ARTICLE

High-resolution physical and transcript map of human chromosome 2p21 containing the sitosterolaemia locus

Kangmo Lu¹, Mi-Hye Lee¹, John D Carpten², Mundeep Sekhon³ and Shailendra B Patel^{*,1}

¹Division of Endocrinology, Diabetes and Medical Genetics, Department of Medicine, Medical University of South Carolina, Charleston, SC 29403, USA; ²Prostate Cancer Investigation Group, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA; ³Genome Sequencing Center, Washington University School of Medicine, St. Louis, MO 63108, USA

Sitosterolaemia (phytosterolaemia) is an autosomal recessive disorder characterised by the presence of tendon xanthomas in the face of normal or mildly elevated plasma cholesterol levels, premature atherosclerotic disease and has diagnostically elevated plasma and tissue plant sterol concentrations. Affected individuals show an increased absorption of both cholesterol and sitosterol from the diet, decreased bile clearance of these sterols and their metabolites resulting in markedly expanded whole body cholesterol and sitosterol pools. The defective gene is therefore hypothesised to play a crucial role in regulating dietary cholesterol absorption, and its elucidation may shed light on these molecular processes. We have previously localised the defective gene to human chromosome 2p21, between microsatellite markers D2S1788 and D2S1352, a distance of approximately 15 cM. Recently, the disease locus interval has been narrowed to lie between D2S2294 and D2S2291/D2S2174. We have constructed a high-resolution YAC and BAC contigs by using known STSs and generating novel STSs from the minimal interval. Eight previously identified genes and 60 ESTs were mapped to these contigs. The BAC contig contains 60 BAC clones and 108 STSs and encompasses a physical distance of approximately 2.0 cM between microsatellite markers D2S2294 and D2S2291. These results will not only facilitate cloning of the sitosterolaemia gene, but also other disease genes located in this region, and accelerate sequencing of the corresponding genomic clones. European Journal of Human Genetics (2001) 9, 364-374.

Keywords: BAC contig; mapping; positional cloning; atherosclerosis genes

Introduction

Sitosterolaemia (also known as phytosterolaemia, MIM number 210250) is a rare autosomal recessively inherited metabolic disorder, which was described in 1974 in two affected sisters.¹ Sitosterolaemic patients develop tendon and tuberous xanthomas, haemolytic episodes, arthralgias and arthritis, and premature coronary and aortic atherosclerosis leading to cardiac fatalities.^{1–5} Affected individuals have very high levels of plasma plant sterols (sitosterol, campesterol,

E-mail: patelsb@musc.edu

stigmasterol, avenosterol) and their 5α -saturated stanols, particularly sitostanol, but their blood cholesterol levels may be normal or only moderately increased.^{1,4} Increased intestinal absorption and decreased hepatic excretion of sitosterol (the major plant sterol) may be responsible for the accumulation of these non-cholesterol sterols in plasma and tissues of affected patients.^{4,6–10}

In addition to the proposed defects of absorption and excretion of sitosterol, reduced whole body cholesterol synthesis has also been noted.^{2,11,12}

Linkage analyses of 10 well-characterised pedigrees localised the genetic defect to human chromosome 2p21, between microsatellite markers *D2S1788* and *D2S1352*.¹³ Recently, we have narrowed this interval to lie between microsatellite markers *D2S2294* and *D2S2291* (Lee *et al*, manuscript submitted). To refine the minimal critical region

^{*}Correspondence: SB Patel, Division of Endocrinology, Diabetes and Medical Genetics, STR 541, 114 Doughty Street, Charleston, SC 29403, USA. Tel: +1 843 876 5227; Fax: +1 843 876 5133;

