Supplementary Table S2: Full list of predicted GIs in UM_CSW genome.

Genomic islands	Genomic size (bp)	GC content (%)	Contig number	Start	End	Gene description	Gene-sequences
GI1	11040	62.3	Contig 2	63446	74485	hypothetical protein	MTQTERLPAYTEAHPPTPSSDELRAQIPGWGADLDPKDRPSNPKLRQDLPTETHWDFPERQPEKWPRERSVEHRFLTPVFGTAQPPSGVSGMLRRYAYKFSEGRAAHWLILLYADRVDTAEHHVRSFLTTRPDNPITETGIKSELTHHGIQSRLGRKRSDLVHIWMDPFIVAGPWILGGQGIASLARKAVKAVGNNGGRGGGR
						low molecular weight T-cell antigen TB8.4	MTGADGAVSTNPEPQPDQDWIVNTKCSFPQFVAALNAEDIRAAQQFAITPAAQSWVGSYLASSPDQRRQMLRRAQSIPDARQYIDTMLRVASTCKNF
						hypothetical protein	MRYEEFISSVAERSGLFQGDAVALTRATLTTLAERISGGEARDLAAQLPSPLQDALLPTEEEAEAFSFDEFVNRVAERSGRDTEAAEAAVDAVLATIRDAVTPGEFDDVLSQLPSEFQRLGTR
						protein of unknown function UPF0118	MGPPGRRFDWHSPFFVGLTASAGVAVTYGAVRGLAAVSSVLLLIGVAFFIALGLEPAVSWLVGKKVPRWAAVAVVLFAVIAVVAGSLAAAIPPLVQQAREFTDQAPHYIQHIGDHSSWIGRLNDRFHLQQRITETVHGSGGAVVGQALKAGEAVFGAVADVVVVVVLTVYFLVDFPRIRATLYRFVPHSRRPRAILIGDQMFAKVSAYVFGNIVISAIAGVATFVWLVIFDVPYPLLLGIFVALLDLIPFGSSVAGVIVAAVALTVSIPVAIATLLFYTLFRFVEDYLLVPKIIGRAVNVPAVATLLAVLVGGALLGIVGALVAIPIAAAIQLAARELLFPTLDKL

hypothetical protein	MVVIGHIADPEEPRHGRRGKAVQSTIARLAAAMVGTAGIEDALSKLTSASLALIPGADCAKISTIENGHLRSITATSQLTSSLDSAQQAAGQGPCLETISAKKAIRCDDLRTDTRWPRFTPRATTAGVYSVLSTPIDIPGETGATLTLFGFRAEAFGPDAEAVGAMLANHAAIALMHVKQERQFKAALATRNVIGQAKGMIMERFGIDAARAFAMLRAISQQTNTPLRELATRLVDGAKR
hypothetical protein	MQRLRAGTSIIYASGSLDGETRAVLCRAVADELTQGSAQLVLELSSATSIDDTAVEALVDATGLAAESDTSICLVTSPTGPIARALAAADLLERFEIFATVDEAQRHP
hypothetical protein	MGIAHRAEPPSLHSVFNLSPLLFDGRRADAVLELAAEAVRSLGKCSSEGAYRVVDGSLVDNSDPDRPLDSGLDAAVTASVGVDHEIVLPDSNWRYAMTLRTTNTVTGVLVVRADNPASCHELVLLKMLAQQAATAMASADTIERERRQHVQLRELTDRHEETIHRLSRSVAELERREHIHKALTNLSGSADAASIADALHDLTSLTVSVEDTFGNLRAWSPAPIPTAYRAVGGGNREDVIRNAGSRGHHSDCGNRIFSLIRARAELLGVVVLHDPQRRADRLDIFALEYAAAVLAVEFSHQRSIAETESRLSRDLVDDLLTGTDNATAYARGEALGYNLHRPHRVTVLQWSAEIDGALIARSATRWATSAGLHPLCARRPSMTILLTEDLPQPQSLYRAISAAVGHGRGWIAIGSVATSPSELPRSFAEARRTLRMQKASVGRHGFRRFDDLGVCRILDPSDNGPEVGEFLAEWLGPLMAYDQEKNAELVNTLARYLDSGGNYDQTARALNIHRSTLRYRLGRIRDVSGRDLQNVETRLNLHLATKVSEMIGETQPIMRAASTGERRPSDEGMPK

putative transcriptional regulator, AraC family	MTELPFDLAPEVRTTDIDEAASALGRVYVTAELVPNKTNSVGMHMNAAQLPLITAGYLGFGADVTVRANDVTAYYIDAPLSGSAVNRWRDGELVKTTTGSVAVFTPGTPCVLDWSGDCGQICLKVSEPQMRRQLEAMLNRPVRKRITFARQFDLSTTAAHDWYHLVSLLAREVGRTDGLLNHRLAAANLQLLLIQGLLQIQPHNYTEALAESEGAASANVAKRAIDLLHAHPETMWSTAQLARQTGVSARALQRAFERSDQPSPMAYLRRLRLHRVYTELAASSPDSVTVTMVAGRWGFVHLGRFAGQYQQLFGETPAETLRHRVADALPTQPQPPL
hypothetical protein	MFISDRTTPQEPDMTTADEDSERDFLAKCVADAYQMLEIIPGVNPNGAALVWLADQFMQVRRRTESHLASTA
hypothetical protein	MKRPNSRFLLAISRTGWLGRTVEILLRGLLASDTGTVKVSMPWSI
hypothetical protein	MWSEYYDVATECAVTLRMLLASMSNVTSIWGSPRGIGRMLPNFETAQ
ThiJ/PfpI family protein	MPNMLQGKTIAILAADGVERVELVHPSAAVHRAGGQTELLSLRNGWIQMRNNDLDPAGTFTVDRAVAGASIDDYGGLLLPGGTVNPDKLRTDERAVAFVRGFVESGKPVGVICRGPLTLIEAGVVAGRTLTSYPSVRTDLRNAGANVVDEEVVVDGNLISSRSPDDLPAFCVAIVKAFAKTPAGTS
hypothetical protein	MAQDTENQHTTGTQTGTENQGTTENDSAGSQAPGGNPDNRETVQNMDAPELRTAVIELLTIIDDSDEMDVVVTISDDGTVVVEAA
ThiJ/PfpI family protein	MTHELRGQRVGILAADGVERIELEQPRGALHGAGAMTEVISLHPGEIQARQFDMEPAGSFPVDRLVGDVSADDYDALLMPGGTVNPDKLRMNADAVAFVKAFADTGKPIAVICHGPWTLAEADVIQGVRMTSWPSIRTDLRRAGADVVDEEVVVDRQFISSRSPNDLPAFCGAIVVQFAKAKQPV

						hypothetical protein	MATKALLVRLEAKPGKEDAVEKFLLSALPLVEQEPGTKPWFAVRFSPSIFGIIDAFPDEAALEAHLGGAVGKALAEKADELFASPPDISKLDVLANKL
GI2	14745	61.7	Contig 17	84910	99654	hypothetical protein	MPGITRSSAHSTSSIAERLTLRPDTTRTASRNTDGAADAGRPRVPEAGKAAGRESGSGRSDGGSGVDRVPAGTNGLVICAKSVRDTVDGIVDAELSASDALGLRFPRPNMVASAAALFEPAEPWAHGRDRLDVEPECGRFRITIGPGVVRLGWTNPVRAEKASERTVGQHMQDVAGEVDRIKAGRDVPDPSGRAITEWSRKSRAAMCRTFAELDYTPMVESGRVPAMVTLTYPGDWESVAPDGASVKRHMMLWRKRFQREWGEPARYIWKMEFQRRGAPHIHLWMAPPYTVGRSGRRFSEWLSQEWAEVVAHPDPEQRTRHLRAGTAIDSLNGLRACDPKRLAIYFTKHSSPNRLGDKEYQHIVPEAWRRPGRGPGRFWGVHGLERATAVVEIAQDAYLTARRIIRRWSRSQAIHSDCTHCFPTALVPRMTRVAVQRTNSKTGKVRHRNVCRRRLLCRQGGFAGGYALVNDGPAFASQLAAALSADRAKAVRDEQLAGRK
						hypothetical protein	MHRQPQLLNDCVHLTSLFSVGKSEFIKLPTIFMILQPQE
						Transcriptional regulator, TetR family	MREVPRKIADRLPPAAALFADRGLNDTKIEDVAATTGIAKATLYYYFAGKEEILAFLLEDVLQQVAHEVAAVVEAEGSAAQRLHAVISAQLRVMAQRPAVCRALIGELGRAARMPVIADMISAAYFAPVETLLRAGAGDGSLVTLDNPAVAAIALFGAVTTSALTYLVLDDALDEDRIARTIHDLVFVGLRPR

