyr Asn Leu Lys Ser Lys Ile Trp Ala Pro Thr Ser Gin

625 630 635 640

Pro Pro Lys Gln Gin Pro Leu Thr Trp Tyr Lys Ala Val Val Asp Ala

645 650 655

Pro Pro Gly Asn Glu Pro Val Ala Leu Asp Met Ile His Met Gly Lys

660 665 670

Gly Met Ala Trp Leu Asn Gly Gin Glu Ile Gly Arg Tyr Trp Pro Arg

675 680 685

Arg Thr Ser Lys Tyr Glu Asn Cys Val Thr Gin Cys Asp Tyr Arg Gly

690 695 700

Lys Phe Asn Pro Asp Lys Cys Val Thr Gly Cys Gin Pro Thr Gin

705 710 715 720

Arg Trp Tyr His Val Pro Arg Ser Trp Phe Lys Pro Ser Gly Asn Val

725 730 735

Leu Ile Ile Phe Glu Glu Ile Gly Gly Asp Pro Ser Gin Ile Arg Phe

740 745 750

Ser Met Arg Lys Val Ser Gly Ala Cys Gly His Leu Ser Val Asp His

755 760 765

Pro Ser Phe Asp Val Gin Asn Leu Gin Gly Ser Glu Ile Asn Asp

770 775 780

Lys Asn Arg Pro Thr Leu Ser Leu Lys Cys Pro Thr Asn Thr Asn Ile

785 790 795 800

Ser Ser Val Lys Phe Ala Ser Phe Gly Asn Pro Asn Gly Thr Cys Gly

805 810 815

Ser Tyr Met Leu Gly Asp Cys His Asp Gin Asn Ser Ala Ala Leu Val

820 825 830

Glu Lys Val Cys Leu Asn Gin Gin Cys Ala Leu Gin Met Ser Ser

835 840 845

Ala Asn Phe Asn Met Gin Leu Cys Pro Ser Thr Val Lys Leu Ala

850 855 860

Val Glu Val Asn Cys Ser

865 870
What we claim is:

1. An isolated nucleic acid molecule comprising a poly-nucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

   (a) a nucleotide sequence encoding the β-galactosidase II polypeptide having the complete amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15 and SEQ ID NO:16 and designated TBG1, TBG2, TBG3, TBG4, TBG5, TBG6 and TBG7, respectively as shown in FIG. 2 or as encoded by the cDNA clone selected from the group consisting of cDNA clones contained in Gen Bank Accession No. AF023847, AF1544420, AF154421, AF020390, AF154423, AF154424 and AF154422;

(b) a nucleotide sequence encoding the mature β-galactosidase II polypeptide having the amino acid sequence at about positions 24-724 selected from the group consisting of SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15 and SEQ ID NO:16 and designated TBG1, TBG2, TBG3, TBG4, TBG5, TBG6 and TBG7, respectively as shown in FIG. 2 or as encoded by the cDNA clone selected from the group consisting of cDNA clones contained in Gen Bank Accession No. AF023847, AF1544420, AF154421, AF020390, AF154423, AF154424 and AF154422; and

   (c) a nucleotide sequence complementary to any of the nucleotide sequences in (a) or (b), above.

2. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7 as shown in FIG. 2.

3. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence in FIG. 2 (SEQ ID NO:4) encoding the β-galactosidase II polypeptide having the amino acid sequence designated TBG4 in FIG. 2.

4. The nucleic acid molecule of claim 1 wherein said polynucleotide has the nucleotide sequence in FIG. 2 (SEQ ID NO:4) encoding the mature polypeptide having the amino acid sequence from about 24 to about 724 in the amino acid sequence designated TBG4 in FIG. 2.

5. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in Gen Bank Accession No. AF023847.

6. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in Gen Bank Accession No. AF1544420.

7. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in Gen Bank Accession No. AF154421.

8. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in Gen Bank Accession No. AF020390.

9. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in Gen Bank Accession No. AF154423.

10. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in Gen Bank Accession No. AF1544424.

11. The nucleic acid molecule of claim 1 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in Gen Bank Accession No. AF154422.

12. An isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide having a nucleotide sequence identical to a nucleotide sequence in (a), (b), or © of claim 1 wherein said polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues.

13. An isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a β-galactosidase II polypeptide having an amino acid sequence in (a), (b), or © of claim 1.

14. A method for making a recombinant vector comprising inserting an isolated nucleic acid molecule of claim 1 into a vector.