Received 26 September 2000; revised 11 December 2000; accepted 19 December 2000

Table 1 Primer sequences used in YAC/BAC contigs

| C3900151F CCAGTGCCATTACTACATTA ANTGCCACTGATCACACAC 207 A2031294 C3900151F CCAGTGCCCACCTTTA GTCAAGAGTACATCACC 206 A2031296 C-4382A5 TCTTGTTTCCCATTCCTGCA CCAGTCACCCACATACACGTCA 218 A2051297 C-328A6.R CCCCCGCACATTCCCACGCTC CAGACACCTCACCTCGCC 306 A2051298 C-328A5.F CCCCGCACTTTCCACGCTC CAGACACCTCACCTCGCCCCAATACACGCT 306 A2051298 C-328A5.F CCCCGCCCTTTACACGC CAGACACCTCACCTCGCCC 218 A0068138 R>900.233.F AACCCATCGCTCTTACACGC ACGACACCTCACCTCCCCC 298 A0391175 R316810.F CCCACACCTTCCACCCCCCCCC 298 A0391174 R318810.F ACACACTTGCACACCTCCACCCCCCCCCACCACCCC 298 A0391174 R318810.F CACACTTGCACACCTCCACCCCCCCCCCCACGCACCCCACCACCCCACCCCCACCCACCCCCC | ^a STS | Forward primer (5′ – 3′) | Reverse primer (5' – 3') | Size (bp) | ^b Accession no. |
|--|------------------------|--------------------------|--------------------------|------------|----------------------------|
| C-590D15.R AMAAGCCTCCCCACCTTTA CTCAMTCTCACACTCTAC 206 A2051295 C-498C3.F CCCCACCACCAATACACGTCAC CACTACACTCTACCTC 203 A2051296 C-528A6.F CCCCCACCACCACTACACGTCAC CACGACACTCCCCCAT 206 A2051297 C-528J0.F TCCCTCACCTCCACGACACCTCCCACT CACGACACCCTCACTCACTCCCCACT 412 A2051296 S790A23.F TCCCGATCGACCTCACTAC CACGACACCCTCCTCATTCAC 412 A20201297 S790A23.F CCGATCGACCTCACTAC CACGACACCCTCCTCACT 298 A0698136 S790A23.F CCGATCGACACCTCACTAC CGTCACGATTCACCACCTTCCC 299 A054115 S790A23.F CCGATCGACCTCACCAC CGTCACGATTCACCACCCCC 298 A059714 S790A3.F CCCATAGATCACCACCAT CTGTCACGTCACCACCACT 288 A059714 S790B3.R ACACCATCGACCATATCTACCACCACT CCTGTCACTCACACCCACTACTACCACACT 230 A064110.F S790B3.R ACACCATCGACCATATATTCCCGC CCCTGTCGACTCACACCCACTA 322 A0649814 S7201.F ACACCATCGACCATATATTCCCGCCCTCTACCCAC 232 A0649814 S72 | C-506D15.F | CCAGTGGCATTTAGTACATTA | AATGCCACTGAATCACACAC | 207 | AZ051294 |
| C-4982.3F TCTIGTITICCATTCIGTC GTCATTTCTACACTCIGAC 248 AZ051295 C-528A.6F CCCACCACCAMIACACTGCA CACITACATTCGCCCCTCCT 218 AZ051297 C-528A.6F CCCCACCACAMIACACTCCACGCT CACITACATTCGCC 266 AZ051298 C-528A.6F CCCCACCACTCCACAGTC CACACACCTCACTCGCTCA 306 AZ051298 R-990A.23.7 ACCCATCCCTCCACAGTC CACACACCTCACTCCCACTCCACCCCACTC 28 A0668136 R-318810.F CCCACACCACTCCCAC CGCAGACGTCACCTCCCACCCCCACCCACT 288 A0597714 R-318810.F CACACACTATCACCCACTCCAC CGTCATACATTCACCACACTTCAC CGTCATACATTCACCACACTTCACC AC059714 R-318810.F CACACACTATCACCCACACTCAC CGTCAGACTTCCACCCCCACCAC 28 A0597714 R-318810.F CACACTATCACCCACACTCACG TCGTCAGATTCACCCCACCACT 28 A059774 R-32811.8 CCTCACACACTTCAACGCCACTTCCACG TCGTCAGACACTCCCAACCACT 228 A059734 R-22011.8 ACCATCGTCATCACCCACACTCGC TCGTCAGACTCCCCACACTCGC 24 A2051307 R-22013.5 ACCATTCACCCCATTCACCCGCACTTATAC CGCAACTCTCCCCCTTCACCCACCACTT | C-506D15.R | AAAAAGGCTGCCCACCTTTA | GTCAAGAGGTAGATGAAATGC | 206 | AZ051295 |
| C-328A6.F CCCCACCAVITACCATGAA CAGTACATCTCCGCGCTCTA 218 A205129 C-3590.62 R CCCTATATTTACCCACGCTC CAGACACACTCTCACCATG CAGACACCCTCACTGCC 266 A2051298 C-3590.62 R CCCTCATATAGCTCACCATG CAGACACCCTCACTGCCTTA R CCCTCACGCCTCACCATG CAGACACCCTCACTGCCTTA R CCCTCACGCCTCACCATG CAGACACCCTCACTGCCTTA R CCCTCACGCCTCACCATG CAGACACCCTCACTGCCTTA R CCCCCCTCACGCATCCCC R CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC | C-498C3.F | TCTTGTTTTCCGATTCTGTTC | GTCAATTTCTACAGTGTAGCT | 203 | AZ051296 |
| C-328A.8.R CCCTGATATTIACCCAGCTC CAGAGAACTGAGTCTGCG 266 A2051299 R-990A23.F AACCGATGGCTCTGTAGCAGTG CAGAGACGCTGGTGCTGATGGAGTG R-900A23.F AACCGATGGCTCTGTAGCAGTG R-900A23.F CTGGATGGACGTCACTACAGG R-1081G2.F CATAAACTTTAAGCGTTATTCA CAGAGACGCTACGGCTGTGCACTA R-1081G2.F CATAAACTTTAAGCGCTATTCC CAGAAACGCTACGCGCTGTCACTA R-1081G2.F CATAAACTTTAAGCGCAA CGTGAGGTTCCCCCCCCC R-316810.R ACACGTAGGAAGTCTCACCT R-316810.R ACACGTAGGAAGTCTCACCA CGTGAGGTTCCCCCCCCCC R-30590h.3.R ACACGTAGGAAGTCACCAC R-205010.R CACACGTCAGGAAGTCGCA CCCCAAGTCAGCACGTCTCACATGCA R-20610.R ACGGTAAGTCACGCCATGTG R-20610.R CACAGCTCCTGCCGCCGCCCCCCCCCCCCCCCCCCCCCC | C-528A6.F | CCCAGCACCAAATACAGTGAA | CAGTACATCTTCCGGCTCTA | 218 | AZ051297 |
| C-56916.F TACCTCAGCTTCAGCATG CACACACCTCCACTGCATA 306 A205129 R-990A23.F CTGCATCGACCTCACAGG ACGACAGCTCAGGCTAACA 412 AQ202225 R-990A23.F CTGCATCGACCCTCACAG CACACAGCTCACGCTGTATC 412 ACGACACACGCTCACCCTCACA CACGACGCTCACGCTCACGCTGTAC 275 AQ240126 R-316810.F CACACACTATGACCACTTCAC CCTGAGCTCTCACCCTCAC 299 AQ341175 R-316810.F CACACACTATGACCACTTCAC CCTGAGCTCTCCCCCCACCC 288 AQ509714 R-32814.R CGTCACACACTTCAGCCACCT CCGCCCCACACGCT 288 AQ509714 R-23011.R ACACACTACACGCACACCCA CTGCGCCCCCCCCC 206 AQ59167 R-2301.8 ACCACCTGACACGCTACCCC CCCCACAGCACTTCAC 215 AQ667844 R-27011.R ACACACTACACCCTACTCCCC CCCACATACATCTCACCTCA 215 AQ667844 R-27011.R ACACACTACACCTCATCCCC CCCCACATACATCTACACCTCACC 214 AQ418643 C-2301.5 ACCACCTACACCCTATAC TTCATCCCCC CCCACATACAT | C-528A6.R | CCCTGATATTTACCCAGCTC | CAAGAAGAATGAGATCTGGC | 266 | AZ051298 |
| R-990A23.F ACGCATGCCTTATACAGG ACGAACCACGTATATTAC 412 AQ702225 R-90A23.F CTGGATGCAACGTCACTA CAGAAGCTTACCACGTATCC CAGAAGCTTACCTCTAC CAGAAGCTTACCCCTTAC 275 AQ740126 R-10810.F CACAACGTATTGACCACTTTAC CAGAAGCTTCCCCCTTAC 299 AQ341175 R-316810.R ACACGTAAGGATATGACCACTTTAC CGTGAGGTTCCCC 299 AQ341175 R-316810.R ACACGTAAGGATATGACGCAG CTGTGGGTGACAACGTCACCTTC 206 AQ359164 R-2201.R CAGATGCTGTGTGAGGCTGTGAGG CCCTAAGCTACACGCACTA 200 AQ35923 R-2201.R CAGGTGCTGTGTGAGGCAGG CCCTAAGCTACACGACTA 200 AQ485023 R-2201.R CAGGTGCTGTGTGAGGCAGG CCCTAAGCTACACGACTA 200 AQ485023 R-2201.S.R ATCTCCTCTAGGCTGTGTG CCCTAAGCTACTGACGACTA 200 AQ485023 R-2201.S.R ATCTCCTCTAGGCTGTGTG TAACCCTACACGACTAGTGCCC 284 AQ322528 R-2201.S.R TGTGCTCTCTAGG TCACCTACACGACTGTGCTGC 284 AQ31329 R-2201.S.R TGTGCAGCCACACGTGACGCACACGCCCCCGAGGTCTGCCCCCCCC | C-569J16.F | TACCTGAGCTTCCTAGGATG | CAGACAGCCTCAGTCGCTA | 306 | AZ051299 |
| R-990A23.R CTGGATGGAACGCTCACTA CAGAAGCCTCAGGTAAGA 268 AQ668383 R-1081G2.F GATAMAGTTMAGCTTATGCT CAAAATGGACACCTTACC 275 AQ740126 R-316810.F CACAACTATGACCACTTGAC CGTGGAGCTTCTCCTTCC 299 AQ541175 R-316810.F CACAACTATGACCACTTGACCAC CTATGGATAGCATATCAGCCAC CTATGGATAGCACACTT 218 AQ57914 R-318810.F CACAACTATGACCACCTTACCAC CTATGGATGAGAGTTCTCAC 312 AQ567344 R-59501.3R ACAACACTGACAGCATTCTAGC CCACGTAGAGAGTTCTACAC 312 AQ567344 R-5020.5F ACACTTCTGCACGCTATAGCAC TGCATGTGAGCCACACACAC 312 AQ511801 R-22131.8 CGCTAACAGCCAMTCTAGG TGCAGCTGCACTGCACAC 288 AQ322532 R-2213.8 TGGACTGCACACAGACGTAC GACAACTGCCACACACAGAC AQ51302 R-2213.8 TGGACTGCACACACAGAC GACACTGCCACACCACACAC 28 AQ322532 R-2213.8 TGGACTGCACACTGCAGAC CCTCACACAGCAGTGTCCCCACA 29 AQ51302 R-2413.7 TGAGCACTCCACAGTCCCACACTCCCCACA 248 AQ326533 AQ26531302 | R-990A23.F | AACGCATGGCTCTATAGAGG | AGGACAGCAGTGTCATTTAC | 412 | AQ702225 |
| R:1081C2.F GATAAAGTTTAGCTATICTC GAMATAGCTTAGCCCTGTAC 275 AQ240126 R:316810.R ACACCTAAGCATATGACCCAT CGTGAGCTICTCCTTTCCC 299 AQ341175 R:316810.R ACACCTAAGCATATGACCCAT GGTGAGCTICTCCTTTCCC 288 AQ507714 R:316810.R ACACCTAAGCATTGACCAGCTA GTGTGAGTGACCTGC 215 AQ56734 R:316810.R ACACCTAAGCACTGACAGCG CCCATAGTGACCACCA 322 AQ516454 R:22010.R ACACGTGATGACGCGCTGATATA TIGATCTAGCACCACCA 320 AQ285023 R:22011.R CACGTGACTTGAACAGC CCATAGTGACGCACCA 301 A2051303 R:22011.R CACGTGACTTGAACCAGC CCATAGTGACGCACCACA 301 A2051303 R:2301.S ATTAGCGACCACTGAATGAC TACGCTGACTGCCACCCCCACGTGATGACCACCACACACA | R-990A23.R | CTGGATGGAACGCTCACTA | CAGAAGCCTCAGGGTAAGA | 268 | AQ668383 |
| R-316810.F CACAACTATCACCACTTGAC CGTGAGCTICTCCCT 299 AQ541175 R-316810.A ACAGCTARGGATACTGACGCA CTATGGATAGCACACT 288 AQ597714 R-312814.R GGTGAGAATCTGACGATACTGACGGCA CTATGGATAGCAGACTCTACCAC 256 AQ539167 R-32814.R ACAACACTGACACTCTAGCAG CCCTGGGTTGACACCACTA 260 AQ285023 R-2011.R CACTGCTGTGTGACAGGC CCCTATGACGCACTA 260 AQ285023 R-203010.R CACGTGCTGTGTACACGGC CCCTATGCTGTGACGCAG 30 AC051300 C-52015.F AGGATTCTGCTCTGACG TAACCCACACTGTGCCCCAC 28 AQ3222538 R-2113.R TATAGCGACACACGCATCTGAC GCCAACGCACGCACGACGACGACGACGACGACGACGACGA | R-1081G2.F | GATAAAGTTTAAGGTTATCTC | GAAATAAGCTTAGCCTGTAC | 275 | AQ740126 |
| R-316B10,R ACACCTAAGCATIAATGAGCCA CTGTTTATCCCCCAAGCACT 288 AQS07714 R-3281-R,R CGTGACAMTCTGACCAAGCTA CTATGGATGAGACTCTCACTG 256 AQS39167 R-959M3,R ACACACCTGACAGTCTATCCG CCCAGATGAGACTCTCACTAGCACA 212 AQ567344 R-202010,R ACGCTGATGAGCAGCTACATGACCCACA 212 AQ567344 R-202010,R CACGTGCGTCTTGACCAGC CCCATAGTGAGCTCACA 321 AQ285023 R-202010,R CACGTGCGTCTGACCAGC CCCATAGTGAGCCACA 301 AQ285023 C-32015,F AGGATTTCTGCTCTCTGAGCACCTACA TTAGGGAGCAACAGTGCC TTAGGGAGCAACAGTGAC 301 A2051301 R-2313,R TGTAGCGACTCCTACTGC TGTAGCGACTCTCTGCGG 248 AQ032238 R-2415,R TGGTAGCGACTCGCATCTCCTTC CTCTCGACAGTGTTGTCGG 248 AQ03238 R-2415,R TGGTAGCGACTGCATGTCCCTTC CTCTACAGAGCTCAGCAGTGTTATCCGG 273 A2051304 R-2415,R TGGTAGCGACTGCACTGCCTCTC CTCACCTAAAGGCGTCAGAGC A2051304 C-32051305 R-35382,F AGGAAGTGTGATAGCGCACA CCTTACCTAAGAGCCTCAGAG 274 A2051304 | R-316B10.F | CACAACTATGACCACTTTGAC | CGTGAGGTTCTCCTTTCCC | 299 | AQ541175 |
| R+22814.R CATAGACATTICACCAACCTA CTATGAACTICAACTICA 256 AQ359167 R>595M3.R ACAACCTACCACCTTATCACG CCCTAGACATCACATCACTACTACTACCAG 372 AQ5166454 R>22C11.R CACGTCGCTGTGTGACGACG CCCTAGATCACACACCACTA 260 AQ285023 R>203101.R CACGTCGCTTGTGACGCAGG CCCTATACTACACACCACCACTA 240 AQ418643 C5201.S.F ATCTCCTCTATAGGCCTCTATG CACGTCACACGACACTACTACCAC 214 AQ213130 R>2113.R TGTGACGCACTCAACAGTACC GACGTACTACAGGCACTCCAC 288 AQ2225328 R>2113.R TGTGACGCAACTCAACAGCACCC GACGTACATGGCACTCCAC 288 AQ2225328 R>2113.R TGTGACGCAACTCAACGCACCC GACGTACTACTGGCACTCCAC 288 AQ2225328 R>2113.R TGTGGCAACTCAACGCACAC CTCCACGACGCACTACTGCCCCCAC 283 AQ2013393 R>2113.R TGTGGCAACTCAACGCACAC CTCTCACGACGCACTCCCCCCCAC 283 AQ225320 R>2113.R TGTGGCAACTCAACGCCACAC CTGTGCGCATCCCCCCAC 284 AQ331303 R>1988C01.L CHACCTAAACAGCCATCCCCACC TTGTGCGCTGCTACACGCACAC 270 <td< td=""><td>R-316B10.R</td><td>ACAGCTAAGGATAATGAGGCA</td><td>GTGTTTATCCCCCAAGCACT</td><td>288</td><td>AQ507714</td></td<> | R-316B10.R | ACAGCTAAGGATAATGAGGCA | GTGTTTATCCCCCAAGCACT | 288 | AQ507714 |
| R*95793.R ACARCA, IGA, AGI, CIALCEG CCAGA INGUINZIA 215 AQ867844 R*46H10.R ACGTATATGAGCGATICTCAAG TCGTAGCACACA 372 AQ516454 R*22011.R CACTEGTICATICTAAGC CCCATATAGAGCATICAACCACA 372 AQ516454 R*22010.F. ACAGTITTCTGCTCTATAGG ACTGATITCTIGCAACACCACAAT 260 AQ285023 R*2113.R ATTGCTCTCTATGGG ACTGATICTTIGAGCAGCACAAT 274 AQ505130 R*2113.R TATGGAGCACACAATCTGAG TAACCCTACATGTGCACACA 285 AQ322223 R*2113.R TGGTACTCCCAATCTCCTTG GTCAAGCACCAACATGTCCCCAA 207 AQ503398 R*2413.R TGGTACTCCACAGTACTCACAG AAGACGCACACATGTCCCCCAA 207 AQ503398 R*38609.L CTACCTCACAGCA AAGACGCACACATGTCCCCCAA 207 AQ503398 R*3616.R ACGGAAGTCAACACTGCACAG TAGACGCACACATGTCCCCAA 207 AQ503398 R*3616.R ACGGAAGTCACCCAACA TTGGTACTCCACATG CCTTACAGACGTACCCACA 207 AQ33398 R*3616.R ACGGAAGTCCACCACAA CCTGCAGACGCACACAA CCTGCAGACCACCACAA 200533930 | R-32814.R | GGIGAGAAIICIAGCAAGCIA | CIAIGGAIAGAACICICAGIG | 256 | AQ539167 |
| Resention AddatiaatGaduCaturitaada ICuruCataiadCata 3/2 AddatiaatGaduCaturitaada R22011.R CACGTGGCTGTGGCAGG CCCATATCTACCAGCACTA 260 AQ285023 R203010.R CACGTGCATTGAAGCGCTAATA TIGATCTACGTAAGCACTACTC 214 AQ418643 C52015.F AGGGTTGTGAGCGCCTATAG TCACCCTCACTCGCCTCC 214 AQ013398 R92113.F TGTAGCGCACCAACGCACTCTGAG TACCCTCACAGGCATTCCC 285 AQ322538 R92113.F TGTAGCGATCCGAACGCATCCC GACCTACATGAGGCCCTTCG 285 AQ322538 R92113.F TGTAGCGAACGCAACTCTACA CACCTACAGGATTCTCCC 233 AQ013398 R98869.L CTACCTAAAAGCCTCACGCAG TTAAGGAGCCTTCGCATCCC 234 AQ0376733 R1516.R AAGGAAAGTCAAACTCCACAGG TTAAAGGAGCCTTCGCATACTCCCAGAG 232 AZ051303 C-35352.F AGGAAAGCCAAAGCCACACAA CTGTGCGCCTCCTCGAGA 244 AZ051303 C-35352.F AGGAAAGCCAACACCACACAA CTGTGCGCCCCCCGAATTCTCCAGAA 2441A27 TCATGCGCATCATGCACACT 24051306 C-2994M11.R GGAAACTGGAAAGCCAACACA CTGTGGCCCCCCCGCAATTCTCCCAC | R-959M3.R | ACAACACIGACAGICIAICCG | CCAGAIAGAIAGIAIGAGIICA | 215 | AQ667844 |