putative cytochrome P450 hydroxylase	MTTVQLRYDPFDATIQDDLYPIYRQLRDKAPLYRAADSNTWVFSRHEDVVSALLDHHTYSSVDGIFPTPPGSDFHASFLPMMILMDPPRHDQLRALVSKAFTPRRIAALSSGIEDLAEDLSDCLDRGAGSVDFAADFAAVLPAMVIADLLGIPREDRAQFRQWSTTMVQSNPARGEMGEGLAAAAAVYAYFTDFLAERRREPRGDLMSALVSAEIDGKRLTDEDLLGFCLLLLIAGHETTSNLLSNAAVVLASDPDIRRRLVADDNLLGPAVEELLRYDSPVQGLSRTLTRDVTLHGVTMSVGDSVLLLFGSANRDERVFADPDVFDIGRKPEHQVAFGRGIHFCLGASLARMEARIALRALLARVPNWEVDLENAQRLRSGLIRGYMSLPISWSAN
Transcriptional regulator, TetR family	MERDGFDMPVDDDVDPRLVRSRTRLLDAATKLLSAGGIEAVTIDAVTKASKVARTTLYRHFSSSTQLLAATFERLLPQVHLPPATGSLRDQLIELLSRQATLFQEAPLHVTTLAWVALGPTSNSTQETYDRHALRTRIIEQYRQPFDLLLQSPEACADLDEFDPELILCQLVGPVAFARLTGLRAIDRQDCERIVDDFLAAHRRKVDEPAS

hypothetical protein	MADNGWVSSGVATLSRPVREQLRKVAAGEGAYSDRLARLAGFSIRHKVLVIGAWVVTAGVLAMLFPQLETVVKQQSVNLVPPDAPSLQTVDRMGTAFGEQGSKTTVFVAFEDPTGLTAPVRQRYKTMVSRLRADSQHVRLVQDLLADPVTAGQAMSQDGKAWYVPVGVAGTLGDPRAAESVRAVRAIVAESFSGTSTIVRVTGPPATFSDLIDSAEQDLIGISIVTAGLIAVILLVIYRSLVTALLPLLVIGVSLAVGRGVLSALGESGMPVSQFTIAFMTAILLGAGTDYSVFLISRYHEQRRQMISADLAVINATATIGRVIMASAATVAFAFLSLVFAKLSVFSALGPACAIAVFVGVAATVTLFPPVLALAAKHGIGEPKADRTRRYWNWIAVAVVRRPAPLLAASLALVLGLAAVALTMHISYDDRQGQPATTTSNEGYRLLDRHFRKDIVITQFMVVESPTDMRTSKGLADLDEMASRVSQLPGVTKVSGVTRPTGARLDQAQLSWQNGQIGNKMAGAVAKGDAHKDDLAKLTSGADQLAGGLAQLDTTLRTALTPLTAVLTQAQSSGSQFQHFRPLLQQLSATTPTIDQAIRTGPGLRQEAQQAQNAIATIDPLVGALNTSPWCATTPECAQIRDQVQILVTLRDGGFFSQLANLGDMYQPGSDNAAGTVADLQSTVTSLNKAFGALGDPADMAGNIRRLQNGISQLASGAQALATGVHTLADSNIEMLSGMSQVATQLQNSARSSAGSDNASGFYLPANTFENRQFADVAKHFMSADGKTARFAIESSYNPYSSQAMDLAQKITEVAGAARPNTSLANATISMAGFPAVNSDIQRLLSADFHLLAIATLVIVGVILVLLLRALLAPLYLLGTVVLNYGAALGIGTLVFQYGLGKEIAWPVPLVAFIILVAVGADYNMLLISRLREESAHNIRVGVLRTVANTGSVITSAGLIFAASMFGLMVGSVGIMIQVGFIVGCGLLLDTFVVRTLTVPAIATLLREASWWPQRKSSTHNGRPHRTT

Hydrolase	MADCYRHAATHPVVLLLHGGGQNRHAWATTARRLYSHGYTVVAYDTRGHGDSDWDPSGQYDIERFVSDLISVRGHVSADSPPAVVGASLGGLIILATHLLAPPDLWAAVVLVDITPRMEFHGARRILSFMAGHPDGFGTLNDAADVIAEYNPRRARPENLDGLHKVLRQRSDGRWIWRWDPAFISSNFDVLQGNLMTGSEEFDAISGFLAEGARRITAPTLLVRGALSDVVSQETVSEFRQLVRHAETTDVTGTGHMVAGDNNDAFTAAVTDFLDRAMRTLT
FIG00825226: hypothetical protein	MGLTDASKGAVMHTQWIEQCTVQRVSLHEGLILDLDGYNELVISRPLRLTLPPAGAWPSDEVLIDPINLSPEERPLLDLAGAICTRAWCDDDGALHLCFSPGHRIDVDPDVAATSWELYGKCHGYVACLPRGRVRAIRHDLPVDENDSDATQPKAMAHQ
hypothetical protein	MAITIPSGRPNWRFMRYVRPPTRDSKLEPLYPLRPATRPVRLGIDVGTIAEPPEEGYITGFLTRDEIEVHLLIPAATEAPSGWTELLDEPPCHTVNFTNVADAGKFCDAAEFSVSTARGESYRAWSKARFFAAYQQLDEHDAPDGLPPLTLDQRHRAAAYAAAAGAVGIDAIVTTAPTAGRTDVADNDVVVSVTPDAAVPLIGHYLRVTSNPVVTVERGMLVGGGSWETTESTATIVNLYDWGTVSGLPYFDAASMFAAAAKGGPEAAEAFTSVRIRLRRAARAFDDLLAALSNPLDGKRNEDVAEATAEAFDRELLYLAAVFDIFGRAYQAMVDPSVDRKKARGSLDSRTFIDKEVRTQYDQSLLGDVTRLRVYAWLCKQLRNHIHDGVLAVDTHPGRSYGNTMNVALNLSVIPELALGADNEMTQHHYDALGVWQTEPVSPFTGSSMVADLATTGFTLIRAALEYIEAFTKLIVRNKPANAPSSSAFLGCVQARPGEVEPAPPKRAVFYQALFGLHPDSV
hypothetical protein	MAAVDQSDNPMGQVLVNAGETFYLHRQSGFLKDFAPNTVLKRFA

	Transcriptional regulator KorSA, GntR family	MLQLGQIDRADDKPPYRQIAGMLREAIRSNRLAPGERLPSETELIEHFGVARMTVRQAVQELRSEGLVVSQHGRGVFIRPTPPIRRLASDRFARQHRAAGKAAFTVEAEKSGYSPQVDNIAVSREKPNSFVAERLRLSPDDDIVVRSRRYLANGRPVETAVSFIPAAFAEGTRIEQVDTGPGGIYARLEENGHVLGHFTEEVAARMPTPEERRQLELEPGVPVLTVLRTAYDANDVAVEVCDTVKVASAYLLEYEFPAR
	FIG00831412: hypothetical protein	MSVPKWLKVSHQQVFPRKAFVVSDVTAVIDYERSTKDNKVQATDRESGAPLWQVEVLDGDPTASKRARTLTVKFATPKQPVCPPNTAGMPFTPVIFEELYILPYVDRSSESGRIAWSFRASGMTADTGKAGSSAGDRVSA
	Transfer protein traSA	MSPYDPVVPSDPYSIPDPDAGIDFFDLTHLLAVGAMWLIAVIGVVAVSLVVWRVRSPATFGRYVATPACRARWLIWASISWPRVAKACGLSTPEHVTRTDAQGKSRTRTVWTHPRLLGVSMFGDCLRMTVRTRTGQTVDDLENAVPAIRDAVGAHSARSTLIGPGTVRVEFVMKQCLSEAETAAFPTSVEPTGVEVGRRENGSAWTLRVAGLHTLTVGCSGAGKGSVFWGIAGGLGPAIKAGTVRLFAVDLKYGIEVSVGSALFSGIATTEANAARLLTNLEELLDSRGRRMAGRARSHTPSTAEPLVVLLIDEFAGLTAYMTDAALRKQVAGSLSRILTKGRAVGIVVAAFMQDPRKEILPMRGLFTQTVALRLRSRDEVAMVLGDGLADAAPAHRINPNEPGTGYVIAEDGSTMRVRADFWPDSLIRSVAQEYGPTSSSRAGSAEN