16. A method of making a recombinant host cell comprising introducing the recombinant vector of claim 15 into a host cell.

17. A recombinant host cell produced by the method of claim 16.

18. A recombinant method for producing β-galactosidase II polypeptide, comprising culturing the recombinant host cell of claim 17 under conditions such that said polypeptide is expressed and recovering said polypeptide.

19. An isolated β-galactosidase II polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:

   a) amino acid sequence at about positions 24-724 selected from the group consisting of sequences SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15 and SEQ ID NO:16 and designated TBG1, TBG2, TBG3, TBG4, TBG5, TBG6 and TBG7, respectively as shown in FIG. 2; and

   b) amino acid sequence as encoded by the cDNA clone selected from the group consisting of cDNA clones contained in Gen Bank Accession No. AF023847, AF1544420, AF154421, AF020390, AF154423, AF154424 and AF154422.

20. An isolated polypeptide comprising an epitope-bearing portion of the β-galactosidase II protein.

21. An isolated antibody that binds specifically to a β-galactosidase II polypeptide of claim 20.

22. An isolated nucleic acid molecule nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group consisting of:

   (a) a nucleotide sequence encoding the β-galactosidase II polypeptide having the complete amino acid sequence
selected from the group consisting of SEQ ID NO:8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO: 15 and SEQ ID NO: 16 and designated TBG1, TBG2, TBG3, TBG4, TBG5, TBG6 and TBG7, respectively as shown in FIG. 3 or as encoded by the cDNA clone selected from the group consisting of cDNA clones contained in Gen Bank Accession No. AF023847, AF1544420, AF154421, AF020390, AF154423, AF154424 and AF154422;

(b) a nucleotide sequence encoding the mature β-galactosidase II polypeptide having the amino acid sequence at about positions 24-724 selected from the group consisting of sequences SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 and SEQ ID NO: 16 and designated TBG1, TBG2, TBG3, TBG4, TBG5, TBG6 and TBG7, respectively as shown in FIG. 3 or as encoded by the cDNA clone selected from the group consisting of cDNA clones contained in Gen Bank Accession No. AF023847, AF1544420, AF154421, AF020390, AF154423, AF154424 and AF154422; and

(c) a nucleotide sequence complementary to any of the nucleotide sequences in (a), or (b), above.

23. The nucleic acid molecule of claim 22 wherein said polynucleotide has a complete nucleotide sequence in FIG. 2 selected from the group consisting of SEQ ID NOs: 1-3 and 5-7.

24. The nucleic acid molecule of claim 22 wherein said polynucleotide has a nucleotide sequence in FIG. 2 selected from the group consisting of SEQ ID NOs: 1-3 and 5-7 encoding the β-galactosidase polypeptide having the complete amino acid sequence designated TBG1-3 and 5-7, respectively.

25. The nucleic acid molecule of claim 22 wherein said polynucleotide has the nucleotide sequence in FIG. 2 selected from the group consisting of SEQ ID NOs: 1-3 and 5-7 encoding the mature polypeptide having the amino acid sequence designated TBG1-3 and 5-7, respectively.

26. The nucleic acid molecule of claim 22 wherein said polynucleotide has the complete nucleotide sequence of the cDNA clone contained in an Gen Bank Accession No. selected from the group consisting of ATCC Deposit No. selected from the group consisting of AF023847, AF1544420, AF154421, AF020390, AF154423, AF154424 and AF154422.

27. A method of modifying cell wall metabolism in a plant which comprises transforming said plant with a DNA construct adapted to modify the activity of a β-galactosidase, growing said plant or its descendent and selecting a plant having modified cell wall characteristics, said construct comprising a transcriptional initiation region operative in plants operably linked to a DNA sequence encoding at least one β-galactosidase.

28. A method as claimed in claim 27, wherein said DNA sequence is selected from the group consisting of the sequences of nucleic acid molecules claimed in claim 1 or claim 22.

29. A plant cell transformed with a nucleic acid molecule as claimed in claim 1 or claim 22.

30. A plant derived from a plant cell as claimed in claim 29.

31. A plant seed derived from a plant as claimed in claim 30.

32. A method for modifying β-galactosidase gene expression in a plant comprising transforming said plant with a nucleic acid molecule as claimed in claim 1 or claim 22, growing the transformed plant and selecting a plant having modified β-galactosidase gene expression when compared with an untransformed plant.

* * * * *