| Int 2011.1K CHARLEGATE CELARACIDATECACCA 200 Ad283022 R220010.E CACCTECATICANGCCTAATA TIGATECACTAC 214 Ad2188023 C-52015.F AGAGTITCTGCTCTATGG ACTGATICTIGAGCCAC 301 A2051300 C-52015.F ACAGTITCTGCTCTATGG ACTGATICTIGAGCCAC 224 A2051301 R-22113.F CCCTAACACCCAATCTCACT GCCAACAGTACCACCACACT CGACCTACACACCTCACT 285 AQ322232 R-2113.F TIGGAGCACCAACAGTACC CGACCAACAGTACTCACCACAC AQ0126226 R2413.R TGGTACTCGACAGTACT CCTTAACAGCAGTCACCACACAC AQ026226 R-2113.F CTGCAACGACACCACACACAC CTCAACGACAGTCTCCCCTC 235 AQ36733 R-16116.R ACTGAACTCACCCACACA CTGTCACCACAGTGTACTCACAC AQ04CGCACACACACCCTTC 235 AQ36733 C-35352.2 ACTGTAACTCACACCACACA CCTGTCACACACAGTCTCCACCTC 235 AQ36733 235 AQ36733 C-35352.2 ACTGCAACTCACACACA CCTGCAACCCACACA CCTGCAACCCACACA CCTGCAACCCACACA AQ363308 24414.7.F TATTTGCCACTACTACGACACA CCTGCAACCCACACA CCTGCAACCCCAC | R-646H1U.K | AGGATAATGAGGCATGTGAAG | | 3/2 | AQ516454 |
| R205010.R CACUTICATICTACIC 214 AQ216843 S2015.F ACGUTICTACICTATGG ACGUTICACCC 214 AQ21631300 C-52015.F ATTCCCCTCAGGCCTCTAG TCACGCCTCTGCCCTTGCCCCTTGCAC 288 AQ322233 R-211.3.F GCTACACACCCAATCTGAG TACCCTACATGATGCCC QACCTACATGATGCCC 243 AQ013398 R-211.3.F TATACGCATCCCAACAGTACC GACCTACATGATGCCCC 243 AQ013398 R-211.3.F TGTACTGCACTCCTTGCTATC CTCACCAACGCTACATGCTCCTGCTGCTATC 233 AQ252238 R-211.8.F TGTGCCACTACTGACAC ACCGCCACATCTTC 233 AQ051303 R-235051.F AGGACATCTTACAGCCTCACAG CTCTCCCAACATCCTCACAGC CCTCACACTCCTCA 224 AQ051304 C-33582.F AGGACATTCTTACAGCCTACA CCTTACATACACCCCA 224 AQ051304 C-33582.F AGGACATTCTACAGCCTACA CCTTACATACACCCCA 224 AQ051304 C-33582.F AGGACATTCTACAGCCTACA CCATACATCCAACACCCCCCCCCCCCCCCCCCCCCCCC | R-/2CII.K | | | 260 | AQ285023 |
| C-2013.F. ADAGMITICING CETERAG ADECRIFICATION ADECRIFICATION R-22113.F. GCCTAACAGCCCAATCTGAG TTAACCCTACATGTTTCCA 285 AQ322538 R-22113.F. GCCTAACAGCCCAATCTGAG TAACCCTACATGTTCTCCA 285 AQ322538 R-22113.F. TGCTACTGCCATCTCCTTG GTCAACGAGTCCTCCA 286 AQ322538 R-2413.F. TGGTACTGCCATCTCCTG GTCAACGAGTCTTCTGCG 243 AQ013398 R-2413.F. TGGTACTGCCATCTGCTGGTTATC CTCTCCAACGATTTCTCCGC AQ013983 R-16116.F. AACTCAAAGCGTACCACAG AACCGCACACACTTCTCCACGC A2051303 C-33582.F. AGGGAAAGTCAAGCCCACAA CTTGTGCTGTCATCTCACACT CCTACATAGTTCCAACGCAA 244 A2051305 C-441A7.F. TATTGCTCATTTTAATGAGCCTA CCTTACATAGTTCCAACGCAA 244 A2051305 C-209M11.R. GGAACTTCTACAGCACAAA CTGTGCCCTCACCT 177 AQ56340 C-2117A16.F. TGTGTGCATCGCAACTAGTGCACTA CCTTACATAGTTCTCACGCT 178 A2051305 C-209M11.R. GGAACAGTGTGAACCTAGGCATCA CCTTACATAGTGCCAACT CGTACAGTGCAACTGCAACTAGTGCATGT 774 A205130 | R-203010.K | | | 214 | AQ416045 |
| C350L3A ATTCLCCTACCAGCCATCTAGG TACCCCTACCAGCCATCTAGG TACCCCTACCAGCCATCTAGG TACCCCTACCAGCCATCTAGG TACCCCTACCAGCCATCTAGG TACCCCTACCAGCCATCTAGG TACCCCTACCAGCCATCTAGG TACCCCTACCAGCCATCTAGGG TACCCTACCAGCCATCTAGGGGGG TACCCTACCAGCCATCTCAGGGGGGGGGGGGGGGGGGGG | C-520L3.F | | | 274 | AZ031300 AZ051201 |
| No.2113.h OLCUMENCICATION PRACTION | D 02113 E | | | 274 | AC322533 |
| R.2415.8. TGCTACTGCATCTCCTTG GTCAAGGAGTCTTCTTGGG 243 AQ013398 Y88809.1 CTACCTAAAGGCTTGGTTATC CTCTGCAAGAGTCTTCTGGG 243 AQ013398 Y88809.1 CTACCTAAAGGCTTGGTATC CTCTGCAAGAGTCTACCAG AQ013398 AQ013398 Y88809.1 CTACCTAAAGGCTGCAAGC CTCTGCAAGAGTCTACCAG AQ051304 AQ051394 C-3535K2.F AGGAAGTCAAGCTCCAAGC TTTAAAGAAGACCTTCCATGC 233 A2051304 C-3355K2.F AGGACATTCTTACAGCTACA CACCTAGTTACTGCAACT 244 A2051304 C-3355K2.F AGGACATTCTTACAGCTACA CACCTAGTTACCAACACTCA 233 A2051304 C-3355K2.F AGGACATTCTTACAGCTACA CACCTAGTTACTACAGGCT 234 A2051304 C-3355K2.F AGGACATTCTTACAGCTACA CACCTACTACTAGCAGCACCTACA CCTTACATAGTTCCAACTCACACA CACCTTACTAGAGCACACTACTACACACACACTCACACA CACCTACTACTACACACACTCACACACACTACTACACACACACTCTACACACACACTCTACACACTCACACACACTCACACACACTCACACACACTCACTACACACACACTCTACACACACTCACTACACACACACTCTACACACACTCACTACACACCTACTCACACACACCTACTCCTACACTACACACCCCCC | R-92L13.1 R-92L13.R | ΤΑΤΑΓΓΓΑΑΓΑΓΙΤΑΙΟ | CACCTACATCATCCCCTTCA | 288 | AQ322533 |
| No.8 CHACCHAAAGGCTTGGTTATC CITCTGCAGACTATTGCCAC 202 AQ051302 R-161J6.R AACTGAAGTGGTAGTATGCAAG AAGAGGGCAGATGTTGCCATG 185 AQ376733 R-161J6.R AACTGAAGTGGTACTCAAG AAGAGGGCAGATGTTACGAAA 244 A2051303 C-3535K2.F ACGGACATGTTACAGGCTAC CAGCTATGTTGCAACTGCCACC 232 A2051305 C-3325015.F AGGGACATTGTTACAGGCTAC CAGCTATGTTCGAACTGCCCC 224 A2051305 C-441A7.F TATTGCTCATTTAATGAGCCTAC CAGCTATGTTCGAACTGCTACC 232 A2051305 C-2094M11.R GGGAGCAAGTGTGCAAGGCA CTTACAAACGGGCTCCCCAATGTC 224 A2051307 C-2117A16.F TAGCAACTGTGCAAGTGCATGC TAACAACAGGCTCCGCACT 38 AQ753530 R-117H6.F TCCTGGCACTGCCATGC TTCAGTACACGGCCCGCCT 38 AQ051308 R2C1 GAACAGTGTGAAGGGTT TTCAGTACACTGCGCGATGC 28 AQ051309 4282 ATGGAAGGGGTTAGGTGGTTG TTCAGTACACTGCGCGATGC 28 A2051308 4287 GCACACTTGTGAGTGGAGTTGT TCAGTACAGTGTGGCTGTA 163 A2051314 <td< td=""><td>R-2415 R</td><td>TGGTACTCCCATCTCCTTG</td><td>GTCAAGGAGTCTTCTTCGG</td><td>200</td><td>AQ013398</td></td<> | R-2415 R | TGGTACTCCCATCTCCTTG | GTCAAGGAGTCTTCTTCGG | 200 | AQ013398 |
| No.1616.R AACTEGACTGGTACTGACAG AACACCGCGACATCTATGCTC 185 AQ36733 C-535K2.F AGGAAAGTCAAGCTGCAGAG TTAAAGAGAGCGTTCACTGC 233 A2051303 C-535K2.R CTTGTGCCATCTACTGCACTC TTGTGCTGTCATGTCCATCGA 244 A2051303 C-325015.F AGGACATTCTTACAGGCTACA CACCTTACTGACGCTCA 253 A2051303 C-4411A7.R TGTTGCGTGCACTACAGCCTA CCTTACTAGAGCTCA 254 A2051307 C-2099MI11.R GGAAACTGTGCAAGTGCACAA TTAACAACGGCACTCGCCC 318 AQ353330 C-2117A16.R TTGGCACTGCACCTAGCTAG TTAACAACGGCATTCCCACC 318 AQ353308 C-2117A16.R TGGCACCAAGCCTGTGACGT TTAACAACGGCATTGCACC 318 AQ353308 422c1 GAACAGACTGTGACGCTTAGCGGT TTAACGACGCTTGCCGCTGTA 163 A2051309 4282 ATGGAAGAGGCTTGAGGTGTGC GCAGCGACAGGTTGCATGACGTA 151 A2051310 4287 GCTGACAATATCCCTTTGACGGTTG GCAGCGCAGAGTTGCATGTAC 151 A2051316 4287 GCTGACAATATCCCTTTACTTC GAAGTGTGCTGACT 174 A2051315 <t< td=""><td>Y888G91</td><td>CTACCTAAAAGGCTTGGTTATC</td><td>CTCTGCAGAGTATTGCCACA</td><td>207</td><td>A7051302</td></t<> | Y888G91 | CTACCTAAAAGGCTTGGTTATC | CTCTGCAGAGTATTGCCACA | 207 | A7051302 |
| C-333K27F AGGAAAGTCAAGCTCCAAGC TTAAAGACAGCCTTCCAGCTTC 233 A2051303 C-335K27F AGGAAAGTCAAGCTCCAAGC TTGAGCACCTTCCTGCAGCTC TTGTGCTGCTCCATTCCATGCAA A2051304 C-335K27F AGGACATTCTTAAGGCTACA CAGCTAGTTACGAGCTC TTGTGCTCCATTCCATGCAA A2051304 C-325015.F AGGACATTCTTAAGGCAAAA CCTTACATGCTCCCAATTCC 224 A2051306 C-441A7.F TATTGGCGGAAAGCCCACAAA CTGTGCCCTCCCCAATTTC 306 A2051307 C-2094M11.R GGAAACTGTGCAAGTCA TTAACAGCAGCACCACA TTAACAGCACCGCT 177 AQ566340 C-2117A16.F TCGTGGCACTGCCAATCTC TTAACAGCACTCCGGATTCG 238 A2051308 R-117H6.F TCTGTGGCACTCACCTAGTAG TGAACAGCACCGACA TCGCGACAACTCCGGATTCC 288 A2051309 4221 AGGAACAGCTGCAACTTC TTAACAGCACCCGGATGCTGGATGTTT 188 A2051304 4225 ATGGAACAGCGTTCAACTGCATTCC CACTGCAGACTTCGCCGTGT 18 A2051314 4257 TGAGACCTTCTGTATCC CACTGGAGATGGCTGTG 117 A2051315 42617 TGAGACCTTCTGTTCTATTCC CACTGGAGAGGTG | R-16116.R | AACTGAAGTGGTACTGACAG | AAGACGGCAGATGTATCCTG | 185 | AO376733 |
| C-533K2.R CTIGTECCACTACTECACTC TTGTGCTGTCCATCACAA 244 A2051304 C-325015.F AGGACATTCTIAAGGCTACA CAGCTAGTATCTGAAACTGCA 253 A2051304 C-441A7.F TATTGCTCATTAAGGCTACA CCAGCATATCTGCAACTGC 224 A2051306 C-441A7.R TGATGGCGGAAAGCCCACAAA CTGTGGCCTCCCCAATTC 306 A2051307 C-2094MI1.R GGAACTGTGCACAGGAGAGA TTACAAACGGAGTCCCGCC 318 AQ753530 C-2117A16.R TAGCAAACTGTGTGCCAATTC TAAAAGGCCACGCAGATTTCTGAC 257 AQ350077 C-2894MI1.R CGCACTACTGTGCAAACTTC ATAACAACGCCAGCCAGATTTCTGA 268 A2051308 42C1 GAACAACATTGTGAAGGGT TTCAGTAACAGCCTGTGTA 163 A2051309 42E2 ATGGAACACATTGCTTACCC CATTGCAACGGCTGTA 163 A2051310 42E7 TGGAACACATTGCTTATCCC CATTGCAAGGCTGTGTA 151 A2051315 4512 ACTGCTGGTTTGGTAAT ACGCACTGCACTGCACTCCAC 158 A2051317 4541 ACTGCTGGTTTGGTAAT ACGCTCTGTGTACTGTACTCT 171 A2051318 45257 < | C-535K2.F | AGGAAAGTCAAGCTCCAGAG | TTAAAGAGAGAGCCTTCAGCTTC | 233 | AZ051303 |
| C-325015.F. AGGACATTCTTACAGGCTACA CACCTAGTTATCTGAAGCTGA 253 A2051305 C-441A7.F. TATTTGCTCATTTAATGAGCCTA CCTTACATAGTTCTACCTCC 224 A2051307 C-441A7.R TGATGGGGGAAAGCCCACAAA CTGTGCGCCTCCCAAATTTC 306 A2051307 C-2094M11.R GGAAACTGTGCCAATC TATAAAGACGAGGTCCTGCCC 1177 AQ566340 C-2117A16.F. TCGTGGCAATCTGCCATTC TATAAAGGAAGGTCCTGCCC 318 AQ753530 R-117H6.F. TCTGTGGACTGCCATTC TATAAAGGAAGGTCCTGCCC 208 A2051308 42C1 GAACAAGATCTGTAAGGGGT TTCAGTAACAATGCCATCGCGATTGC 208 A2051308 42C1 GAACAAGATCTGTAAGGGGT TTCAGTAACATTGCATATTTTCT 138 A2051309 42B2 ATGGAAAGGGTTGGATGTC CACTGCGCATCGCGATTGC 208 A2051308 42E7 TGACAACGTTGCTATATC CACTGCGCATTCC CACTGGCGGTGG 314 A2051314 42B7 GCTGAGAATTCGCTTCATCC CACTGCGAGGTTGGCTGTG 314 A2051314 42B7 GCTGAGAATTCGCTTGTATCC CACTGGAGGCTGGGTGT 151 A2051315 42512 ACTTGCTTGGTTAAT ACCGTTGCCC GATTGCACCCGG 158 A2051316 42F7 GAAGTGGCTAAGGGCTAAGGAGCTTAAT GTGAGCCACTGCGCCGG 158 A2051317 87A10 GGTTCGTTTTGTGTAAT GTGAGCCCCCGG 158 A2051317 87A10 GGTTCGTTGTTATGGC CACCTGGCACTGGCCCGG 158 A2051318 45D2 CACTGCCTGAGAAGGCT GCCCCCTGGGATGGAGAGATGCT 221 A2051318 45D2 CACTGCTGAGATGTGC CCCCATGGAGTGAGATGCCC 21 A2051318 42D2 CACTGCTGAGAGGCAACGG AGGAGGATAAT GTGAGCCACTGCACAGATGGTTC 221 A2051318 42D2 CACTGCTGAGATGCC TGCCACTGGACTGCAC 215 A2051319 42D12 CTTACACATTGTAGACTGC TGCACTGCAGGAGTGAGTGCACA 215 A2051329 42D12 CTTACACATTGTAGAGCGC AGGAGTAATGGAGCGCCCT 196 A2051332 42S9 GCATTCGTGAGCGCC 45B11 GAACGTGGAATGCC TATCTCAGGAGAGGTGACAC 214 A2051326 45B11 GAACGTGGAATATATAAGAGCC TATCTCAGGAGCAC 118 1A2051326 45B11 GAACGTGGAATATATAAGAGCC TATCTCAGGAGAGAGTGCACA 214 A2051328 7A5 TGCAGGTTATGGAGGCCCACTTC GCTCCTGAGAGAAGTGTCACA 215 A2051332 7A5 TGCAGGCTTATGGAGGCC AAGGGGCAAGACTTCAGGGCCC 219 A2051332 7A5 TGCAGCGTATGGAGCGCCACTC GCTCTCAGAGAGCTCCCCC 234 A2051333 87A3 CCCCAGAGCATAGG GCTCCAACCACCGCC 178 A2051333 87A3 CCCCAGAGCATAGC GCTTCCACACCACGCC 209 A2051333 87A3 CCCCAGAGCTTAGCG ACTGCC ACTGCACACCGCC 209 A2051333 87A3 CCCCAGAGCTCATGC GCCTCCAACGCCC 209 A2051333 87A3 CCCCAGCGTAGCCCACTC GCTTCCAACGCCC 209 A2051334 42E12 CCAGATTGCGCCATGCG GCTTCCAACGCCC 209 A2051334 42E1 | C-535K2.R | CTTGTGCCACTACTGCACTC | TTGTGCTGTCCATTCCTAGAA | 244 | AZ051304 |
| C-411A7.F TATTTGCTCATTTAATGAGCCTA CCTTACATAGTTCATCCTC 224 A2051306 C-4411A7.R TGATGGGGAAAGCCCACAAA CTGTGGCCTCCCAAATTTC 306 A2051307 C-2094M11.R GGAAACTGTGCAAGTGAAGA TTAACAACAGGAGTCCCTGACC 318 AQ255353 R-117HoF.R TCTGTGGCTGACCTAGTAG TGGGCGCAACATTTTCGAG 257 AQ350077 C-285H9.F CCCACTACTGTGCAAACTTC ATAGAAGCCCGCGCAACATTTTCTGAG 288 A2051309 42C1 GAACAGAGCTGTGACATGCT ATAGAAGCCGTGGGAT 163 A2051309 42B2 ATGGAACAGGCTTGCATGC CAACTGGCGGGTGTA 163 A2051314 42B7 TGAGACACTTCGCTTCTATCC CACTGGAGAGTGGATGTTA 151 A2051315 42B7 GCTGAGAATTACACTTTACTGC CACTGGAGAGTGCACTCT 211 A2051315 4512 ACTGCTGACATGTGAAGGAGTTAAT ACGGACAGTCCTTT 217 A2051315 4512 ACTGCTGACTTAGCG CACTGCAACTGCACCCCGG 158 A2051315 4514 ACTGCTGATGTGACATGG CCTGCAGATGTTCC 21 A2051315 4527 GCAGACTTTCAGGAGTGACGC< | C-325O15.F | AGGACATTCTTACAGGCTACA | CAGCTAGTTATCTGAAGCTGA | 253 | AZ051305 |
| C-411A7.R TCATGGGGGAAAGCCACAAA CTCTGGCCCCAAATTTC 306 A2051307 C-2094M11.R GGAAACTGTGCAAGGAAGTGAAGA TTAACAACAGGAGGTCCCGCT 177 AQ566340 C-2117A16.R TAGCAAACCGTGCCAAGTGAAGA TTAACAACAGGAGGTCCTGACC 318 AQ753530 R-117H6.F TCTGTGGACTGACCTAGTAG TGGGCCCACAGATTTC ATAACAGCCATTCC 208 A2051308 2281H9.F CCACATACTGTGCAAATTC ATAACAGCCATTTCCTTGCGTATTTCT 183 A2051308 42E1 GGAACAGGTGTGGATGTTG TGAGTCCGCGTTGTA 163 A2051310 42E7 TGAGACCGTTTGCTTACTCC CACTGGAGAGTGTGATTTCT 172 A2051314 42B7 CCTGGAAATACCATTTACTCC CACTGGCAGACGTGTA 151 A2051314 42B7 CGTGAGAATGTCAAGGTAAGTAGT GTGAACCCCCCCGG 158 A2051314 42512 ACTGCTGATGTGTATGGG CACTGCTGAGGTGGATTAT GTGAACCAGGATACTT 217 A2051318 4252 CACTGCTGAATGTGAACTGC CCCATGGTTTGCAGATACTT 269 A2051313 4253 CACTGCTGAATGTGAACAGC AGGGAGAGGTGGATTTC 269 A2051329 | C-441A7.F | TATTTGCTCATTTAATGAGCCTA | CCTTACATAGTTCTCATCCTC | 224 | AZ051306 |