GI3	10251	62.8	Contig 22	38227	48477	Coenzyme F420-dependent oxidoreductase	MAKLSVATPVVMLPAANGSWEKDASIEDLAQIAEAADRLGYHHLTCAEHIGLPAAERERRGTRYWDPLATFGYLAARTQRIRLTTYVLVLGYHHPLEIAKRYGTLDKVSNGRLILGVGVGSLKEEFELLGASFDDRGARGDDALRALRASLSVPEPTYHGEFYSFSGMVVDPCAVQDRVPIWIGGRTLRSLRRAVTLADGWAPFNVSPQQAREWLGRFDIQPGFEVLLPPLAPLNPIDEPDRTRHIIADTVAHGATIVSTMFRHRSLQNYLENLHALAEIYPPDGGAA
						hypothetical protein	MNRVKAGVFSLTPPAPADDDGSYLRWHLLDHMPEQYQLPGIVHGLRWIADGDYADHRLAANGPLSQIGNAVHYLVSDPVEETLDDFVALGSALRENGRFPISRPYLQVAGLRLLQWHSAPQALISGEVVPFRPHRGILLIVEEPADGRSNEWLQWLHAEHYPALLATPGTVGAWTFGTSTGWIHLPAGWRTDQQYITVVYLDADPLATTDAVAALVEDRWRSAAVRPVFAGPLRTMISWDAWPTS
						Transcriptional regulator, TetR family	MTDQEPRLRQRTDGRLNRSRDPAILNAALAALAEHGYHATNMDDIAARAGVGKAAIYRRWSSKAALMTDALVYWRPDLLSADVPDTGSLAGDLDALVERIKRSDSDLVSNDLVLRVAMEAAHEPELASALDDLMLGKGKRTMTAILTHAAKRGEISGDRDWSLVADVLVAMGLLQVARGQTVDANFVRQVIDTLVLPAVRTTNNA
						putative cytochrome P450 hydroxylase	MDPAMFFGTDAIQNPYPLYEQMRAASPVHRIGDSDFYAVCGWDAVVEAVARVDDFSSNLTATVVFEGDTGSVEPYSMMPLGDPMHSLATADDPAHAVHRKILLPHLSAKRVRVIEEFAQETARRLLSESLVDGRIEWMSTVANRLPMMVVARLLGLPHDDVDTLIRLGFATTTLLDGLVTPAQLEAAGTAVMELSAYVLDHFDKAATDAQPGLMSDLAARCASGQLDQPVALSMMLTLFSAAGESTASLLGSAAWILVSRPDIQEQVRAHPDLLGPFIEETLRYEPPFRAHYRHVRRHTALAGVELPADAHLLLMWGAANRDAAQFESPDEFRLDRGGAKGHVTFGKGAHFCIGAALARLEARVVLGILLERTDWIEAAEIGEWLPSILVRRLERLHLATR

Transcriptional regulator, TetR family	MAGVDWLVGQDRGSEATARIHAAAADLVSERGWEGFTIGALAAKVHCSPATIYRHAGGRTAIRNAVVNIHAARVIESMRDAIEGLTGAERVVSATAIALQRIRSDPLTQMIRSTRPPIDDDWMPSSEVVVQCAQEVLGQHDPDPLAQQWLIHVFLALWIWPMKDAEAELQMLQRFLGPPYDAAGS
hypothetical protein	MMQGDLEIAYEIIGEGEPYILTGGGRFSKDAGGVRELAEAIADRGKKVVIYDRLNCGASAVCFEGSSESVMQAESLARLIVDLDLGPTVLLGGSGGARISLLAAARHPEIARAVVMWWPSGGAFGSMSLAEVYNFPSIRAAWNGTMADVADDPTWAEVITRNPRNRDRFLAQDKQTFIETMERWAHVYCACGNPLVAGLTDEEAAAITLPVLVFRSGKSDYFHTRATSERLAANLQNAEIIDPVWGDREFLDRYEDSRDNGVSVWINWKQLAPQIVDWTEKVLD
L-carnitine dehydratase/bile acid-inducible protein F	MIKVMEGVRVLELAQYTFVPAAGAILADWGADVIKVEHPVRGDTQRGFLYMGGIKLEADRHPLIEHPNRGKRSVGVDVSTPGGQEVLYELAKTSDVFLTNYMPQARQKNKFDVEHIRAANPNIIYARGSAYGDKGPERLVGGFDGTAFWTRSGVGHALTPEALGGALPQGIPAFGDSIGGMNIAGGISAALFHREKTGEAVEIDVSLLSTAWWAAGASVTQGMETGETMRSTMPDETPPIFNPFLGNFLTSDGGTINLCIVSPTGYIRDTFEHLGLPELADDPRFSDVMPLIRNAAEAAELVAKAIRGKSFDYWREHLKTMKGQWAPFQSLLELASDEQAIANDMIVEYEASAGDKPFRVVRGPVQFNHEPLETTRAPQASEHTEIVLMELGMEWDRIEALKDAGAIA
hypothetical protein	MDDPVTDVEDTVESFPGIESGGKSTEVGHFMAAMRRLQDVVVSTDPDNELWGSAAREVNDLCERFELHRVPAGFGLRGRGPHLPGLGHPLMPPWMMTEYGPDGVTMQGHFSQYHVGSNNAVNGGVIPLLYDWQFGMIVSAVDRRDSRTAYLHIEYRKVTPINQPLTSRGRIETIDGRKVFVTATMTDSDGSVLSEASGLMIQLKSNQP

						Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	MYPGHWGITKPDHPAVIMAGTRETITYGELDRRSRCLAGFLWSRGLRRGDHVAILMHNDVRYFEMLWAALRSGMYVTPVNWHLKPDEVAYVIGDCEAKLLFTSARQSDLVEDIDWRTLPALTTRVMADGTTTGFEPYESVVAFEGEPAEESSGAAMLYTSGTTGRPKGVLRPLPDTKPGDLDAASHRLAALFGIDEQSTYLSPAPQYHAAPLTFTLAAHRAGATVVMMEKFDARDALKAIDMHGVTHSQWVPTMFVRMLALSPEDKQAADLSTHSVALHAGAPCPPEIKRQMIDWWGPILVEYYGGSDGGGFTVIDSLDWLAHPGSVGRSRTGAVYIVDEDADAVVGPGQTGTVYFSGDPGFVYKGDPEKTERSRRAGGELSTMGDIGYLDADGYLYLTDRKEFTIISGGVNIYPREAEDVLILHSDVQDVAVFGIPHHEFGEEVKAAVQLKPGVTATPELAGQLIAFTRARLASFKCPRSVDFLEQLPRTPTGKLRKADLRSAYAQGNSTP
						hypothetical protein	MANKMFRKYGYDAVTLEQIADACVMSVRTILRYFNTKEALALAGEHELLESFRAQLGVRETDAVTCWRNFRNETLPLMETPESRERMRAIYETPPLLAEFLRIGEGYQQCLTAAIAEDWGEVASLEATVFASVLVSTTSAAFRRWYVSEEPFDLVALQDIVDRVCVAFDKGRQSADGRRVGRVSAANTAPTRRAPTGRR
GI4	11329	61.3	Contig 22	102354	113682	hypothetical protein	MKMYAARILSIMRLAGGISSPALGRTNFTSSSKRSSFNSFQIWILPPGKVL
						hypothetical protein	MREGFSWATATDSGRGSQRRSWAAAVDIFVDMAAADGGAATSAHFSATRIARLAISIVQVWLYEWWHGPRDSAIEEDAERLVSFLSLWLRDED