| C-2094M11.R GGAAACTCTGCGACTGAAGA TTAACAACAGGAGTCCCGCT 177 AQ566340 C-2117A16.R TAGCAAATCCTGTGCCATTC TATAAAGGAAGGTCCTGACC 318 AQ753530 R-117H6.F TCTGTGGACTGCCTAGTAG TGGGCCGAAGATTTCTGAG 257 AQ350077 C-285H9.F CCCACTACTGTGCAAACTTC ATAAGAGCCATCGCGGTTGC 208 A2051309 4282 ATGGAAGAGGTGGATGTTG TGAGAGCGT TTCAGTAACATTGCATATTTTCT 138 A2051310 42 E7 TGAGACAGGTCTGCTTCGCTACTG CAGAGTTGCCGGTGTA 163 A2051310 42 E7 GGAAGGGTTGGATGTTG CAGAGTGCTGC GATTTCCAGGAGTTGGCTGTG 314 A2051314 42 E7 GGAAGTGCTTGCTTCTATCC CACTGGAGGTTGCAACATTGGTA 151 A2051315 45A12 ACTTGCTTGGTTTGCTAAC GGAGTTAACATTGCAACTTGCATCT 217 A2051315 45A12 ACTTGCTTGGTTTGGTAT GGCGCCCCCCGCGCGTGA 158 A2051317 87A10 GGTCGGTGTAAGAGAGTTAAT ACAGTCCTTTGCGACCCAG 158 A2051317 87A10 GGTCGGTGGAAGGGAGTTAAT GGCAGCCCCCCGG 158 A2051317 87A10 GGTCGGTGGAAGGGAGTGAACGC CCCCATGGACAATGGTTC 221 A2051318 45D2 CACTGCTGAATGGAACTGC CCCCATGGACAATGGTTC 262 A2051311 42C5 CACTTCATGTGAACGAGACTGC TGCTATGGACAAATGGATTC 262 A2051313 4584 ACTGCTGAATGGAACTGC TGCTCATGGACGACAATGGTTC 262 A2051313 4584 ACTGCTGAATGGAACTGC TGCTACTAGAGGAGTGCACA 45D11 GTACACATGTGAACTGC TGCTCACGAGAAGTGCTCA 215 A2051323 42C9 GTGTAGCCTATTCAGAAGAGA CGTCTACAGAGCGCT 196 A2051323 4259 GTGTAGCCTATTCAGAAGAGC AGTCATCAGAGCGCAT 181 A2051326 455D11 GTACACATGTGAACTGC TGCTCACAGAAAGAGTTCACA 215 A2051323 45810 ACCAGGATAATAAAACC TATCTCACCACCCCAGCGC 219 A2051323 4582 CATTCTCAGGGAGAAATTAAAACC TATCTCACCACCCCACACTG 187 A2051326 45810 ACCAGGATAATAAAACC TATCTCACCAACCCCCAACTG 187 A2051326 45810 ACCAGAAATGACACCTTC CATCTGCAGCCCAAATGCACC 214 A2051338 87A3 CCTCCAGGAGAACAATGG GGCAAATCCACAGCTGCAA 222 A2051331 87A7 CAGCCCTCAGGAGCAATTC CATCTGCAGCCCAACTG 219 A2051333 87A3 CCTCCAGGAGACATGCC CAATGG GTGCTAGAAGTCGCAGCC 219 A2051338 87A3 CCTCCAGTGGCAATGGG GTGCTAGAAGTCGCCCCCA 173 A2051334 42E2 CCACATTGCGCAATGGG CTACACCT CCAAGGCCCAAGAGT 234 A2051333 87A3 CCTCCAGTGGCAAATGGG CAATCTCCACGAGCC 299 A2051330 87A3 CCTCCAGGGCAAATTGCG CTTCTCCCCCCC 209 A2051330 87A3 CCCCACATAGCGCATGCGCC ACAGGCGTGCACACC 187 A2051334 42E10 CAACGTGGTAACAGCC AAGGGGGAACCTCC 234 A2051334 42E11 CGACCTGGGCAAATT | C-441A7.R | TGATGGGGGAAAGCCACAAA | CTGTGGCCTCCCAAATTTC | 306 | AZ051307 |
| C-2117A16.R TAGCAAATCCTGTGCCATTC TATAAAGGAAGCTCCTGACC 318 AQ733530 R-117H6.F TCGTGGACTGACTGAGCA TGGGCGCAAGATTCGACC 208 A2051306 L225H9.F CCCACTACTGTGCACCTAGTGTG TCAGTAACATTCC ATAAGAGCCATTCCGATTGC 208 A2051309 4221 GGCAACAAGCTTCGTAAGGGT TTCAGTACATTGCATATTTTTCT 138 A2051310 42E2 TGGAACAGGGTTGGATGTGTG TGAGTGCTGCGCGTA 151 A2051316 42E7 TGAGACAGGGTTGGATGTCC GATTGCAAGAGGTTGGATTCC 217 A2051316 42E7 GCAGCAGGAGGAGTTAAT GTGACCACTGCACGCAGG 158 A2051316 42E7 GAAAGTAGGCTAAGAGAGTTAAT GTGACCCATTGCACTCC 21 A2051316 42E7 GAAAGTAGGCTAAGAGAGTTAAT GTGACCCATTGCACACATGCT 158 A2051317 87A10 GGTTCGTTTGATGTGTGACC CCCATGGTTGGACGCAGA 158 A2051313 42E5 CACTGCTGAAGTGGACTGC GGCACATGAGAGGCCCT 196 A2051319 42D12 CTTACACATTGTGAACTGC GGCACATGAGAGCCACA 11 A2051328 42E9 | C-2094M11.R | GGAAACTGTGCAAGTGAAGA | TTAACAACAGGAGTCCCGCT | 177 | AQ566340 |
| R+117H6.F TCTGTGGACTGACCTAGTAG TGGGCGCAAGACTTTCTGAG 257 AQ350077 C-285H9.F CCCACTACTGTGCAAACTTC ATAGGACCATCCGGATTGC 208 AZ051308 42C1 GAACAAGAGTTGTGCAACTTC TTAGAGACATTGCCATATTTTTTT 138 AZ051309 42B2 ATGGAACATTCTGCTTCATTCC CCACTGCGCGTGTA 163 AZ051314 42E7 TGGAACCTTTTCTCTCTTCTATCC CACTGGAGATTGACTTTA 151 AZ051315 42B7 GCTGAGAATATCACTTTACTCC GATTTCCAAGGTTACAATGTGTA 151 AZ051316 42F7 GAAAGTAGGCTAAGAGACTTAAT GTGAGCCATCGACCCCAC 158 AZ051316 42F7 GAAAGTAGTGTAACTGC CCACTGCACCCCAC 158 AZ051317 87A10 GGTTCTGTTTCATGTATGG CAACTAGAATTGGATTCC 221 AZ051313 45D2 CACTGCTGAATGTGAACTGC CCACTGCATGCAATGGTTT 269 A2051313 45B4 ACTGGCAATGTGAACTGC TCCCACAGAGAAACAGGGATTGGATTCACA 215 A2051323 42C9 GTGTAGCCATTATATAAGACGCA AGCTGAGAAACAGAG AGCTGAGAAACAACGAC 181 A2051326 45E 11 </td <td>C-2117A16.R</td> <td>TAGCAAATCCTGTGCCATTC</td> <td>TATAAAGGAAGGTCCTGACC</td> <td>318</td> <td>AQ753530</td> | C-2117A16.R | TAGCAAATCCTGTGCCATTC | TATAAAGGAAGGTCCTGACC | 318 | AQ753530 |
| C-285H9.F. CCCACTACTCTGCGAAACTTC ATAAGAGCCATCCGGATTGC 208 A2051308 42C1 GAACAAGATCTGTGAAAGTTC TICAGTAACATTGCATTTTTCT 138 A2051309 42B2 ATGGAAGAGGGTTGGATGTTG TGAGACCTTTCTCCTTCTATCC CACTGGAAGATGAGTGTG 163 A2051310 42B7 GCTGAGAATATCACTTTACTCC CACTGGAGAGTCAATGAGTTG 151 A2051316 42B7 GCTGAGAATATCACTTACTTCC GATTTCCAAGGTTACAATGCGTT 217 A2051316 42F7 GAAAGTAGGCTAAGAGAGTTAAT GTGAGCCACTGCACCCAG 158 A2051316 42F7 GAAAGTAGGCTAAGAGAGTTAAT GTGAGCCATGCACCCAG 158 A2051318 45D2 CACTGCTGAATGTGCAACTGC CCCATGCATGAGTGAACTC 221 A2051313 42C5 CACTTCATCATGTAGAACTGC TCTCACGAGAAAGGGTTGACAC 156 A2051313 42D12 CTTACACATTGTTAGAACTGC TCTCACGAGAAAGAGTGCACA 215 A2051326 45D11 GAACGTGGAATATATAAGAGG GACAATCTTCACGAGCCACA 181 A2051326 45B10 ACCACCTTTAGGAACAGG GAAATACTCAGAAAACC 214 A2051333 | R-117H6.F | TCTGTGGACTGACCTAGTAG | TGGGCGCAAGATTTCTGAG | 257 | AQ350077 |
| 42C1GAACAAGATCTCTAAGGGGTTTCAGTAACATTGCATATTTTCT138A205130042B2ATGGAAGAGGGTTGGTCTTCTGAGATGCTTCGCGGTCTA163A205131042 E7TGGAGACTTTCTCCTTCTATCCCACTGGAGAGTTGGCTGTG314A205131442B7CCTGACGATTATCACTTTACTCCGATTTCCAAGGTTACAATGTGTA151A205131642F7GAAAGTAGGCTAAGAGAGTTAATGTGGAGCCAAGGTTACAATGTGTA151A205131642F7GAAAGTAGGCTAAGAGAGTTAATGTGAGCCACCTCCACCCAG158A205131787A10GGTTCTGTTTCATGTGTATGGCAACTAGAATGAGTTC221A205131845D2CACTGCTGAATGGAACTGCCCCATGGTTTGACAATGGTTT269A205131942D5CACTTCATGATGGAACTGCCGCTACTATGGAGCCCT196A205131942D12CTTACACATTGTTATGAAGTGCACGTCTCAGAGAAAAGATGTCACA215A205132242C9GTGTAGCCTATTGAGAGGGACGTCTCAGAGAAAAGATGCACAG181A205132342C9GTGTAGCCTATTGAGAGGGGAAATGTCTCAGCAACCTG187A205132445B1GTCAGCATAATATAAAGACCTATCTCACCACCCACACTG187A205132645B2CATTCTGAGGGCCAGATTTAGATGTAATATAGGCC219A205132945B1GTCAACATGATAGAAGAGGGGCAAAATGCCC219A205133042C8GCCAAAATTGACAAGGGTCCTACAGAACACAGGAT300A205133387A3CCTCACGGGCAGAATGCAGTCTACACTGACTGCAGCAT300A205133387A3CCTCACAGGAGAGCCACATGGCGACTCTCTGAGAGCAGCCC178A205133987A4CCAGACATACGGGACAGAGGCGACCACCACTCCACACTGCCACA </td <td>C-285H9.F</td> <td>CCCACTACTGTGCAAACTTC</td> <td>ATAAGAGCCATCCGGATTGC</td> <td>208</td> <td>AZ051308</td> | C-285H9.F | CCCACTACTGTGCAAACTTC | ATAAGAGCCATCCGGATTGC | 208 | AZ051308 |
| 4282ATGGAAGAGGGTTGGATGTGTGAGTGTCTGCCGGTGTA163A205131042 E7TGAGACGTTTGCTGTCATCCCACTGGAGAGTTGGCTGTG314AZ0513144287GCTGAGAAATATCACTTTACTCCGATTTCCAAGGTTACAATGTGTA151AZ05131545A12ACTTGCTGGTTTGGTATTACAGTCTCTTTGTGATCTT217AZ05131647F7GAAAGTAGGCTAAGAGAGTTAATGTGAGCCACTCGCACCCAG158AZ05131787A10GGTTCGTTGTATGGACTGCCCCATGGTTGAACTC221AZ05131845D2CACTCGTGATGTGAACTGCCCCATGGTTGAAGAGGGATTGGTTT269AZ0513134584ACTGCTGATGTGAACAGGAGGATGATAGAGGGATTGGTAT269AZ05131342D12CTTACACATTGTTATGAAGTGCACGTCTCAGAGAAGAGTGTCAC215AZ05132342C9GTGTAGCCTATTCAAGAGGACAGTCACTGTCTCAGAGAAC214AZ05132645D11GAACGTGGAATAATATAAGACCTATCTCACCACCACCACACAG181AZ05132645B10ACCAGAAAATGGCAACGTTCCATGCTGACAGTGGCA219AZ05133187A7CAGCCCTAAGAGACACTTCGCCTAACTGAGTCAGAGTGCCACG219AZ05133287A5TGAACGTGAGAGAGCACTCCGCCTAACATGAGTGCACG22AZ05133287A5TGAACGCTGAGAGAGAGGGTCCACATGCATCAGGGG257AZ05133487A2CACATATCTCGAGAGAGGGTTCCACATCACTGCACAGAGT300AZ05133587A3CCCACAGGTGAGAGGGTTCCACATCACTGGTCCAAGG178AZ05133587A4CCACACGTGTAGGGGTTCCACATCACTGGTCCAAGG178AZ05133787A5TGAACGGGTGAGAGAGGTTCACACTGCCCAAGGG178 </td <td>42C1</td> <td>GAACAAGATCTGTAAGGGGT</td> <td>TTCAGTAACATTGCATATTTTTCT</td> <td>138</td> <td>AZ051309</td> | 42C1 | GAACAAGATCTGTAAGGGGT | TTCAGTAACATTGCATATTTTTCT | 138 | AZ051309 |
| 42 E7TGAGACCTITICTGCTICTAICCCCACTGGAGAGTTGGCTGTG314A20513154287GCTGAGAATATCACTITTACTCCGATTTCCAAGGCTTACAATGGTAT151A205131545A12ACTTGCTTGGTTTTGGTAATACAGTCTCTTTGTGATCTT217A205131642F7GAAAGTAGGCTAAGAGAGTTAATGTGAGCCACTGCACCCAG158A205131787A10GGTCTCTTTTCATCTGTAGGCAACTAGAATGGACTACATCC221A205131845D2CACTGCTGAATGTGAACTGCCCCATGGTTTGACAAAGGATTGCTT269A205131342E5CACTTCATCATGTAGAACAGGAGGATGATGACAGCGCCT196A205132342D12CTTACACATTGTAAGACGCGTCTCAGAGAAAGATGCACAC215A205132342D9GTGTAGCCTATTCAAGAGAACAGTCAGTCTCACCGCCCACCACTG181A205132645D11GAAGGTGGAATAATATAAGACCTATCTCACCACCCCACACTG187A205132745E11GTCAGCTTTAGGATAGGGGAAATACTCAGAAACCACACAC214A205132845E2CATTCTGAGGCCCAGATTTAGAGTGAGACCACTTCCATAGTATACTTGCAAGCC219A20513284582CATTCTGAGGCCCAGATTTAGAGTGAGACAATGAGATCGTGCCAAGCCAACCACT212A205133245810ACCAGAAAATGCACCCTTCCATAGTATACTTGCAAGCC219A205133187A7CACCGCTCAGAGCACAATAGATCGTGCCAAGCCATCACGAGT300A205133387A3CCTCAGTGGAGAGCTCCATCGTCTCACAGCCAGGCGTACAGAT300A205133387A3CCTCAGTGGAGAGATGCCAGATTCACACTGAGGC29A205133742E10CAAGACGTGGAGAGCTACGCAAATTTCCTAGGAGCCC209A2051334 <t< td=""><td>42B2</td><td>ATGGAAGAGGGTTGGATGTTG</td><td>TGAGTGTCTGCCGGTGTA</td><td>163</td><td>AZ051310</td></t<> | 42B2 | ATGGAAGAGGGTTGGATGTTG | TGAGTGTCTGCCGGTGTA | 163 | AZ051310 |
| 4267GCIGAGAAIAICACTITACTICCGAINICCAAGGITIACAAIGTGIA151A205131645A12ACTIGCITIGGITTIGGITATACAGTCTICTITGIGATCTT217A205131642F7GAAAGTAGGCTAAGAGAGATAATGTGAGCCACTGCACCCAG158A205131787A10GGTICTGTTTCATGTGTATGGCAACTGCATGAGATGGACTAGATACTC221A205131845D2CACTGCTGAATGTGAACTGCCCCCATGGTTTGACAAAGTGTTC262A205131142C5CACTGCTGAATGTGAACTGCTGCTACAGAAGGGGATTGGTTT269A205132345B4ACTGCTGAATGTGAACTGCTGCTACAGAAGAAGATGTCACA215A205132642D12CTTACACATTGTATGAAGTGCACGTCTCAGAGAAGAAGTCTCACA215A205132645D11GAACGTGAATATATAAGAGCTATCTCACCACCACACTG181A205132645D11GAACGTGAATATATATAAGAGCTATCTCACCACCACACTG187A20513284582CATTCTGAGGCCACAGTTTAGATGTATATATAAGAGC144A20513284582CATTCTGAGGGCCAGATTTAGATGTAATATATAGGG272A205133187A7CAGCGCTCAGAGACAATAGATGCTGCCAAGGCGCACTGCAA222A205133187A5TGACAGGGTGAGAGTCCATCGCTTACACTGACAGATTGCAGA222A205133287A5TGACAGGGTGAGAGATGCAGTTCCACATGACTGCAGAGT300A205133287A8CCTCACTGGAGCAGTTGAGGGTTCCACATGACTGCAGGG257A205133487A2CACAGTGTAGAGGTGCAAGGGTTCCACATGAGGTGCAGGTC153A205133442E10CCAAGAGGTGAGAGCCCCTTTTTCACACTTGCAGAGGTC234A205133742E10CCAACGTGGTAGAGGGTTC | 42 E7 | TGAGACCTITICTGCTTCTATCC | CACTGGAGAGTTGGCTGTG | 314 | AZ051314 |
| 43A12ACTICUTICUTICUTICAT217A205131642F7GAAACTAGGCTAACAGCAGTTAATGTGAGCCACTGGCACCCAG158A205131787A10GGTTCTGTTTCATGTGTATGGCAACTAGAACTGGCACCCAG128A205131845D2CACTTCGTGAACTGCCCCCATGGTTTGACAAATGGTTT269A205131345B4ACTGCTGAATGTGAACTGCTGCTACTATTGCAGCGCATT196A205132942D12CTTACACATTGTTATGAACTGCACGTGCAGCAGATAGAGCGATTGCACA215A205132642C9GTGTAGCTATTCAGAGAACAGTCAGCTCACGGCCA181A205132645D11GAACGTGGAATAATATAAGACCTATCTCACCACCACCACTG187A205132745E 11GTCAGCTTTATGGATAGGGGAAATACTCAGAGAAC214A205132845B2CATTCTGAGGGCCAGATTTAGATGTAATCTCAGAGTCCTC242A20513304268GGCAAAATGACACCTTCCATAGTATGAGTGCTACTTGCAACGC219A205132845B10ACCAGGAAAATGACACCTTCCATAGTAGAGCGATTGCAA222A205133187A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCATTGCACACGATT300A205133287A3CCTCAGTGGAGAGATGCCAAATTCCTAGCAGACGAAT300A205133387A3CCTCAGTGGAGAGTCCATCGCTTACACTCACTGACTGCAC178A205133487A2CACATATCTCTGAGTAGAGCTTCCACATGACTGCACG178A205133442C2/T7CAGACTGTATGAGGTTCCACATGCC209A205133442C2/T7CAGACTGTGTAGAGGGTTCCACATGCCCC209A205133442C2/T7CAGACTGTGTAGGGGTTCCACATGCCCCC209A205133442C2/T7 </td <td>42B/</td> <td></td> <td></td> <td>151</td> <td>AZUSI315</td> | 42B/ | | | 151 | AZUSI315 |
| 427/GAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG | 45A1Z | | | 217 159 | AZUSI316 |
| 07A100GTTCTGTTCAGTGGAACTGCCCCATGGTTGGAAATGCT211AC03131845D2CACTGCTGAATGTGAACTGCCCCATGGTTGGAAATGGTTTC262A205131142C5CACTTCATCATGTAGAACGGAGGATGATAGAGGGGATTGGTTT269A205131345B4ACTGCTGAATGTGAACTGCTGCTACTATTGCAGCCCT196A205132942D12CTTACACATTGTTATGAAGTGCACGTCTCAGAGAAAGAGTGTCACA215A205132645D11GAACGTGGAATAATATAAGACCAGTCAGAGAAAGAGTGTCACA214A205132745E 11GTCAGCTTTATGGAAGGGGAAATACTCAGAAATCCAGAAAC214A205132845B2CATTCTGAGGGCCAGATTAGATGTAAAATACTTGCAAGCC219A205132845B10ACCAGAAAATGACACCTTCCATAGTATGAGTGCTACTTGCAT272A205133042G8GGCAAACTTTGGCTCATGGGTCTAGCAAGCCATCCAA222A205133187A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACGG178A205133387A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGCAGG178A205133487A3CCTCAGTGGAGAGCCATTGGCTATGCTTCGAATGCCAG178A205133541HM9CCCACCACCAGCATGTAGAGGTTCCCACTCGAAGCCC209A205133742F10CAAGCATGTGTAGAGGCTATCCTTCTGCCCTC201A205133742E10CAAGCATGTGTCCCATTGGAGTGCCATTGCAGCAGCAC183A205133442D10CAAGCCTGTGTAGAGGGTTCCACATCACAGGGTC234A205133742E10CAAGCCTGGTGTGCCATGGCGCAAGCCTTCCCATT201A205133742E10CAAGCATGTGTATGGGCTATCCTTGAGTAAAGCCC209< | 4ZF7 97A10 | | | 221 | AZUSISI7 A7051218 |
| 42C5CACITICATCATGTAGAACAGGAGGATGATAGAGGATTGGTTT262AZ05131345B4ACTGCTGAATGTGAACTGCTGCTACTATTGCAGCCCT196AZ05131342D12CTTACACATTGTTATGAACTGCACGTCTCAGAGAAAGATGTCACA215AZ05132342C9GTGTAGCCTATTCAGAGAACAGTCAGTCTTCACGGCCA181AZ05132645D11GAACGTGGAATATATAAGACCTATCTCACCACCCCACACTG187AZ05132745E11GTCAGCTTTATGGATAGGGGAAATACTCAGAAACC214AZ05132845B2CATTCTGAGGGCCAGATTTCATAGTATGAGTGCAGCACAAACC219AZ05133045B10ACCAGAAAATGACACCTTCCATAGTATGAGTGCTACTTGCAT272AZ05133145B10ACCAGAAAATGACACCTTCCATAGTATGAGTGCTACTTGCAT272AZ05133245810ACCAGAAAATGACAACATTGGTGCTAGAATCATCAGTTTGCAT272AZ05133187A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGTGAGAGTCCATCGCCTTACACTGACTGACGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133487A2CACATTATCTCTGAGTAGAGGTTCCACATCACTGGTCATC153AZ05133442C2/77CAGACCATGGTTGCCATTGGGGTTCCACATCACTGGTCATC234AZ05133742E10CAGACGTGGTTGCCATTGGAGTTGCCTTTTCCTCCCCTC201AZ05133442D3AACCACTCTTTACCCAGGGGGAAGCCTTCTTTAAATGGCATA237AZ05134142D3AACCACTCTTAAATGGGCGCAAGCCTTCTCAAAAGGCCA187AZ05134342D1GCAGGGAGAGTTACAGGGCGCA | 45D2 | | | 262 | AZ031318 A7051311 |
| 12:50CENTERTOR ON CONCONCINCE12:50 <td>4205</td> <td>CACTTCATCATCTACAACAGG</td> <td>ACCATCATACACCCATTCCTTT</td> <td>269</td> <td>AZ051311</td> | 4205 | CACTTCATCATCTACAACAGG | ACCATCATACACCCATTCCTTT | 269 | AZ051311 |
| 1351Instruction of control of the control | 45R4 | | TGCTACTATTGCAGCCCT | 196 | AZ051319 |
| 42C9GTGTAGCCTATTCAGAGAACAGTCAGTCTTCACGGCCA181AZ05132645D11GAACGTGGAATAATATAAGACCTATCTCACCACCCACACTG187AZ05132745E 11GTCAGCTTTATGGATAGGGGAAATACTCAGAATCCAGAAAC214AZ05132845B2CATTCTGAGGGCCAGATTTAGATGTAATACTTGCAAGCC219AZ05132945B10ACCAGAAAATGACACCTTCCATAGTATGGTACTTGACTC242AZ05133042G8GGCAAACTTTGGCTCATGGGTGCTAGAATCATCAGTTTGCAT272AZ05133187A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACAGAT300AZ05133387A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGCTGGG257AZ05133487A2CAACATTATCTCTGAGTAGAGCTTATGCTTCAAGGAAAGTTGGG257AZ05133587A3CCTCACGTGAGCAGATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133587A4CCTACATGCTGTATAGAGGTTCCACATCACTGGTCATC153AZ05133587A5TGACAGGGTGAAGACTCCTTTTTCACACTTCAACTGGTCATC153AZ05133487A2CAACATTATCTCTGAGTAGAGGTTCCACATACTGGTCATC153AZ05133587A3CCTCACAGCATATGGGGTTCCACATCACTGGTCATC153AZ05133587A4CCTACATGTGTTATGAGGTTCCACATACTGGTCATC153AZ05134142C2/T7CAAGACTGGTTAGAGGTTCCACATCACTGGTCAC234AZ05133742F10CAAGACTGGTTGCCATATGGGATCATCATGGGGCCC209AZ05134142D8CCTACATGGTTCCCATTGCATTGCCTTGATATACTTGCAAGCCC | 42D12 | CTTACACATTGTTATGAAGTGCAC | GTCTCAGAGAAAGATGTCACA | 215 | AZ051323 |
| 45D11GAACGTGGAATAATATAAGACCTATCTCACCACCCACACTG187AZ05132745E 11GTCAGCTTTATGGATAGGGGAAATACTCAGAATCCAGAAAC214AZ05132845B2CATTCTGAGGGCCAGATTTAGATGTAATACTTGCAAGCC219AZ05132945B10ACCAGAAAATGACACCTTCCATAGTATGAGTGCTACTTGACTC242AZ05133042G8GGCAAACTTTGGCTCATGGGTGCTAGAATCATCAGTTTGTCAT272AZ05133187A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACAGAAT300AZ05133387A3CCTCAGTGGAGAGACCATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133587A42CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133587A2CACATTATCTTGAGTAGAGGTTCCACATCACTGACTGACTGC153AZ05133742F10CCACCAGCAGTGTATGAGGTTCCACATCACTGACCC209AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133742E12CCACAGTGTTGCCATATGGCATCTCTTTCCTCCCCCC201AZ05133942 D8CCTACATGGTTCCCATTGCATTGCCTTGATGAGCCATA237AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGCCAGTA237AZ05134342H11GCACTCCAGCCTGGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTCCCAAAAAGCGCC175AZ05134142A3AGCGGAATCTAGCCACCCAGTCCTCCCAAAAAGCGCC175AZ051312 | 42C9 | GTGTAGCCTATTCAGAGAAC | AGTCAGTCTTCACGGCCA | 181 | AZ051326 |
| 45E 11GTCAGCTTTATGGATAGGGGAAATACTCAGAATCCAGAAAC214AZ05132845B2CATTCTGAGGGCCAGATTTAGATGTAATACTTGCAAGCC219AZ05132945B10ACCAGAAAATGACACCTTCCATAGTATGAGTGCTACTTGACTC242AZ05133042G8GGCAAACTTTGGCTCATGGGTGCTAGAATCATCAGTTTGCAT272AZ05133187A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACAGAT300AZ05133387A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGCTGGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGTCATC153AZ05133742F10CAGACCTGTTGCCATATGGCATCTCTTTCACCACAGGCC209AZ05133942 E2CCAGATTGGCTATGGAGCAGATGTATAGCC209AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGCCAGTCA183AZ05133887A1GATTACAGGCATGACGCCGAGTGTAGCAACCCAGTCTCCAAATGGTC175AZ05133687A1GATTACAGGCATGCGCACCCAGTCCTCCAAAAGCGCCCAGTCCTACAGCAC175AZ05133187A1GATTACAGGCATACGCCCAGTCCTCCAAAATGGTC175AZ05133687A1GATTACAGGCATCACGCCCCAGTCCTCCAAAAAGTGCT175AZ051336 | 45D11 | GAACGTGGAATAATATAAGACC | TATCTCACCACCACACTG | 187 | AZ051327 |
| 45B2CATTCTGAGGGCCAGATTTAGATGTAATACTTGCAAGCC219AZ05132945B10ACCAGAAAATGACACCTTCCATAGTATGAGTGCTACTTGACTC242AZ05133042G8GGCAAACTTTGGCTCATGGGTGCTAGAATCATCAGTTTGTCAT272AZ05133187A7CAGCCCTCAGAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACAGAT300AZ05133387A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGTCATC153AZ05134442C2/T7CAGACTAGCATCCTCTTTTCACACTTGACAGAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCTCACACAGGTC209AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134342H11GCACTCCAGCTGGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCGCCAGTCCACACATACAGGTC175AZ05133687A3ACGCAATCTGCGCTACTAGCGCCAGTCCACACAT210AZ051336 | 45E 11 | GTCAGCTTTATGGATAGGG | GAAATACTCAGAATCCAGAAAC | 214 | AZ051328 |
| 45B10ACCAGAAAATGACACCTTCCATAGTATGAGTGCTACTTGACTC242AZ05133042G8GGCAAACTTTGGCTCATGGGTGCTAGAATCATCAGTTTGTCAT272AZ05133187A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACAGAT300AZ05133387A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGTCATC153AZ05133442C2/T7CAGACCATAGCATCCTCTTTTCACACTTCACACGAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134342B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGCCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAAATGGTC175AZ05133142A3AGCCAATCTGCGCTTACTACGGCCACTCCAACAATGCGC175AZ051312 | 45B2 | CATTCTGAGGGCCAGATTT | AGATGTAATACTTGCAAGCC | 219 | AZ051329 |
| 42G8GGCAAACTTTGGCTCATGGGTGCTAGAATCATCAGTTTGTCAT272AZ05133187A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGTGAGAGTCCATCGCCTTACACTGACTGACAGAT300AZ05133387A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGCAGTC153AZ05134442C2/T7CAGACCATAGCATCCTCTTTTCACACTTCACACAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134387A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAATGGTC175AZ05133687A1AGCCAATCTGCGCTTACTACGGCCACTCCAACATCAGCAC209AZ05133687A1AGCCAATCTGCGCTTACTACGCCCACTCCAACAATGGTC175AZ05133687A1AGCCAATCTGCCATCACCACCCACTCCAACAATCAGCCC209AZ05133687A1AGCCAATCTGCCTTACTACGCCCACTCCACACACT210AZ05133687A1CATTACAGGCATCACCCACCCACTCCACACACT210AZ051312 | 45B10 | ACCAGAAAATGACACCTTC | CATAGTATGAGTGCTACTTGACTC | 242 | AZ051330 |
| 87A7CAGCCCTCAGAGACAATAGATGCTGCCAAGCCATCCAA222AZ05133287A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACAGAT300AZ05133387A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGTCATC153AZ05134442C2/T7CAGACCATAGCATCCTCTTTTCACACTTCACACAAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134387A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAATGGTC175AZ05133687A3AGCCAATCTGGCCTTACTACGGCCAGTCCTCCAAAATGGTC175AZ05133687A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAATGGTC175AZ05133687A1GATTACAGGCATCAGCCACCCAGTCCAACATCAGCACCT210AZ05133687A1GATTACAGGCATCAGCCACCCAGTCCAACATCAGCACCT210AZ051312 | 42G8 | GGCAAACTTTGGCTCATGG | GTGCTAGAATCATCAGTTTGTCAT | 272 | AZ051331 |
| 87A5TGACAGGGTGAGAGTCCATCGCCTTACACTGACTGACAGAT300AZ05133387A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGTCATC153AZ05133442C2/T7CAGACCATAGCATCCTCTTTTCACACTTCACACAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCACCCAGTCCTCCAAAATGGTC175AZ05131242A3AGCCAATCTGCGCTTACTACGGCCACTCCAACATCAGCACT210A7051312 | 87A7 | CAGCCCTCAGAGACAATAGA | TGCTGCCAAGCCATCCAA | 222 | AZ051332 |
| 87A3CCTCAGTGGAGCAGATTGCAAATTTCCTAGGAAAGTTGGG257AZ05133487A2CACATTATCTCTGAGTAGAGCTATGCTTCTGAATGCCAG178AZ05133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGTCATC153AZ05134442C2/T7CAGACCATAGCATCCTCTTTTCACACTTCACACAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGCCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCACCCAGTCCTCCAAAATGGTC175AZ05133642A3AGCCAATCTGCGCTTACTACGGCCACTCAACATCAGCACCT210A7051312 | 87A5 | TGACAGGGTGAGAGTCCATC | GCCTTACACTGACTGACAGAT | 300 | AZ051333 |
| 8/A2CACATIAICTCTGAGTAGAGCIAIGCTTCTGAAIGCCAG1/8A205133541HM9CCCACCAGCAGTGTATGAGGTTCCACATCACTGGTCATC153AZ05134442C2/T7CAGACCATAGCATCCTCTTTTCACACTTCACACAAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGCCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCACCCCAGTCCTCCAAAAAGGTC175AZ05133642A3AGCCAATCTGCGCTTACTACGGCCACTCAACATCAGCACCT210AZ051312 | 87A3 | CCTCAGTGGAGCAGATTGC | AAATTTCCTAGGAAAGTTGGG | 257 | AZ051334 |
| 41HM9CCCACCACCACGCACTGTATGAGGTTCCACACTGCACATCACTGGTATCA153AZ05134442C2/T7CAGACCATAGCATCCTCTTTTCACACTTCACACACAGGTC234AZ05133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGCAAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCACCCAGTCCTCCAAAAGGTC175AZ05133642A3AGGCAATCTGCGCTTACTAGGCCGACTGAACATACAGCACT210AZ051312 | 8/A2 | CACATIAICICIGAGIAGAG | CIAIGCIICIGAAIGCCAG | 1/8 | AZ051335 |
| 42C2/17CAGACCATAGCATACCTCTTTCACACTICACACAAGGTC234A205133742F10CAAGACTGGTTGCCATATGGCATCTCTTTCCTCCCCTC201AZ05133942 E2CCAGATTTGACCAAAAGCCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGCCAC187AZ05134342H11GCACTCCAGCCTGGCAAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTACAAAGGTC175AZ05133642A3AGCCAATCTGCGCTTACTAGGCCGACTGAACATACAGCACCT210AZ051312 | 41HM9 | | | 153 | AZUS1344 |
| 42F10CAAGACTGGTIGCCATAGGCATCLETTCCCCCCC201A205133942 E2CCAGATTTGACCAAAAGCCAGATGTAATACTTGCAAGCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCCAGCCTGGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAAATGGTC175AZ05133642A3AGGCAATCTGGCTTACTAGCGCGACTGAACATACAGACATC210AZ051312 | 42C2/17 | | | 234 | AZUS1337 |
| 42 L2CCACATITICACCAAACCCACATITIAATACTICCAACCC209AZ05134042D8CCTACATGTGTTCCCATTGCATTGCCTTGATGCCTCCCA175AZ05134142D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAAATGGTC175AZ05133642A3AGGCAATCTGGGCTTACTAGCCGACTGAACATACAGACAT210AZ051312 | 42F10 | | | 201 | AZU31339 AZ051340 |
| 42D3AACCACTCTTAACTCCAGGGGCAAGCCTTCTTAAATAGGCATA237AZ05134242B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAAATGGTC175AZ0513364243AGGCAATCTGGGCTTACTAGGCGACTGAACATACAGACAT210AZ051312 | 4208 | CCTACATGTGTGTCCCATTCCA | TTGCCTTGATGCCTCCCA | 175 | Δ7051340 |
| 42B1AGGTGGATGTCTACAATGGTCGGTTTGCATATAGCCAGTCAC187AZ05134342H11GCACTCCAGCCTGGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAAATGGTC175AZ05133642A3AGGCAATCTGGGCTTACTAGGCGACTGAACATACAGACAT210AZ051312 | 4200 | | | 227 | AZUJI 341 AZ051 241 |
| 42H11GCACTCCAGCCTGGGCAAAGAGGTGAAGCTTACTGGAA183AZ05133887A1GATTACAGGCATCAGCCACCCAGTCCTCCAAAAATGGTC175AZ05133642A3AGGCAATCTGGGCTTACTGGGCGACTGAACATACAGACAT210AZ051312 | 42B1 | ΑGGTGGATGTCTACΔΔΤGGTC | GGTTTGCATATAGCCCAGTCAC | 187 | A7051342 |
| 87A1 GATTACAGGCATCAGCCAC CCAGTCCTCCAAAAATGGTC 175 AZ051336 4243 AGCCAATCTGCGTTACTACG CGACTGAACATACAGACAT 210 AZ051312 | 42H11 | GCACTCCAGCCTGGGCAA | AGAGGTGAAGCTTACTGGAA | 183 | AZ051338 |
| | 87A1 | GATTACAGGCATCAGCCAC | CCAGTCCTCCAAAAATGGTC | 175 | AZ051336 |
| | 42A3 | AGGCAATCTGGGTTACTAGG | CGACTGAACATACAGACACT | 210 | AZ051312 |
| 45F3 CAAGTACTGTTCTAAGGGCT TATGATAGAGGTATGCACTGG 168 AZ051320 | 45F3 | CAAGTACTGTTCTAAGGGCT | TATGATAGAGGTATGCACTGG | 168 | AZ051320 |
| 45D2 TGGCCACTATCATTATTAGAAA CTCTTCAGAGAGTTTGGACC 255 AZ051321 | 45D2 | TGGCCACTATCATTATTAGAAA | CTCTTCAGAGAGTTTGGACC | 255 | AZ051321 |