hypothetical protein	MYRWVAAVMTAGYIVTAAGTGLKIFPTTPITGLVACIIFTNAIFVAFWACWFLEGPGKKRTTLEKADEFTMIWACVAAVGSEMTWELPWLVADALGRAHITPHDRWVYIWYYYARVDERYLISDGALWGMEAVVILAAVLLLVQWFRLRSAPTHDPKRISALWWSLLAMVVMLTIFVVYYAAEVRHGFANVQRGFWDITVIFVYENLPWLAAPIISIPLTAKQLAYLYREQGRRTALAEEGALADDRADVRSLV
hypothetical protein	MHASAPSRKVADLAERLQTIHFGREQSAEGGTIVSPSARIGDC
Protein of unknown function DUF1486	MTIFKQEIKAAMTPDMTTDSALRRSREELVLRHVAAENAHDIDAAMATFTHPRYEIMPTGMVFDGAESVRTMLLQQWAQLPMLQYSAEAIYHGEHGLVVETRTTAPGRSIDMLSMNLFGFRGADLVLERCYFDQSLLAAQLEALAPD
hypothetical protein	MDGPSALRRRRFKFTETPVYGTDDRSLGLRPVGVPAEASMQQ
hypothetical protein	MTRLTRNGSMISIEIDPTRVSMRSKGGATPSEHRRWGVRMLGPLAHSLSRRRIHE
TesB-like acyl-CoA thioesterase 4	MASPQEPYFTENGKTFVPSDVARGGWGPTVGGHAIGGLLARAVESERADETLIPARFTVDILRRVALAPVRVEASVVRSGRRMQAVDAVMTQDGELVARAATLFLRQSEQPQEIPWTPTRSMPPLPEELTALPKGVPMQIVPYNDGVEMSGQLPWQQDGLRQAWIREIRSLVGDEELTPFIRAAMAVDVTASMTGFTRTGLGFINADYTLTLCRLPKGSFVGLAGLTHHSAAGVGVGAATLFDAEGPIGTGVTTAAANTTFRGGAL
Transcriptional regulator, TetR family	MRRSNLNAVARRANVSRSTLYRRFPTKEELLYAVVGRLTESIFAELQHHTTGMTAKQLVVEVFCIAVRETTSNPVLHQLMVAEPDTLAALLGFFGPGMSAVLDRAVELAVAQAINAGAKMPPADLRIVVELMIRLTTSLMNSPSAILDADDPAAARQFAQKFFAPMIW

PROBABLE CHOLESTEROL DEHYDROGENASE	MGRNLILELLAQGHQVRSLDLTRNPDADSRIEVAIGDICDPRVVAEVVDGIDTVFHTAAVICLLGGRAATSQYREHSFAVNVEATQSLVSHARDAGVSRLVYTSSNTVVMGDEPIVGGNESLPYTTRFADLYTETKVVAERLVLQHNGSGGLLTCAVRPSGIWGPGDQTMFRKLLEAVASGQVRALIGSRKTVMDYTYIGNLVHAEILAANQLGPNGRASGQAYFITDGEPVNVFEFARPVVEACGQHWPRLRVPGGLVRAVMTWWQWLHFRLGVTAPPLEPQIVTRLCVDNYYSIDKARRDLGYEPMFTREQGLASSLLYYTDMFKTMKHQAG
hypothetical protein	MQNFNRPVAAGRKRVYVAMFVLGCLVTGIPFFGVGFGLMPPEPTYEIWGALIYIPLLVPLFTAWRDAPGENGTWTERAAEFTMVWFPVTASSQITWETPWLIGDLTGVMHGAGPQDRWVWLWWGYGAADTRYLTSDAGLYGMEVVAVIGGLALFTAWFYLLRAKHTGDTMKRIKGLFVGIVACSMMLAVVTIYYVSETRAGFRNLEQGFWEGFMFKFIFMNIWWVFAPLITIPFMIKQIDYLYKSMPRSIHQDVQPGAELQHAV
hypothetical protein	MPNRRSDALPGIRGARREMYRQQILTAAEYEFARTGFAETRVATIAATADVSLATLYKNFAGKDEIWNALVEQRTNELVGLARAATVGVQSPLERIIVGARAQVDFFAQHPNFLRLNVKENWSWATATEAARGGQRDAWRVGIAVMVREAEAAAASGELRHLRPQIVAGLAVSALQVWLTDWVNSGVQRSPSQIADELQDYLRRLLVVPRATDTTPLRD
hypothetical protein	MVRATGVIACSVFTIPMMIVFVTAWRDAPGENRTLLERANEFQTIWFFAAAAGSEITWETPWLIGDLMGWMHLTPKDTWGFAWWYYGDVDQRYLTSDGGLWGLELAVVALAVILLVQWFRLRSAATHDPKRINALWWAFFAMAGMLSLFATYFLMEVRHGFSDFKLGFWEITVILGFENLPWLIAPVISLPFVAKQLTYLIRSAPAATGTPLRSGAVSSSAAGL

						hypothetical protein	MNHINIVVEEFDAAVMHFQRVFGAQLVLDLPQPEWHACLLDIGDVLFEVFSPPQFLLNARYGPHYVGIEYQTGMDAARRAVTMRNIRLVRDIGHAIHTHPADTLGIAWEFFDGDFRVMPNLKWLEPLHSPERWRAHPIGYTGLRRLSVAVHDCDRAAQFLADFFGAMIVYKESRKSLSATIIGMLLADTTIELISPTGEGAISEHLHRYGDGIRSTVFEVADLERTHAHFKAIGIDLVDGDGPSTLAIEPRDNLGVMMEFCAS
						Transcriptional regulator, IclR family	MVNLLAGLRGETLSISEIAGQLRFNLSTCHSMLAELVKAGVLVRHPGTKRYALGPTLVTWGVASAFDSYRVLEFAEPEMSRLRHQLHVSCVARALVGPDTVVLARRDTDGPIAAFTPVGYRLAAVPPIGAEFIAWESDATIDEWLNRPQRALDTTERTRMHALLAEIRRDRYRFLWADPAEQTPLATELIQTSASEDEPGGDRLFELLQRRGYGRGPRTGTLSAITVPIFGADAKPVLALLLTFNPVPIDISEVEHYLRPLIHTCQRITTTIGGRAPN
GI5	14057	58	Contig 34	10279	24335	hypothetical protein	MPVAIGNPCVEEHVASENVESIETLRRNRVGRRARTSEHESDALE
						hypothetical protein	MQFSQRRLAVSRWGGCPTYWSTKFIASLEIRHKSPFEQ
						hypothetical protein	MFWQKVMINVACVEGQDPDVFDGGRVIAQSRSNRKGEALHHLVGVRLVLTERESD
						hypothetical protein	MILDLIQTVAAVSVPIVVAVIGYKLNHRLKLYEASQWRNQELIKARLQYFGQLAPMLNDLMCYLTFIGRWKELTPPDLIAIKRDTDRLFYSVAPLFSLAAVTAYQDFLGACFTTHTGWGTDARICSGFVRRREASREPWLAEWGQLFTHQDGDAIPESSMTAVRATYDKLLATLVEDVELLQPRDRYADSNLVANAR
						hypothetical protein	MPFHPEVSISRESAKFARNYGHLKEETVVINLPGVHRIW

hypothetical protein	MDALEGPSKISDSRQVALPKRLMDSLGWDKGDYVMFRLSDDDPEVMKVVLESVVLRQLRRGEDAERTMREQRAETTTDESPQDRG
hypothetical protein	MTGLVRDPNVGFFDDQLTREAVVKLAACNRLTFFVGAGASLDLNLPTWSALVSDLLKDTISANFGPDSPVETICAELVNQLFQVPAASVIDSLLFDEARASLSGRPARSTIHAKMLSTRNERLRKSLYKRDRFGDITVGPTLVGKMLELAILLRRAGVDVHIVTTNYDNAFEEAALHEPLFSIMSTERLRLVLFAESPPEAGDLDEGDIPVVHIHGLLRAPSGYAAAGHNPIEGKIVFSEVDYIDWETGDFRTYLTNRVQSGGLLTIGASLRDNNIVARLRDGTRSNNEPRYALMPSESEFKYLSDKNIAEKYWMPFVEMAARRGFIFGVSILRPDFYGQVFQFLQELGVNVSSRINPGTAYVEYRARIRQWAVAWSTVRESDEPGMRGNIETACRHLAVKFEALPQFDHCKVEVWARCEPDDRTMRRWCSSQSTWRPDAWPHTTLIRPLSKQNAVMAFSSRGTVINKTPGSADRRWTHCLTVPVFLDGEPWLGLPVGALEILLHESGAATSDECLNYLRANATPWVKEISKLGEKILTPEPWVD