Continued

European Journal of Human Genetics

|--|

| ^a STS | Forward primer $(5' - 3')$ | Reverse primer (5' – 3') | Size (bp) | ^b Accession no. |
|------------------|-----------------------------|--------------------------|-----------|----------------------------|
| 42B3 | AACAGTCAGCTTCTCAAAGG | ATGGAGACTTCTTTAGGAGG | 217 | AZ051325 |
| 45A8 | CATCTTCATCATCAAGCAGTG | AAGTACTGTGCCAAGGCCTG | 240 | AZ051322 |
| D2S4009 | GATCCAGTGTCATTATGCATAC | GCCAGTTGTTAATATTTTGCC | 219 | G64673 |
| D2S4010 | CAGCGGTAGTCTCTATGATA | TCAGAAGGTTCCTTATACAAGGC | 172 | G64671 |
| D2S4014 | TGCAGACTGTAATTGTGGGCT | GACTCCAGATGAGATCTATGACTG | 297 | G64669 |
| D2S4015 | CTCAAATCTCTGACTCCAGATC | GGCTATCCACTCAATAATTC | 297 | G64672 |
| D2S4016 | GATAAGCAAGCTGGTCACACTC | ATTTGAGCTTCAGAGGTCAA | 253 | G64670 |
| D2S4019 | ATGATCTGCATGAGGGTCAAGG | GAGTATATTTAGAAATTTCCATAA | 102 | G64675 |
| D2S4020 | TAGTCTTAATGTTTCCCTTGG | GAGACTAGTTTTCTGACTCAAG | 189 | G64676 |
| D2S4023 | GAGATTCTTTTATTCTGATTTTTTGAG | ATGATCTGCATGAGGGTCAA | 127 | G64677 |
| DESTORS | | | 12/ | 001077 |

Table shows the primer sets for all the unique STS and microsatellite repeats identified in this study and not available in the public databases. ^aPrefix C is CITB-SHP-C BAC library; R is RPCI-11 BAC library; Y is CEPH YAC library; D2S is a microsatellite marker. ^bPrefixes AZ and G are from this study.

and isolate candidate gene(s), we have constructed highresolution YAC and BAC contigs by using known STSs and by generating novel STSs from this region. We have mapped a number of ESTs to this interval, building a partial transcript map that should aid identification of the defective gene. Additionally, this may facilitate the identification of other disease loci mapped to this region, such as a QTL for serum leptin levels,¹⁴ as well as a locus for gingival fibromatosis.¹⁵

Materials and methods

Selection and STS contents of YAC clones

YAC clones were identified through the YAC databases developed by CEPH^{16,17} and the Whitehead Institute¹⁸ using all of the known markers and STSs in sitosterolaemia region (*D2S2291, D2S2174, D2S1830, D2S1485, D2S2298, D2S119, D2S2294, D2S414*). The YAC clones were purchased from Research Genetics, Inc (Birmingham, AL, USA). Single YAC colonies were grown at 30°C for 48 h in 15 ml of selective YPD medium. Total YAC DNA was prepared as described previously.¹⁹ The STS contents of the YACs were determined by using PCR amplifications.

Inter-Alu PCR

Inter-Alu PCR was performed using YAC DNA as template and the following primers: CL1, (5'TCCCAAAGTGCTGGGATTA-CA), CL2 (5'CTGCACTCCAGCCCTGGG) and used as CL2 alone or CL1 and CL2 combined primers.^{20,21} The PCR products were isolated and cloned into plasmid, pBluescript (Stratagene, La Jolla CA, USA) using TA cloning, as previously described,²² and sequenced using T3 and T7 primers. The sequences were scanned against the databases, using BLAST²³ (http://www.ncbi.nlm.nih.gov/BLAST/) and the RepeatMasker program (http://ftp.genome.washington.edu/RM/web RepeatMasker.html). Unique sequences were used to design primers for further mapping (Table 1). Confirmation of chromosome 2 specific sequences was verified by PCR, using chromosome 2 specific humanhamster hybrid somatic cell line DNAs (Corell Cell Repository, Camden, NJ, USA).

BAC clone screening

PCR-based library screening The CITB-SHP-C Human BAC library,²⁴ (Research Genetics, Inc., Huntsville, AL, USA) was screened by a PCR-based assay of DNA superpools and plates according to the vendor recommended procedures. Positive clones were obtained from Research Genetics, Inc., plated on agar plates containing 12.5 μ g/ml chloramphenicol and colonies screened by PCR for STS content verification.

Hybridisation-based library screening High-density gridded filters of BAC libraries (RPCI-11) were obtained from Roswell Park Cancer Institute (Dr. Peter de Jong's laboratory, Buffalo NY, USA), and screened with radioactive probes from the IMAGE cDNA clones of ESTs mapped to the YAC contig. Positive clones were obtained from the Roswell Park Cancer Institute.

Selection from database

All known STS, EST and Alu PCR sequences were checked by a Basic BLAST against the Alu database (http:// www.ncbi.nlm.nih.gov/blast/blast.cgi) and masked by RepeatMasker²³ and unique sequences thus identified were used as probes. Sequenced BACs in the public databases were identified by a BLAST 2.0 alignment search of the HTGS database^{25,26} (http://www.ncbi.nlm.nih.gov/blast/ blast.cgi) and the complete BAC sequences were obtained from GenBank (http://www.ncbi.nlm.nih.gov/genbank/ query form.html). BACs with known end-sequence information were determined by searching the BAC End Sequence Database at TIGR (http://www.tigr.org/tdb/humgen/). The overlapped BACs or BAC contigs were obtained by searching the Washington University Human mapping database (http://genome.wustl.edu/gsc/cgibin/fpchuman.single.pl) for likely matches to specified clones.

Sequencing of YAC and BAC ends

Isolation of YAC ends was performed using a modified vectorette method, using primers as previously de-

scribed.²⁷ YAC DNA (0.1 μ g) was digested with 10 units of *RsaI* and *AluI* in 30 μ l reaction buffer. Five microlitres of digested YAC DNA was ligated to vectorette adapters using 10 units of T4 DNA ligase in a total volume of 50 μ l by incubation overnight at room temperature. The YAC end fragments were purified by Qiagen PCR column kit and directly sequenced using the left or right internal primers.

To obtain BAC end-sequences, BAC plasmid DNA was prepared using alkaline lysis procedure and tip-500 columns (Qiagen).²⁸ The quality and quantity of DNA samples were tested by HindIII digestion pattern on agarose gels, as well as by the presence of expected STS markers. Direct BAC end sequencing was performed using an automated ABI 373 DNA sequencer. Three micrograms of BAC DNA and 50 pmoles of primer were used in a total volume of 40 μ l. The following primers were used: T7 (5'-TAATACGACTCACTATAGGG-3') and SP6 (5'ATTTAGGTGA-CACTATAG-3'). PCR reactions were carried out under the following cycle conditions: initial denaturation at 96°C for 4 min; 100 cycles of 96°C for 10 s, 50°C for 10 s, 60°C for 4 min. The end sequences of some BAC clones were obtained by searching BAC End Sequence Database at TIGR.

To identify candidate genes, known genes and ESTs previously mapped to the region between D2S177 and D2S337 were selected from the Human Transcript Map.²⁹ The selected ESTs and genes were tested by PCR amplification against our YAC and BAC contigs and positive clones further characterised, as described above.

Results

Construction of a YAC contig

The initial goal was to construct an extensive YAC contig spanning the sitosterolaemia candidate region on chromosome 2, between markers *D2S2174* and *D2S2294*. Based on the publicly available contig maps (contig WC2.4) from Whitehead Institute/MIT Center for Genome Research (WI/ MIT) (http://carbon.wi.mit.edu) and the CEPH-Généthon (CEPH) (http://www.cephb.fr/infoclone.html), 30 YAC clones were identified, using following microsatellite and STSs markers (*D2S414*, *D2S2294*, *D2S119*, *D2S2298*, *D2S1484*, *D2S1486*, *D2S1485*, *D2S1830*, *D2S2174* and *D2S229*). Additionally, information from a published partial YAC contig was also available.³⁰ All YAC clones were screened for the markers and confirmed by testing three colonies of each

| | MARKERS telomere | | D2S2230 | D2S2172 | D2S1754 | D2S2306 | D2S2613 | D2S2259 | 171070 | 8/61/1 | D2S4010 | D2S2294 | DZ34011 | 1 22636 | D25119 | D2S4009 | D2S2298 | stSG63433 | D2S4001 | A007E35 | D2S4002 | D2S1435 | A009V46 | 45B10 | 45B2 | 42E 2 | 42D3 | Y888G9.L | 42 B1 | <u>A010A13</u> | D2S1485 | D2S1486 | D2S1830 | D2S1484 | 45C2 | D2S2035 | D2S2174 | D2S1488 | 42G8 | 45011 | 45A12 | 42C5 | 42B7 | D2S2280 | D2S2291 | D2S1487 | D2S2240 | D2S1489 | D2S2378 | D2S2182 | DZSZZZ/ |
|-----------|---------------------|-----|---------|---------|---------|---------|---------|---------|--------|--------|---------|---------|---------|---------|--------|---------|---------|-----------|---------|---------|---------|---------|---------|-------|------|-------|------|----------|--------------|----------------|---------|---------|---------|---------|------|---------|---------|---------|------|-------|-------|------|------|---------|---------|---------|---------|---------|---------|---------|---------|
| YAC CLONE | S (| Mb} | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 934 e2 | | 1.5 | + | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 962c12 | | 1.4 | + | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 802c5 | | 0.9 | + | + _ | 7 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 738h9 | | 0.8 | | | + | + | + | + | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 748b9 | | 1.8 | | | | | + | + | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 899b1 | | 1.4 | | | | | ÷ | + | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 888g9 | | 0.9 | | | | | | | U۰ | • • | + | + • | • | • | + • | ÷ | + | + | + | + | + | + | ÷ | + | ÷ | + | + | + | | | | | | | | | | | | | | | | | | | | | | | |
| 930a1 | | 1.7 | | | | | | | | - | + | + • | • | + | + • | + | ŧ | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | | | | | | | | | | | | | |
| 919c10 | | 1.1 | | | | | | | | _ | | | | + | + • | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | | | | | | | | | | | | | | | | |
| 919 e8 | | 1.8 | | | | | BA | C R- | 489 | G24 | L I | | | | | | + | + | ÷ | + | + | + | + | + | + | + | + | + | + | + | ÷ | + | + | + | + | + | + | + | + | | | | | | | | | | | | |
| 742b10 | | 1.1 | | | | | | | | | | | | | | | + | + | ÷ | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | | | | | | | | | | | | | | | |
| 957b10 | | 1.2 | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | | | | | | | | | | | | | | | |
| 922b12 | | 1.8 | | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | + | | | | | | | | | | | | | | | |
| 921b6 | | 1.6 | | | | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | + | + | + | ŧ | ŧ | + | + | + | + | | | | | | | | | | | | | | | |
| 755c8 | | - | | | | | | | | | | | | | | | | | | | | | + | + | + | + | + | ÷ | + | + | + | | | | | | | | | | | | | | | | | | | | |
| 746g4 | | 0.8 | | | | | | | | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | + | | | | | | | | | | | | | | | | | | | |
| 751h4 | | 1.1 | | | | | | | | | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | | | | | | | | | | | | | | | | | | | |
| 913g4 | | 1.6 | | | | | | | | | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | | | | | | | | | | | | | | | | | | | |
| 939g6 | | 0.5 | | | | | | | | | | | | | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | + | | | | | | | | | | | | | | |
| 981a4 | | 0.8 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | + | + | + | + | + | | | | | | | | | | | |
| 761 e1 | | 1.2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | + | + | + | + | + | + | + | + | + | + | + | + | |
| 892a2 | | 0.8 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | + | + | |
| 972c5 | | 1.2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | + | + · | + |

Figure 1 YAC clone contig encompassing the sitosterolaemia locus. The markers are oriented along the X-axis from telomeric end at the left to the centromeric end at the right, and the YAC clones are indicated along the Y-axis. There are two gaps in the contig (boxed areas, see Text). The distal centromeric gap is closed by a linking BAC, R-489G24, giving a contiguous contig from *D2S1754* to *D2S2227*. ESTs mapped to the YAC contig are underlined, microsatellite markers are italicised and STSs are in normal font.

clone for their STS contents (Figure 1). Two gaps were identified at the telomeric ends, and no further YACs were identified, despite additional library screening. However, we subsequently identified a BAC, R-489G24, positive for markers D2S2259 and an EST T71978, that linked YACs 888g9 and 899b1, thus closing one gap (Figure 1). Since the sitosterolaemia locus is located towards the centromeric end, no further attempts were made to close the more distal gap. Using the YAC contig, new markers were generated, employing a combination of YAC end-sequence analyses and inter-Alu PCR. To confirm that the identified STSs were from chromosome 2, all of these markers were screened for their presence in human chromosome 2-specific somatic cell hybrid cell-lines. Sixty ESTs from the databases (Unigene and GeneMap98, see Materials and methods) were screened by PCR against the YAC contig. Six of these were positive for the YAC contig (Figure 1), of which five mapped to within the region of interest. The sixth EST (T71978) was found to be positive on a linking BAC (see above). By performing inter-Alu PCR using YACs 919c10, 930a1 and 761e1 as templates, we generated 35 additional unique sequences from Alu PCR clones for obtaining STS markers and eight microsatellite repeat markers that allowed for further fine-mapping of the sitosterolaemia locus (Lee *et al*, manuscript submitted). A total of 76 STSs, including 17 new STSs generated from YAC insert-end sequences and inter-Alu PCR products and nine EST markers, were used to order the clones (Table 1 and Figure 1) and span an approximate distance of 5 cM. This physical map provided a resource for the construction of a BAC contig.

Note also that our YAC contig is a telomeric extension of a published adjoining YAC contig, thus providing a continuous map that spans chromosome 2p15–2p21 (D2S1364–



Figure 2 BAC clone contig encompassing the sitosterolaemia locus. The markers are oriented along the X-axis from telomeric end at the left to the centromeric end at the right, and the BAC clones are indicated along the Y-axis. Prefixes of the BAC clones are as follows, R; from RPCI-11 BAC library; and C; from CITB-SHP-C BAC library. ESTs are underlined, microsatellite markers are shown in italics and STSs are in normal font. A box indicates BACs, for which almost complete sequence information is available in the genome databases. Four gaps (indicates by vertical lines) were identified, but these BACs are indicated, as they contain markers placed on the YAC contig framework (Figure 1). Of these the more centromeric is spanned by YAC 888g9. BAC R-436K12 (not shown) is linked to the contig published by Kirchener *et al*,³¹ and links our YAC contig at the centromeric end.

Table 2 ESTs and genes mapped to the YAC/BAC contigs in this study

| NCBI No (GeneMap '99) | Aliases or synonyms | UniGene No. | Genbank Accession No. | Image Clone ID | Known genes | Mapping data |
|--------------------------|-------------------------------|---------------------------------|--------------------------|--------------------|---|--------------------|
| WI-20996 | stSG41980, T48876 | Hs.19280 | R26389 | 132199 | KIAA0544 protein | R-50107 |
| SGC33875 | | | T98917 | 122669 | | R-78114 |
| stSG48396 | | Hs.98023 | AA854974 | 1394041 | | R-459K11 |
| SGC34340 | WI-6575, SGC34340 | Hs.78909 | X78992 | | ERF-2 protein | R-339H12 |
| STSG16054 | A009V10 | Hs.17711 | R98822 | 207006 | | R-339H12 |
| WI-14187 | G21943 | Hs.16063 | AA515534 | 925219 | | R-339H12 |
| stSG52431 | | Hs.165571 | AI566776 | 2168475 | | R-489G24 |
| T89476 | sts-T89476 | Hs.16587 | AA934036 | 1551421 | | R-489G24 |
| stSG15818 | | | R83265 | 194194 | | R-489G24 |
| WI-8407 | | Hs.108441 | Z29481 | | 3-Hydroxyanthranilic acid di-oxygenase | R-489G24 |
| BCD1971 | | | M79071 | no image clone | ,,, | R-489G24 |
| T71978 | sts-T71978 | Hs.168439 | AA534545 | 925906 | | R-489G24 |
| stSG58568 | | Hs.58598 | AI359618 | 2013757 | | R-489G24 |
| stSG30561 | | Hs.58598 | AA169121 | 594556 | | 506D15 |
| H96893 | stSG21270 | Hs.32241 | AI274775 | 1986682 | | R-489K21 |
| stSG21136 | | | H58682 | 205857 | | R-489K21 |
| | | | AF151818 | | CGI 60 protein | R-489K21 |
| T99836 | | Hs 18176 | T99836 | 123200 | | R-489K21 |
| WI16988 | A007F35 | Hs 142718 | AA034046 | 429916 | | R-1081G2 |
| stSC63433 | //// 255 | Hs 190354 | AA700586 | 433330 | | R-1081G2 |
| stSG32054 | stSG1757, SHGC-8019 T17102 | Hs.182490 | M92439 | 00000 | Leucine-rich protein mRNA | R-1081G2 |
| | 117102 | Hc 128203 | AI223013 | 1838809 | | R-2201/23 |
| | | Hc 225721 | AI873/// | 2362150 | Trans prepultransferase (TPT) | D 550M23 |
| | | Hc 225721 | AA880371 | 1/71263 | | D 550M23 |
| | | 113.223721 | AA009371 AA457300 | 83810/ | | D 550M23 |
| | | | AAQ20060 | 127/287 | | D 550M23 |
| | | Hc 197045 | AA020000 AA027600 | 13/420/ | | D 550M22 |
| 1004127/ | c+5C 51006 | TIS. 10/943 | AA95/099 AA164202/ | 1471137 | Protoin phosphatasa 20 | R-3391V123 |
| H99661 | 313031090 | Hs.5687 | AA164363/ AA565932 | FFZC | Protein phosphatase 2C | R-2413 |
| stSG3387 | A003R48 | | R11895 | 25315 | | R-2415 |
| stSG52154 | | Hs.112916 | AA620873 | 1049335 | | R-2415 |
| M95548* | SHGC-9884, stSG4626 | Hs.198294/ Hs.154834 | D82326/ M95548 | | Amino acid transporter, SLC3A1 | R-2415 |
| M95548* | SHGC-9884, stSG4626 | Hs.110 | AB007896 | | KIAA0436 mRNA | R-2415 |
| A009V46 | | Hs.174862, Hs.220859 | H95593 | 242930 | | R-194L1 |
| A010A13 | WI-18144 | Hs.124990 | H58934 | 207758 | | R-194I 1 |
| D29089 | | | D29089 | no image clone | | R-194I 1 |
| stSG8383 | | | H60063 | 205767 | | R-194I 1 |
| | | Hs.132799 | AA922097 | 1543611 | | R-194I 1 |
| | | Hs.129473 | AA994134 | 1628550 | | R-19411 |
| | | Hs.213492 | AI928677 | 2466254 | | R-194I 1 |
| | | Hs 124990 | H60592 | 207898 | | R-19411 |
| | | | W80452 | 415494 | | R-194I 1 |
| stSG26329 | | | H57813 | 205424 | | R289F6 |
| 515626525 | | Hs 136519 | AA601487 | 1100969 | | R-44205 |
| | | 113.1 50517 | T87425 | 115418 | | R-44205 |
| \\/I_3495 | C02557 | Hc 188588 | AA583683 | 1088083 | | R-44205 |
| VVI-J-7/J | 602337 | 113.100500 | ΔΔ835723 | 1372934 | | R-44205 |
| | | | H64341 | 210718 | | R-44205 |
| | | | AA838130 | 13855/0 | | P 44205 |
| | | Hs 170/22 | ΔΙΔ50Λ50 | 21 <u>/</u> 0052 | | R_44205 |
| | | 113.170420 Hc 170420 | 01737030 ۸\\/204717 | 217222 2722/80 | | D 44203 |
| | | 113.1/0420 Hc 222172 | ANN/022704 | 2122400 | | D 44203 |
| c+SC 16110 | | 113.2331/2 Uc 07404 | AVVUZZ/UU | 240013/ 720207 | | D 20V21 |
| N24004 | | 113.77070 | ND1005 | 127201 | | Π-07ΝΖΙ D 572Γ1 |
| M/I 2076 | SHCC 17227 | LC 246042 | NZ4073 | 200/92 | | N-370F1 |
| vv1-37/0 c+SC/0702 | 3110C-17237 | П3.240042 Цс 1 <i>6764</i> 0 | 1N/ 3743 | 273200 | | R-J/0FI |
| 313U477UZ | | ПS. 10/040 Цс 78024 | 110//30 LICI 102011 | 220030 (hN/SH2) | Mismatch ropair protain | R-430NIZ |
| vvi-10/91 | SHGC-10660 | П3./0У34 | 1166003411 | (nivisitz) | (MSH2) mRNA | K-430K12 |
| stSG60189 | | Hs.122384 | AI015254 | 1641212 | | R-436K12 |
| embl-AA007353 | sts-AAUU/353 | Hs.256042 | AAUU/353 | 429281 | | R-436K12 |
| | | | | | | Continued |

Continued

European Journal of Human Genetics

Table 2 (Continued)

| | ucuj | | | | | |
|--------------------------|------------------------------------|----------------|--------------------------|-------------------|-------------|-----------------|
| NCBI No (GeneMap '99) | Aliases or synonyms | UniGene No. | Genbank Accession No. | Image Clone ID | Known genes | Mapping data |
| SGC34683 | SHGC-34683, stSG28638, stSG9035 | Hs.117085 | AA677756 | 430606 | | R-436K12 |

All ESTs and genes that were mapped to the YAC and BAC contigs (Figures 1 and 2) are shown. For clarity, only the BAC ID is shown in the far right column. BAC R-436K12 is not indicated on the BAC contig (Figure 2), but is contiguous with the centromeric end. Only a representative EST or Image clone is indicated, where multiple clones were identified. The asterisk indicates a GeneMap ID, M95548, which identifies two separate genes that share the 3' UTR (see text). Additionally, there are two GeneMap98 IDs for the same gene (PP2C) that have been consolidated.

D2S1754, $\sim\!14$ Mb) and contains several human disease loci. 31

Construction of a BAC contig

To construct a BAC map, we used the following strategy; (1) identify BAC clones using both a PCR-based and filter hybridisation-based BAC library screenings, (2) screen all positive BACs for STS content by PCR, (3) search BAC-related databases for updated information, and (4) perform chromosome-walks using selected STSs generated from BAC ends. Initially, PCR was used to screen the CITB-978SK-B human BAC library using six repeat polymorphic markers (D2S2294, D2S119, D2S2298, D2S1830, D2S2174, and D2S2291), a YAC end-sequence (from 888g9L) and an EST marker (T71978). Twelve BAC clones, positive for D2S2294, D2S119, D2S2291, 888g9L and T71978, were identified. For hybridisation-based BAC library screening, high-density filters were hybridised with a mixture of five probes consisting of ESTs T99836, T71978, A007E35, stSG63433, A010A13, previously mapped to the YAC contig. Eight more positive BAC clones were obtained from RPCI-11 BAC library. The BAC end sequences of identified clones were determined by direct automated sequencing or by searching the BAC end sequence database at TIGR. BAC end sequences of the inserts of BAC clones were used to develop further STS markers. All STS markers were tested by PCR amplification against all identified BAC clones, to verify true positives. By searching the databases in an iterative manner, we identified 18 sequenced BACs. In total, we used 118 markers, composed of 29 microsatellite markers, 53 new STSs from BAC/YAC end sequences and inter-Alu PCR sequences, and 36 EST markers. The constructed BAC contig contains 60 BAC clones, which contains a high density of STS markers, at an average of about 20 kb for each marker, and covers a physical distance of about 2.0 Mb (Figure 2). A significant number of these BACs have been sequenced (boxed, Figure 2), but about 500 kb sequence is not publicly available.

Mapping of known genes and ESTs to the YAC/BAC contigs

We have constructed a transcript map (Table 2) of the BAC contig using two methods. From GeneMap'99, based upon

two radiation hybrid panels,^{32,33} we selected 80 genes and ESTs between anchor markers D2S177 and D2S2291. All ESTs were verified by PCR against the BAC contig. Of the 80 markers, only eight known genes and 30 ESTs mapped unambiguously to our BAC contig. The eight known genes are KIAA0544 protein,³⁴ ERF2 protein,³⁵ 3-hydroxyanthra-nilic acid dioxygenase,³⁶ CGI 60 protein,³⁷ leucinerich protein,³⁸ protein phosphatase 1B (formerly PP2C),³⁹ Na+independent neutral and basic amino acid transporter (solute carrier family 3, SLC3A1),⁴⁰ and KIAA0436.⁴¹ In the second approach, using the known human genomic sequences from STSs and sequenced BACs between D2S2294 and D2S2291, we identified a further 30 ESTs by a BLASTN search of the EST databases. All of these 30 identified ESTs contain unique sequences, >95% matched to genomic sequences, and have not been previously mapped to a chromosome. A summary of the mapped ESTs to our BAC contig is shown in Table 2. We computed the expression patterns for many of these ESTs (Table 3). Additionally, we screened each of the mapped ESTs against the databases, looking for homologous ESTs/genes identified in other species, on the assumption that highly conserved expressed sequences may reflect proteins that have highly conserved and critical functions, such as selective sterol absorption. Only sequences that had >100 bp of sequence identity and >70% homology are reported (Table 3). Although such analysis is limited by the lack of depth of the EST databases for the other species, we identified 11 ESTs that appear to have homologues in non-human sequence databases (Table 3), although none from the Drosophila database were identified.

Discussion

Positional cloning techniques, combined with computerassisted data analyses of the sequence rich databases generated by human genome projects,^{42,43} has considerably facilitated the identification of disease genes. The availability of complete and detailed clone contigs of candidate regions make for efficient positional cloning projects. We first constructed a YAC contig of this region and used it as a resource for the construction of a deep BAC contig. At the centromeric end of our YAC contig, there is a YAC, 972c5,

Table 3 Expression pattern of ESTs and genes

| NCBI No (GeneMap'99) | GenBank Accession No. | Expression pattern | Known gene | Human | Mouse | Rat | Bovine | Porcine | Zebrafish | Chicken |
|-------------------------|--------------------------|------------------------------|--|---------|--------|-----|--------|---------|-----------|---------|
| WI-20996/ | R26389 | Multiple tissues | KIAA0544 protein | 57 | 4 | 0 | 0 | 1 | 0 | 0 |
| SGC33875 | T98917 | Fetal liver, spleen | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| stSG48396 | AA854974 | Testis | | 12 | Õ | Õ | Õ | Õ | Õ | Õ |
| SGC34340 | X78992 | Multiple tissues | ERF-2 protein | >75 | 16 | 15 | 6 | 0 | 1 | 0 |
| STSG16054 | R98822 | Fetal liver, spleen | | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| WI-14187 | AA515534 | Multiple tissues | | 30 | 0 | 1 | 0 | 0 | 0 | 0 |
| stSG52431 | AI566776 | Brain, eye, heart, pancreas, | | 18 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | uterus, thymus | | | | | | | | |
| embl-T89476 | AA934036 | Bone, germ cell, prostate | | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| stSG15818 | R83265 | Fetal liver, spleen | | 40 | 40 | 7 | 1 | 0 | 0 | 0 |
| VVI-8407 | 229481 | placenta, spleen, uterus | dioxygenase | 6 | 0 | 0 | 0 | 0 | 0 | 0 |
| BCD1971 | M79071 | Brain | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| embl-1/19/8 | AA534545 | Colon, kidney, liver, lung | | 22 | 23 | I | 0 | I | 0 | 0 |
| STSG38368 | AI359618 | Multiple tissues | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| 212020201 | AA109121 A1274775 | Multiple tissues | | 22 | 0 | 1 | 0 | 0 | 0 | 0 |
| C21136 | AIZ/4//J | Eetal liver spleen | | 5Z 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 313021130 | ΔF151818 | Multiple tissues | CCI 60 protein | 72 | 8 | 6 | 0 | 0 | 0 | 0 |
| T99836 | T99836 | Fetal liver spleen | cal bo plotein | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| WI16988 | AA034046 | Fetal liver spleen | | 6 | Ő | õ | õ | Ő | Ő | Ő |
| stSG63433 | AA700586 | Fetal liver, spleen | | 2 | ŏ | ŏ | ŏ | ŏ | ŏ | ŏ |
| stSG32054 | M92439 | Multiple tissues | Leucine-rich protein mRNA | >100 | 5 | Ō | Ō | 0 | 0 | 0 |
| | AI223013 | Testis | | 4 | 1 | 0 | 0 | 0 | 0 | 0 |
| | AI873444 | Ovary | Trans-prenyltransferase (TPT) | 3 | 2 | 0 | 0 | 0 | 1 | 0 |
| | AA889371 | Ovary | <, , , , , , , , , , , , , , , , , , , | 3 | 2 | 0 | 0 | 0 | 1 | 0 |
| | AA457390 | Retina | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AA828868 | Ovary | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AA937699 | Skin | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| A004I37/ | AA164383/ | Multiple tissues | Protein phosphatase 2C | 48 | 26 | 8 | 0 | 0 | 0 | 0 |
| H99661 | AA565932 | | | | | | | | | |
| stSG3387 | R11895 | Brain | | 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| stSG52154 | AA620873 | Testis | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| M95548* | D82326/ | Brain, kidney, pancreas, | Amino acid transporter, | 36 | 38 | 2 | I | I | 0 | 0 |
| N 40 5 5 40+ | M95548 | uterus, colon | | 100 | 17 | 2 | 0 | 0 | 0 | 0 |
| NI95548° | AB007896 | Multiple tissues | KIAAU436 MKINA | 100 | 16 | 3 | 0 | 0 | 0 | 0 |
| A009V40 | H2024 | Fetal liver, spleen | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| 29089 | 1130254 P20080 | Enidermis keratinocyte | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| stSG8383 | H60063 | Fetal liver spleen | | 1 | õ | õ | õ | õ | õ | õ |
| | AA922097 | Testis | | 1 | õ | Õ | õ | Õ | Õ | Õ |
| | AA922097 | Testis | | 3 | 5 | Ō | 1 | 0 | 0 | 0 |
| | AA994134 | Tonsil | | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AI928677 | Brain | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | H60592 | Fetal liver, spleen | | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| | W80452 | Fetal liver, spleen | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| stSG26329 | H57813 | Fetal liver, spleen | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AA601487 | Adrenal gland | | 1 | 2 | 0 | 0 | 0 | 0 | 0 |
| | T87425 | Fetal liver, spleen | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| WI-3495 | AA583683 | Kidney, nose | | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AA835723 | Germinal center B cell | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| | H64341 | Fetal liver, spleen | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AA838139 | Ovary | | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AI439038 | Lung | | 2 | 0 | 0 | 0 | 0 | 0 | 0 |
| | AVVZU0/1/ | Lung | | 2 1 | U 2 | 0 | 0 | 0 | 0 | 0 |
| stSC/6/10 | AVVUZZ/UD | Lai Tostis | | ו כ | د م | 0 | 0 | 0 | 0 | 0 |
| N24095 | N24095 | Melanocyte | | 2 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| WI-3976 | N75945 | Whole blood | | י ז | 0 | 0 | 0 | 0 | 0 | 0 |
| stSG49702 | H87795 | Retina, colon | | 2 | õ | õ | õ | õ | õ | õ |
| WI-18791 | HSU03911 | Multiple tissues | Mismatch repair protein | 63 | 13 | 1 | 1 | Ő | Õ | 3 |
| | | | (MSH2) | | | | | | | |