ATP-dependent DNA helicase UvrD/PcrA	MTFKLAGKAKDAYEQDGRVMVLGGPGSGKTTLALLKAKRLLSSLEPEQDVLFLSFSRAAVRQVVLRCRDVLSAPERDRISVQTYHAFCMGILRAHGRLLTGRPARILFPDREKVAKAGFEGDWNTERDRLATEESIYVFDHFASGAARLIGDVAKVAELVADRFPVIILDEFQDTSDSQWELVQALAVRSTVIILADADQRIFEYDDNVDPNRLNQVREVLKPAEFDLVGENHRSPDAGILQFADAVLRNRALPETDDVQVISVYPRALESTAHAAVRWLYSRLRKQGIQRPSIAVLARANGLVSDVSDWLSTARTYNGSELKPVAHDVLWDVELSAAAAQIVASILEWPLHDAGRAVAQTLNAVANFYDVKNAINNTPIASAVRARDSYRANADSLTGGGTVRLKAAKHLVEQYPAMSAYGGDPAADWKTARDILAGADALKEIVQAARFIRLFGATDEIGGRLSEQWDRSGGYDNAVELVRRALDQGRLVTELREPQGCVLMNIHKSKGKEFDGVLIVEAQYKGVLFDGREEPPHMSSRRLLRVGITRARHKVIIVRPKRAPPLASP
hypothetical protein	MKACYLHIQNFRGIEDGRVSFQPHTLLVGGNNVGKSTICEAMDLVLGPERLYRRPVIDEHDFYHGKYLDEDDSPLDVRIDLVLTDLTSEERRRFGDQHLRLWDNKTDSFIDEEAGGSELADAESVEWGLPVCFIGRYDKEEDDFVGETFFCHPEPVQDELDVDQIASLGGGRAVFRRTHKRLAGYVYLRALRTGSRALSLQRGSLLDTILQLGGEGSAEMWKDTLNHLSTLEPAIGDVPQLETVRELLRKRLGAFVNLAPGEHSAAFFASDLTRQHLREVVRLFVATQPSEHLVPYVRQGTGSINLLVFALLTIIADLKGSQSVIFAMEEPEIALPPHTQRRVTRYVLQQMGQSIVTSHSAPVIEQFEPESIVMLHRDGTKLAGAPIDLTKIKRKTYLTNRRQFAEAILARGVLVVEGSTEAVVFPAISSVLERVRAGYTHLDFAGVSMFACSGDGDVDRFGPIFTALGKKSYGACDKPNTAPSADVMANRAAFDHFWESAESGIEHVLVRGIPVEILRKFLDAVSARGDYPAAHPYNPAMSDDDVPALAFKVLKARKGEAFAYAAVLIEQCETEADLPPELITILITIDDELRAEPEEPIEPESASESTEGE

						hypothetical protein	MVDWIRLTGGEHPDVAAQLSRAGLTAADAELRISFGRIDPTRDTIIMRVIRGHLRVNDAVRQVQEFRRSQATGS
						hypothetical protein	MDDKAARLLSCLAEYELTLPEDVARQLISDRLDGIAKRMRIGRQAARSYVTNDVIQSLADEVVGSRSGTDDEASNVVSLSDRRRSRRSNRR
						putative ATPase	MSTRKLPSHISQLILTDYRVKPESVVYAAFNQSQLDGSQPYVARRIIRTSATLDVLALVPEHLQVSRRVSSRGTEVLLFSGDAETVFLEYATEQRLEVYVSASDHNHATAIANQITSRIPEAAPPKDTVGVTVWHSGARGASPSLKSIRVPSWADIRRNYPGAVAEPLNDLFTTVRPHGYGKLILWHGEPGTGKTTALRALSREWKDWCSIDYVADPERLFFDTGYLLEVIAKTSSDELQDSNTARWRLLIAEDSDEFLRVSARREAGAALGRLLNLSDGILGQGSNTLILLTTNERLDRLHPAVVRPGRCLAQVEFTRFPAAEANRWLPENHARVSEPKTLAELIEHRDATKQIATGIAPVTNIGAYL
GI6	12666	59.7	Contig 34	30105	42770	NADH oxidase (putative)	MTVRHIVAVGGSDAGISAALRARELDPTTEVTVVVADAYPNFSICGIPYYVSGEVTHWTNLAHRTAADLAATGMRVFTDTRATGINVEDHTLDVLDPTGAPEQLSYDALIVGTGAISVRPPIAGISGPDALGTVDGVHLLHSMGDTFAVMDSLQQRDPATAVIVGAGYIGLEMAEALTTRGIRVTQIEALPEVLPTVDPELGTLVHAELERNGVEVLTNTAVSAITRTDTGALSVIATHDGQTISRTVDFVLVVVGVRPDTELAADAGAELGVKGAVVVDEAMRTNLPDIFAAGDCVHTHHRLLGITWLPLGTTAHKQGRVAGENALGGQAHFAGSLGTQVVKVFDVVAARTGLREHEALAANRGWTPVSSQSSPDDHKAYYPGATPVHIRITGDLHSGRLLGAQLVGHRSAEIAKRVDTYATALYHEMTVGGLSELDLSYTPPLGSPWDAIQMAAQTWARHHQPATATAASQPGTRRNLAAPN

dNTP triphosphohydrolase, broad substrate specificity, subgroup 2	MHAIIPNDPDQAAKRSTRPYAKPKSGGDARSEGQRDRDRVLYSTAWRRLGGVTQVVTPFSELPLLHNRLTHSEKVAQVARSIAEAKILSANETTQKLVQKLGGFDLDVCEAAALAHDLGHPPFGHIGEKILDDVARNKEILDLDDGFEGNAQSLRILVIGKIRSNKYEGNNPTYATIAAVSKYPWKRAGKRADHDEIIDTDSEYRKHWKKFSAYDPQYELLTQAREFLPPCISEKEIQSLEASVMDVADDITYAVHDLEDFYLAGILDVSWIREDLEKYIRGEQKTGDVFFELASRLGVDYSGWFNDDEMIRAAEWAERELKSGFNRRQVNYSEVEAQARKRGSDFIDTFINAIEVRDEPFWTGGPYVGLDERNWHRVQILKEITRSYVVQRPDIALLQRGQQEVLEKFVKLLYEWSENDRARLPARLKREIEIAEEQLDGGSTIGYGKDQHAPRGKNKNRAILDYICTLTDLECLQIYYKLSGIQVHRPGMVSF
hypothetical protein	MPEQERRGALINELKIVRKTGLHRLRERIDELPELSKLATATMGAGSADDIERMLRHAFTSYAEGAQGTAIGILFGLEIGRRGARPSVLREVAAKRLGYDSVETFRKRPEYNAIAYFADLLHRLAVDANSQKPADVNKIDHIMSLITELTIPEYWELTRRVRQLFAAGSQMS
hypothetical protein	MAQMDRQSIVRAREDFNRVREHLHQLELSRLAENAEADMREECTIHAMRNAREVDDAARHFEGGREGLAMTIREMQAAYNTGEIAHIAMRGCHR
hypothetical protein	MAALPCFEDLWNGVTCLAAAPPPVDPLIERLRQISQQLVEEKRAMDRRAQQVIRAMQAVASAAGQDGSTVQDDTSVERWVVTE
hypothetical protein	MNSSGPGHFVDRLIGWCFSILAAAIALYCAVKLLADILPALVVVVGALTLISLVVGAGIVVFRTIRNRW
hypothetical protein	MREAEADAEVKAAIADWAPDELQDEDNPSKGSDVEVGHD