Continued

European Journal of Human Genetics

| Гab | le | 3 | (Continued) | |
|-----|----|---|--------------|--|
| | - | - | continucay | |

| NCBI No (GeneMap'99) | GenBank Accession No. | Expression pattern | Known gene | Human | Mouse | Rat | Bovine | Porcine | Zebrafish | Chicken |
|-------------------------|--------------------------|---|------------|-------|-------|-----|--------|---------|-----------|---------|
| stSG60189 | AI015254 | Testis | | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| AA007353 | AA007353 | Lung | | 3 | 0 | 0 | 0 | 0 | 0 | 0 |
| SGC34683 | AA677756 | Fetal liver, spleen, neuroepithelium | | 5 | 0 | 0 | 0 | 0 | 0 | 0 |

Expression profiles were determined for the ESTs and genes, based upon the identification of the EST or gene transcript in various cDNA libraries. Thus this profile is a minimal expression pattern. Additionally, homologues for the ESTs and genes were searched for (see Materials and methods) and the number of ESTs thus identified are indicated in the columns on the right. No homologues (based upon parameters specified in the text) were found in the *C. elegans* or *D. melanogaster* databases.



Figure 3 Summary of the YAC, BAC and EST mapping data. The figure shows a summary of the data presented in this study, indicating the genetic distance, physical distance and Genbank Accession numbers for mapped ESTs and genes located in the region of interest. Note that the genetic distance, based upon publicly available databases, spans ~ 10 cM, but spans ~ 5 Mb. Although only the Genbank IDs are shown, all identified EST can be obtained by utilising the Unigene or the GeneMap'99 identifiers shown in Table 2. Accession numbers in bold represent known genes, the remaining represent putative ESTs. The exact order of the ESTs at any given map location can not be determined at present and are thus grouped, indicated by the vertical lines.

which contains markers *D2S2182* and *D2S2227*, which are also located in a published adjoining YAC contig.³¹ Thus combined with this published YAC contig, this provides a continuous map that spans chromosome 2p15-2p21 (*D2S1364–D2S1754*, ~14 Mb) and contains several human disease loci.³¹

Sixty-seven new STSs were identified by inter-Alu PCR and YAC/BAC end sequencing. The high-resolution physical map generated in this study spans ~ 2 Mb with complete coverage of the minimal region of sitosterolemia. The data presented

here have been parsed for multiple ESTs for single genes represented in the databases and we have attempted to summarise data that are found scattered in a number of different databases, increasing the utility of this information. A summary of the results is provided in Figure 3. Based upon the radiation hybrid mapping databases, our initial YAC contig spans approximately 10 cM. However, this area appears to span only 5 Mb in physical length, suggesting a lower than expected recombination frequency (Figure 3). Assuming that all the non-redundant ESTs mapped to the BAC contig are unique transcripts and taking into account the small number of genes known to map into the BAC contig, we estimate that the gene density is approximately 1 gene per 50 kb of genomic DNA (Figure 3, 40 ESTs and genes mapped with the 2 Mb area).

One of our findings is the mis-assignment of BAC R-35M22. This BAC was previously assigned to chromosome 4 (Genbank accession number AC016338, Birren *et al*, direct submission), but is positive to DNA sequences from BACs R-24I5 and R-194L1. Additionally, it also contains ESTs *A004137*, *H99661*, *stSG3387*, *stSG52154*, *M95548*, and *M95548*, 9 of 14 exons of KIAA0436 protein and exon 2 of Na+-independent neutral and basic amino acid transporter, thus placing it firmly on chromosome 2, in the interval *D2S119-D2S2291*.

Our integrated BAC contig allows for more accurate placement of genes and ESTs than the corresponding region in Genemap'99. In the *D2S119-D2S2291* interval from GeneMap'99, 43 ESTs listed, 39 of which are unique. However, only nine of these 39 ESTs actually map to the *D2S119-D2S2291* interval into our BAC contig, 30 of 39 map outside of this region. Of the 40 ESTs we have physically mapped to the *D2S119-D2S2291* interval of our BAC contig, 31 of these were previously assigned to lie outside of this region. Therefore, the accuracy of GeneMap'99 for the *D2S119-D2S2291* interval is only 25%, which is similar to the 30% reported by Kirschener *et al* for the *D2S123-D2S2251* interval, but much lower than 75% in the *D2S291-D2S123* interval reported by the same authors.³¹

In summary, we have developed 67 new STSs, constructed an integrated YAC and BAC contigs for sitosterolaemia region and mapped eight known genes and 48 ESTs to the contig. These results will facilitate the identification of the sitosterolaemia gene and other disease genes located in this region. Additionally, this information may be useful in ordering some of the sequenced BAC contigs and accelerate sequencing of the corresponding genomic clones.

Acknowledgments

We are grateful to Dr Anand Srivastava for expert advice and critical review of our work, to Starr Hazard and the BioMolecular Computing Resource for assistance with the software. This work was funded by a Scientist Development Award from the American Heart Association grant 9730087N (SB Patel) and by the National Institutes of Health, NHLBI Grant HL60616 (SB Patel) and MO1 RR01070-25 (MUSC GCRC), and by an intramural award from the University Research Committee, Medical University of South Carolina (SB Patel).

References

1 Bhattacharyya AK, Connor WE: Beta-sitosterolemia and xanthomatosis. A newly described lipid storage disease in two sisters. *J Clin Invest* 1974; **53**: 1033–1043.

- 2 Miettinen TA: Phytosterolaemia, xanthomatosis and premature atherosclerotic arterial disease: a case with high plant sterol absorption, impaired sterol elimination and low cholesterol synthesis. *Eur J Clin Invest* 1980; 10: 27–35.
- 3 Kwiterovich Jr P, Bachorik PS, Smith HH, *et al*: Hyperapobetalipoproteinaemia in two families with xanthomas and phytosterolaemia. *Lancet* 1981; 1: 466–469.
- 4 Salen G, Shefer S, Nguyen L, *et al*: Sitosterolemia. [Review]. J Lipid Res 1992; 33: 945-955.
- 5 Shefer S, Salen G, Bullock J, *et al*: The effect of increased hepatic sitosterol on the regulation of 3-hydroxy-3-methylglutaryl-coenzyme A reductase and cholesterol 7 alpha-hydroxylase in the rat and sitosterolemic homozygotes. *Hepatology* 1994; **20**: 213–219.
- 6 Salen G, Ahrens Jr E, Grindy SM: Metabolism of beta-sitosterol in man. *J Clin Invest* 1970; **49**: 952–967.
- 7 Gregg RE, Connor WE, Lin DS, Brewer Jr H: Abnormal metabolism of shellfish sterols in a patient with sitosterolemia and xanthomatosis. *J Clin Invest* 1986; 77: 1864–1872.
- 8 Salen G, Shore V, Tint GS, *et al*: Increased sitosterol absorption, decreased removal, and expanded body pools compensate for reduced cholesterol synthesis in sitosterolemia with xanthomatosis. *J Lipid Res* 1989; **30**: 1319–1330.
- 9 Nguyen LB, Salen G, Shefer S, et al: Decreased cholesterol biosynthesis in sitosterolemia with xanthomatosis: diminished mononuclear leukocyte 3-hydroxy-3-methylglutaryl coenzyme A reductase activity and enzyme protein associated with increased low-density lipoprotein receptor function. *Metab Lin Exp* 1990; **39**: 436–443.
- 10 Bhattacharyya AK, Connor WE, Lin DS, McMurry MM, Shulman RS: Sluggish sitosterol turnover and hepatic failure to excrete sitosterol into bile cause expansion of body pool of sitosterol in patients with sitosterolemia and xanthomatosis. *Arterio Throm* 1991; **11**: 1287–1294.
- 11 Nguyen LB, Shefer S, Salen G, *et al*: A molecular defect in hepatic cholesterol biosynthesis in sitosterolemia with xanthomatosis. *J Clin Invest* 1990; **86**: 923–931.
- 12 Patel SB, Honda A, Salen G: Sitosterolemia: exclusion of genes involved in reduced cholesterol biosynthesis. *J Lipid Res* 1998; 39: 1055–1061.
- 13 Patel SB, Salen G, Hidaka H, *et al*: Mapping a gene involved in regulating dietary cholesterol absorption. The sitosterolemia locus is found at chromosome 2p21. *J Lipid Res* 1998; **102**: 1041-1044.
- 14 Comuzzie AG, Hixson JE, Almasy L, *et al*: A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. *Nat Genet* 1997; **15**: 273–276.
- 15 Hart TC, Pallos D, Bowden DW, *et al*: Genetic linkage of hereditary gingival fibromatosis to chromosome 2p21. *Am J Hum Genet* 1998; **62**: 876–883.
- 16 Cohen D, Chumakov I, Weissenbach J: A first-generation physical map of the human genome. *Nature* 1993; 366: 698–701.
- 17 Chumakov IM, Rigault P, Le Gall I, *et al*: A YAC contig map of the human genome. *Nature* 1995; **377**: 175–297.
- 18 Hudson TJ, Stein LD, Gerety SS, et al: An STS-based map of the human genome [see comments]. Science 1995; 270: 1945–1954.
- 19 Cruts M, Backhovens H, Wang SY, *et al*: Molecular genetic analysis of familial early-onset Alzheimer's disease linked to chromosome 14q24.3. *Hum Mol Genet* 1995; **4**: 2363–2371.
- 20 Strong TV, Tagle DA, Valdes JM, *et al*: Widespread expression of the human and rt Huntington's disease gene in brain and nonneural tissues. *Nat Genet* 1993; **5**: 259–265.
- 21 Qin S, Nowark NJ, Zhang J, *et al*: A high-resolution physical map of human chromosome 11. *Proc Natl Acad Sci USA* 1996; **93**: 3149–3154.
- 22 Nemeth-Slany A, Talmud P, Grundy SM, Patel SB. Activation of a cryptic splice-site in intron 24 leads to the formation of apolipoprotein B-27. *Atherosclerosis* 1997; **133**: 163–170.

- 23 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Biol* 215: 403-410.
- 24 Kim UJ, Birren BW, Slepak T, *et al*: Construction and characterization of a human bacterial artificial chromosome library. *Genomics* 1996; 34: 213–218.
- 25 Ouellette BF, Boguski MS: Database divisions and homolgy search files: a guide for the perplexed. *Genome Res* 1997; 7: 952–955.
- 26 Center TSCaTWUGS: Towards a complete human genomic sequence. *Genome Res* 1998; 8: 1097-1108.
- 27 Riley J, Butler R, Ogilvie D, *et al*: A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res* 1990; **18**: 2887–2890.
- 28 Kirschner LS, Stratakis CA: Large-scale preparation of sequenceready bacterial artificial DNA using QIAGEN columns. *Biotechniques* 1999; 27: 72–74.
- 29 Deloukas P, Schuler GD, Gyapay G, *et al*: A physical map of 30,000 human genes. *Science* 1998; **282**: 744–746.
- 30 Schell U, Wienberg J, Kohler A, *et al*: Molecular characterization of breakpoints in patients with holoprosencephaly and definition of the HPE2 critical region 2p21. *Hum Mol Genet* 1996; **5**: 223–229.
- 31 Kirschner LS, Taymans SE, Pack S, *et al*: Genomic mapping of chromosomal region 2p15-p21 (D2S378-D2S391): integration of Genemap'98 within a framework of yeast and bacterial artificial chromosomes. *Genomics* 1999; **62**: 21–33.
- 32 Stewart EA, McKusick KB, Aggarwal A, *et al*: An STS-based radiation hybrid map of the human genome. *Genome Res* 1997; 7: 422–433.
- 33 Gyapay G, Schmitt K, Fizames C, *et al*: A radiation hybrid map of the human genome. *Hum Mol Genet* 1996; 5: 339–346.
- 34 Nagase T, Ishikawa K, Miyajima N, *et al*: Predicition of the coding sequence of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro. *DNA Res* 1998; 5: 31–39.

- 35 Nie XF, Maclean KN, Kumar V, McKay IA, Bustin SA: ERF-2, the human homologue of the murine Tis11d early response gene. *Gene* 1995; **152**: 285–286.
- 36 Malherbe P, Kohler C, Da Prada M, *et al*: Molecular cloning and functional expression of human 3-hydroxyanthranilic-acid dioxygenase. *J Biol Chem* 1994; **269**: 13792–13797.
- 37 Lai CH, Chou CY, Ch'ang LY, Lui CS, Lin W: Identification of Novel Human Genes Evolutionarily Conserved in Caenorhabditis elegans by Comparative Proteomics. *Genome Res* 2000; 10: 703–713.
- 38 Hou J, Wang F, McKeehan WL: Molecular cloning and expression of the gene for a major leucine-rich protein from human hepatoblastoma cells (HepG2). *In Vitro Cell Dev Biol Anim* 1994; **30A**: 111–114.
- 39 Marley AE, Kline A, Crabtree G, Sullivan JE, Beri RK: The cloning expression and tissue distribution of human PP2Cbeta. *FEBS Lett* 1998; **431**: 121–124.
- 40 Miyamoto K, Segawa H, Tatsumi S, *et al*: Effects of truncation of the COOH-terminal region of a Na+-independent neutral and basic amino acid transporter on amino acid transport in Xenopus oocytes. *J Biol Chem* 1996; **271**: 16758–16763.
- 41 Ishikawa K, Nagase T, Makajima D, *et al*: Prediction of the coding sequences of unidentified human genes. VIII. 78 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* 1997; 4: 307–313.
- 42 Collins FS: Positional cloning moves from perditional to traditional [published erratum appears in Nat Genet 1995 Sep;11(1):104]. *Nat Genet* 1995; **9**: 347–350.
- 43 Boehm T: Positional cloning and gene identification. *Methods* 1998; 14: 152–158.