hypothetical protein	MTDQLDVPLTADSLFWQQFFFQQPFQEAKAVGLLRHWAAQDHAPQLILEARSDVTGVEYLVGSQLRHAQQVRRSVEDLVEGSIVTSYEPGDREQISTARRLQLTTTTRPLEPSDPVAAGRSILHALTAVQRGERMVIQIVLGPRYQPLLGPPAPNRADQPVLSKVLHGVLPEQRPDSRQALAHKLGQHGFRAALRIGVDAATADRRRTLLLGLAAAIRTAEAPGVHLKLRPEKPREINAPRSTWSMFTRSQHLTVTEVARLTGWPVSDRDEPFPGQPPRHPKPVRPSGAVQTGERVVATANAPGAEGPVGYSVTDALRHTWVIGPNGVGKSTLLLNLIVQDLEAGRPIVVVEPKDLVRDILARVPEHRKEDIVLVDPFDVAPVGINPLDSRYRDGRQPEVVADALFGTFKAIHGDSLGPRSADILRNSLDVLARRDDASLVMLPLLLTNPGFRRSLTQHTMRDDPYAAGPFWQWFDSLSPEAIATTVAPLSNKLRPLLTKQLRAVLAQRSPKFNIRQVLTENKVLLVPLQKGVIGPESAQLLGALVVAELWQAIRERAGTSEGTRTPLMVYIDEVQDYLRLPTDLGDALATARGLGAGFHLAHQYEKQLPPAMLDAFRNNARSRICFQLRPGDAKEMVAGQSVLSVEDFGMLPAYHVYASLVRDNAVQPWASGITSAPPPKTSDPDEIRRLSRERYGQPLDDIETGFSELLSGAAKAEKSDEMGGPKRRRRQA
hypothetical protein	MRREDNHDHAGPIGDPIRLSNKPPLTSINDKSGLANDGAHTGTTPPKQTKKRRTQSHDIEALRERLSERDLAILRSVAEHQFLTVRQIEAFYFADHPSATGGRLARRALARMRNLRLLGATNRRVGGVRSGSAGMVHYIDVVGDQLLDGRSGRGSRRFFNWSQRFVHHRLAIADTHLALIEADRQAQLELVECLVEPASWRRFVGLGGARLALKADLYIEIATTPGSDFVNPWFVELDLGTESIATVLKKCRDYEAYRRTGTEQADGSGFPLVAWSMTHPDPTKAERRRLALRDAIDNDRNLPAALFRIIAPDKLIPLLTRGGAV

DNA methylase N-4/N-6	MGDALDRLRELPDSSIDMALTSPPYFRLRNYGEDGQIGLEQHVDQWADRLAAISAEIHRVLVPTGTFWLNLGDTYATHPSQGADRKSLLMAPERLALRLQRAGWTVRNKVVWAKTNPMPTSIKDRLNCTYEFVYVFAKQPSYFFDLDAVRQPHRSSLRKPYVAERRDAELWRGPNSDTATGLDAIKAAGRVGHPLGKNPGDVWLLPSSSFRGAHHATFPVRLASRAIQAGCPEARCTRCRLPWRRRVIRAIGGTAIRAALVPTCECRTTSEPGLVLDPFIGSGTTAIAAEELARDWLGIELNPDFAAMAERRILDARGRRQPEEKAA
hypothetical protein	MSTPDESAKGVKTLGIRLQPDVHAQLSFIAQLREGTITDEIQIAITGHIARSKEDPELRGRADAARAEIEREAAARQKAIASLFTTEPPGSSGDSDDTAGQSRPARRGQKGDSPPEN
hypothetical protein	MSPKAKSLQAIGHVLGIPASELFAIAPWVTPQELPTIRPYLRTKYRELPPAAVQEIETYFNDVARKHGISFDPNDGPLDGEDE

GI7	12727	59.9	Contig 35	75351	88077	Superfamily II DNA/RNA helicases, SNF2 family	MLLEELKPGLRIAGLIPGQVVTVIFAQSHGADTVELTFKTASGGLDQQLIFRKDENKLSVAQTGGRPFSAPASEFKLVAEAQRITLAGLYDPMLAVATSDVRPLPHQISAVYGELLPRTPLRFLLADDPGAGKTIMAGLYIKELILRDDIKRCLVVAPGGLVEQWQDELFFKFGLRFDLLTNQAIDAYINDNVFETHPLLIARMDQLSRNQELQIQIQDTEWDLVVIDEAHRMAAHYFGGKLEKTKRFQLGELLGRLTRHLLLMTATPHSGKVEDYQLFLTLLDRDRFEGKNSQNPETTGIMRRMIKEDLLTFDGKRLFPERKAETVPYELTQLEENLYEQVTDYVREGMNRADRVGGKRKNTVGFALTVLQRRLASSPEAIYKSLVRRTQRLHRKKQEILNGTYRETEPTVDLEGLDADDYNAEEVEQLEEELLDAATASQTVEELDAELAELADLTALAKQVRDSGTDRKWSELRTILEDHALITDNDGAPRKFIIFTEHRDTLDYLTTRIRTLIGKPDAVKAIHGGVRRGERRLITEEFTKNPETQILIATDAAGEGLNLQAAHLMVNYDLPWNPNRIEQRFGRIHRIGQEEVCRLWNIVASNTREGDVYTRLLIKIEEQRKAYGGKVFDVLGDAFHETPLRELLLKAIRYGDDPQVREKMHEVVDATVSEGLPELLKERALAADHLAEADLAELRAAMDEARARRLQPHYIELAFKAAFTRLGGRIAKRERGRYEIPHVPPQIRASKYQPIATKYDRVTFDLEHVRPESQTRADLLAPGHPLHDAVMEETTKLLGGALNYGTVLVSSTLEEPHLLVGVIEEVCDATGASISRRFGYAFVDSYGTVSHAGPAPYLDCVAAPDTPAVTAARQLPWLAEAEDRAMSWIITNQLPGYLAEVQPRRAAELAKTRALVIKRLEGERDRLLFDSAIAAEKEHAGDKPKESSESLNRKAVDLDVRLRKRLELLDQQELMSTKPPRIITAALVLPISMLEHEIPASAPVHAKETKDVERRGVDLVLAYEKALGRQPAEQTFTNKGFDILSTTPNGDTYRIEVKARLEGATDFYVTHNEVLLGKNAVPRYRLALVKVDPRGPEHDQVRYLDHPFATTELGDFEATGIRGDWN

Adenine-specific DNA methylase containing a Zn-ribbon	MEGTLSDSAPKRKLIEVALPLEAINRESAREKSIRHGHPSTLHLWWARRPLAAARAVLFAQLVDDPSARPDEFPTEELQRKERNRLHKLIERLVVWESIRNEELFAEAHAEILKSTDGNPPPILDPFAGGGTIPLEAQRLGMEAYASDLNPVAVLINKALIEIPPKFRGQPPVFPGLADSEIRSWKGAEGLAADVRAYGRWARDEAEKRIGHLYPKATLPDGSNATVIAWIWARTVTCPNPACGIDMPLVRSWWLGKKKGKESYIVPTIVDDRTAPGGRRVRFEIGNDPSTGPMSSNDGTVARTTATCVACESPVPLPYVRAEGREGRLRTELMAIVAEGQRRRVYLAPTALQSDAARIGRPASVPVGSLPPHGLSPDSHQTIRIYGFYEWADLFTNRQLVALTTFSDLVAVAHDRVLADGGPSDYADAVATYLGFAVSRLANKSSSTCSWDSSTKMEAVRGVFARQALPMTWDFAEANPWGGSGGDYLEDVEWVARAVERVPATRPGFAQQADARRVVMPANTLISTDPPYYDNIGYSDLSDFFYVWLRRSLRGIYPDLLSTMLVPKADELVANPHRHDGREGARQFFEDGFKSVFASARATATADFPITVYYAFKQSESDDDGQASTGWETLLEGMIQSGWEITSTWPMRSELSNRMLASGTNALASSIVLSLRPRRADAPTTDKRSFITALEAELPDALRKLQQGRIAPVDLPQATIGPGMAVFSRYAAVLEPDGKHMSVRSALARINEILDRVLNEQEADFDPTSRFAIAWYRAHGYNFGKYGDAELLANARGVVVQALDRAGVLISRAGRVNLLKPNDLGPNYDVVADLHTSNWEALHHLVRVLDAEGIERAGEFLRAALGRTDGIIDADLIKELAHLLFRVAETNGWTKDALSFNGLVTSWPEIVDTAHTKTPRTSQGSFNFEEDAD

FIG00863870: hypothetical protein	MALSNRDRINKMFEIVAPPLDDYISLVIGQGDPAVGAAWPKLVRAKDGTSSAKTYNAHDPQVQFRMLTESNITAGFKKGWYPFNKTLGKAGESFAIELREVRNNWAHNGMFTDDDAYRALDTGERLLRLIGAANEADEVRSMRLNLRRVTADKEDKKVLRAAVDNPEAAGLKAWRDVLPPHDDVATGNFAASEFAADLYKVAFGGEQDSGYANPVEFFQRTYLTEGLSDLIGRAVRRLSGDDNAPPVINLQTNFGGGKTHSMLALWHVAAGLPIGEFPQATQDLLSANKYSGGNVNRVAIVGNHFSPAGIERDGTHINTTWGELAWQLGGPEAYAYVAKADADRTPPGDALAKLLAAYAPTVILIDEWVAYARSLVGRDDLAGGTFDDQFTFAQLLTEAAKATSGVLVVISIPASESGDAADRVVAGNAEEVGGAHGLEALKRLQNVVRRVADQWRPASSTEAYQIVKQRLFTQPDAVALASINATARAYVDMYRKYTDDFPRESRDTAYEDRIKRTYPIHPELFDRLYEEWSSLERFQRTRGVLRLMSTVIHALWVGEDASPLIMPGSVPLATGSVNSELTQYLQDSWKAIIDADVDGPNSEPARIDKDKPLFGQRSLTKRLARTVFFGAAPTIGAAHKGLETQRVFLGTAIPGDVPGNFHAALTQLADRATYFYTGSGKYWYDLQANITRTARDQAERLHKEDVWAEIVRRLQSQARSRGDFTGVHVCPETNGDIPDTDEARLVILHPKVAHKKGTDSPAKEFAHMATEHRGSANRTNRNTLVYLAADEARLEELDTAARDYLGWTHVLNNAADLDLTENQKNQACQRQTTAEQTVTARLLQTFTWALVPAQPDPGAPFIIRETKVEGQSESLAERVSRRLGNDGDLSIRQAAATIRLAINKVPQIWKDGHVSLGSLWALYCQYPYMPRLRDRTVLNEGIVDMPMHWDVEAFALATAFDKSAHRYINLWISGDKNAAPAPTDAVLLVQPEVALKQREEEPPHGGGDKPPYGGEPGSAKPPSIDVAFPPAKKRFYGVKTLSSDKIALDFKNIADEIISNLREPSTSLVVKIEIEATDASGFDESKIRTVSENARTLKFD
hypothetical protein	MVTPGACDTVIMPIGPAATALRADTTTSFEGKRHFPAPVWAV

						FIG00828583: hypothetical protein	MSSHRTTFNHVGLCVSDRERSRRFYEGVLGFQFWWELDPPDGPTAQLVQLPEPLGVHATYLVRDGFVLELMDYSKRQVHAGSERVMDQIGLTHISFSVSDLPGALKKVEEFGGAVVKETVTEAMAMIRDPDGQLLELLSGGWLAALPLRS
GI8	37271	62.6	Contig 37	19212	56482	hypothetical protein	MRFTYFAAAVGIAIALLGASPAHADYSGFTRCVGGIKQLPLNEPDPQNFQRVGVIEQDLKSGVSPAAEAQKVAGMGFDQQTAATIVQCIIEENP
						hypothetical protein	MHGKQIAIAGAAIVLAGCGGSGKTATTVTSAVEITKTVTVTAPPPPFAPRTVIETDGTYRVGIDIVAGTYRSGGPSPEGGSDCYWARLSSLNSTHIIDSDIGTGPQVVMIAPTDKAFLTRSCQTWHKTD
						hypothetical protein	MMSGHYLKTTTDPGGGRAFTENWYFTPCGDGCADMSSPASGVSGRAMLVDGQWTLDSTEDIVCKGGVTEGNAANAHYTWDPNTLAGTVQVVQNRGVCDHPPQSYSLSFRLTKAP
						Integrase	MTIRRREAKSGTRWDVEWLLPDGTKRSKTFKTLRAAKLFDGEVRSAREDGEAVDPRGGKTELRLVYKSWLASRPDLSPKVRRGYEDNWRLRIEPKFGNWPIGRILREDVQEWINGMTADGLGPRTVRWTHSVLRMTLDHAVTDKKLRGKNPAANVRFPAMGETSHVYLTAVEVAKLAELCDTASGESSRAAKQGDVVLILAYTGLRFGELTGLNVEDVDLPARRIRVRRSITQLSGRLLEGPPKSRAGRRSVPIPERLIATLERRIAGRTASEPAIVSPKGARLGLENWKRAVGWRTRIGELGHPTMRVHDLRHTYASLARSAGADMKLLQVTMGHASIVVTAHTYADLYDSDLDRVADALDGLGD
						hypothetical protein	MEQALLQLLVRADPVPLSLLHFVITGTWPPEAMQPRFDVSFPKRSGSGAVPI

hypothetical protein	MPKKGRGRRPVWCSARCRRQASAERIAARNAGAAARVVEVPRAHRPDPNARLPLPSMHTLQRLFLSSDYQCQTLLEDLARRYTSGAMGEQLRAAVQRFVAAIALEQALADDPVYRRAREDVERLRERLCREVERAERRDRELARLRRDAEELWRLRARVAELESALAGVAHPLHQAGQHDQVPVSRQQRRAAQRAARKTC
Death on curing protein, Doc toxin	MRRGELWFAATPGGDRPVLILTRDPVADRIGAFVVAALTRTRTGLVSELALTAPDDRLPSDCVVNFDKFIRCRAPRSDAGSLACLRRGYTRCVEHFERPPAAETSIRTRALISLVCLQER
Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	MSISLLLEMAASTNPHRGAVVSGDARLTVDELSVLADGGAGVITSCAARHVAYVGTGGALLPLLLFASARAGVPFTPLNYRLSGESLRELIARLPEPLLIVDVAFAGQVAGVGALTMHSYDFITAARTAEPVSNFIDPDDVAVMLFTSGTTSRPKAVQITHNNLTSYVTGTVEFDSAAAEDATLICVPPYHIAGIGAALSNLYAGRKMVYLRNFDERKWLELIGAEGVTTATVVPTMLARIIAVLEDKPVELPTLRNLAYGGSKVALPLVRKALQLLPHVGFVNAYGLTETSSTVAVLGPEDHRAALCSPDPAVARRLGSVGKVVPGIEAQIRGEDGSILGPGQRGELFVRGAQISGNYADIGSVLDADGWFATNDIAEIDEEGYLFIGGRSDDTIIRGGENIAPAEIEDVLIEHPHVRDVAVVGTEDPQWGQTILAVVVPAEGTDPDPEELRAHVRKHLRGSRTPDRVVFRNELPTNATGKVLRRQLIVEFSSVTDEFA
hypothetical protein	MRINRDPAIDERLVPQLARLRGRDFFYAVDTIIAGLQARLAAKSAS

Butyryl-CoA dehydrogenase (EC 1.3.99.2)	MVNAQSAEVADIIKAVRSFVRDHVVALEPKIEADDDIPPQLRQACAEMGLFGTAIPEEYGGLGADVETEVLLAFELGWTTPALRAMFGTNNGIAGEVLISAGSDEQRSAWLPRLASGAAIASFALTESEAGSDPSTMKTSARRDGGDWIIDGAKRFITNAPYADIFVVFARTNPEVPAAQGISAFLVPKGTRGLQIGPRDAKMGQAGEVTAPVYMDGVRVPGEAMVGGEGQGYAAAMRSLGPGRIRIAALCVGMAERLLEESLAYATTHHQSGRPIAEFQLIQAMLADSQTDIYAGRALALEAARAFDAGTDRRTGPSVAKYFCSEMVGRVADRAVQLHGGTGYMKGVTVERLYRDARVFRLYEGTSQIQQLIIAGQSVRQYRNRAGIPS
Transcriptional regulator, GntR family	MPKRYGTKEKDQAVEYIIDRLLTGKLRTGDRVDRNEIARDLGLSRVPVTEAIGQLEHDGIVTSRYHRGTFVERFDPATVLEHHEVYGLLSGMASARAATYPTPRLIDKLNELIRVLRTATDSEIFEQTTWEYRRTINHEYAGPRLQAAIRTSQSFIPKAFWVNFEDNQRFMLPSYEREHAAILRHDSTAARAACEDRSMTTARVVVTELVRRGVFESTDAVAPVDGLLSP
hypothetical protein	MGHEEFRLLARDFIARNVTSYYDWITDGIFPAQSVPPTHQIGVMGFRDS
putative cytochrome P450 hydroxylase	MNEWYYFTHRDDVLNALRTPEIYSSKKAFDMLGSPLPLVPISYDPPEHTRFRKVLQPFFSPHALGKMMPSLQQQVIDIIDNVAAKGECDVVTDIAIPYPSQVFLTLYGLPLADRDKLVNWKDTVIALADLPGVPGEDKLGPALELFAYLTNAITERKANPGPDILSQVLTGEEPLDEAEAIGLSFLFVLAGLDTVTATIGFALEKLACDPALRGQLREDPKQTNVFVEEIVRLESAAPILPRITTEEVTVGGFRLPPGSPIRLCIGAINRDGSDDISTDRMVLDGKVHRHWGFGGGPHRCLGSHLARIELTLVVNEWLSRIPDFEVKPGFTPQIEWPANTAALISLPVQWKS

Putative GntR-family transcriptional regulator	MTQFGVSRPTLREAFRILESEQLIEVLRGARGGARARRPDPGAAGRYTGALLQSRGTTLDDVYQARAALEESAIGLASGRGLSRQIRELEELSTRGAGLIEDPAAYAEFDVVFHHSIMALAGNTTLSVLADQLYSIIEAHNRAHLADHPAGYDVPACRTHQRGHNKLVRLLIEGDLEGAQRHWRRHLDGVQKYMIEDSRVAVVDVLSAAL
L-carnitine dehydratase/bile acid-inducible protein F	MVDDLTGFEEAVALKKPMTGVRVIEVAEYTFVPSAGAVLADWGADVVKIENPVTGDAQRGLVSVLSRSASRQGVPFAPMMEAPNRGKRSIGLSLALSESRPAFEELIRRSDVFLTSYLPRTRAKLRIDVTDVRRVNPDIIYVAGSGFGHHGPDRDRGGYDATAFWARGGSAQSVTQIDDDSPAPMPAGAYGDNIGGLTIAGAVAAALFGRQNTGETSSIDVSLLSVGAWAMQANINLAMLYGCPLPVIGGRAPAPGNPLTGTYRCKDGRFIQLSMLQPTRYWPEFCSLFGLASAATDPRFASMEALAQHNDEAIRLVAEAIGSRPFEECNRLLDQSSGPWAPLQDAWEVANDEALIANGRIADIIDSQGHPQRLVANPAQFDGAPTRLNRAPLFAEHTDEVLRDLGVDETELLQLKYAGAIT
Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	MEPAPAPPRQFATWSQRILHHLLSDSDRDAIVSMNHERIESWTRAELVARTAGASDYLDVHGVQAGDCVPALLTSQPHSVALLIAGALTRRPLAPLAPRVTHREVFASLKNISGPVLVTEPQYRDFASELAKATGKRIAVVDEIGAGNVALGADARRPEDIACMLHTSGTTGLPKPVPVPEGPFERRAEILGQLCGFNGDARYVTAALFHHVAGLGNLMPALANGASIILIPPLSVDLWRNLASAAPTHTTLVPSVIEMLLEEGVPAPPSLRVLAYGGSSIHPDTVRRMQKVMPAVDLINLFGQTEGSPVTVLNADDHRRAAAGQEGLLSSVGRPAPGVELRLNEPDANGIGEIWARSDHSFVVDEQGWQRTGDLGRVVDDFVYLVGRRGDKIIRGGENVFPLEVEQILETHPDVVSAGVVGKPDRRLGETIVAFVVPVDVEEPPDSENLRVHCRAQLAGFKVPVEWIFVDRLPRNPNGKLLRGELSRELARRTSCGAKGHRVIWQ

Alpha-methylacyl-CoA racemase (EC 5.1.99.4)	MGPLRGTRIIELAGIGPAPFGCMLLADLGAEVVRVERAMSAAALASIDDGGRSQADPRRYIPHRGRRSISLDLKSAEGRDALLRLVETADVFVEGFRPGVTERLGVGPDDCLARNSRLVYARMTGWGQSGPLASTAGHDINYIALAGVLDNFRRTGERPMPPLNLVGDMGGGGMMLAFGITAALVHAARTGEGQVIDAAMVDGAALQMAVVLGGRAQGRWPGEPGSNYSDTGAPQYEVYECADGKFVAVGALEGPFWSEFLRILDQPAGAVPSRDDPALWPQGKEALAKVFRTKTRDEWAAMFDGSDACVTPVLSIDEAPAHPHAKARSAFITVGGIEQPAPAPRFSATPADVPSAPPDRGEHTEEVLRELGLTDEEIVAATGSQRV
MaoC family protein	MTETGSVVVGYSPPPRTDKPLKVTDFVRYQGASGDMNPIHHDPEYAARAGYDQVFAVGMLAAGRLATYVTDWLGPENVRKFGVRFKEQAWPGDELTYTATVVDSRTDENERVVDLDLVCTRHTGGVHLTGWATVALKTPTTDETDDGGPPSSD
hypothetical protein	MAEPFDIPIELGKSREFAKATGSKHPAHYSPGGVTPPTFLMTAILWQGPDNNAWPKDRDMRRVLHASQEFAFPKGPPRIGMQLTGQQRIEGSHTKEGKRGGTLTFTTLLTEFRDEKGDLVAEVRNTFVETSKPTGGDH
Alcohol dehydrogenase (EC 1.1.1.1)	MMKATALVLETPNHLIERHIEIPAPAAGEAILRVLACGLCGSDHELFTGAMPGSLPLIPGHEVVGVVERATDEFLSARGLEQGEVVALEVFQRCEQCDACRRGAYPLCRTYGLARSYGNTSIEWGSGLWGGFATHLLLGCDALVHRVPPGLDAAYATLFNPLGAGVRWAHVLPQVQQGDVVAVLGPGLRGLSSVAVVSIAGAAFTLLTGAGSRDRERMELGRTLGATETVDVTTTDARALLKERTGGLADVVVDVTAAAPAAFLQAIDLARPGGTVVVAGTRGLHALKDFNPDRFVLKELTLLGARGVDGTAYARALELLATDDRFEAIPRVTCALDAGSVADLLGDMAHGEAPPLHAVIVP

Putative uncharacterized protein BCG_1584c	MSEDEQRERHGKSLRFMFTKVPFNKVVGLELIDWSEPDVAVVRLPFSEAIDNSGGTAHGGAIATLADVAGSAAAWNGHDYEKGTRGATVSVTINFTGAARGEAVIATARCVRRARELNFTEVSIVSESGRPVASGTVIYRIAT
3-ketoacyl-CoA thiolase (EC 2.3.1.16) @ Acetyl-CoA acetyltransferase (EC 2.3.1.9)	MPEAVIVEAVRTAIGKRNGALSGVHPTTLGAVVLGELVARTGVDPADVDDVQWGCVTQLGDQSSNVGRFTVLAAGWPETVPAVTINRACGSSQQALDHAAYAVMSGQQEIVVAGGCESMSRVPLGAGRATGMPYGPAAYDRFPGADFNQITGAQAIAEKWGLSRRRLDEYSSRSHELSAEAADHGKFDEQIVAVETSDGHFRVDEGLRRGTTPETLAKLSSVSGPDGVIHAGNSSQISDGAAAMLVTTPEKAKSLNLRPLARYVRGVAVGADPRLVLTGPIPATQQVLAKSGLTIGDVGAFEVNEAFAPVPLAWLAETGADPARLNPLGGAIAVGHPLGASGAILMTRLVHHMRDKGIRYGLQTMCEGGGTANATIVELFGD
Enoyl-CoA hydratase (EC 4.2.1.17)	MPELLSDAEAGVLTLTLNRPQSRNAVSYSLLDNLIEAFEAADSDPSIRAIIVTGSGDSFSYGTDLAAGGGGLDVNAPGFKPLRGTKRDVGGVLALRLFSCTKPVIAAVNGTAVGVGVTMILPMDVRIAADSARFGLPFTRRGIVPESCATWFLPRIVGIATAIEWCVTGRVFDAQEAAARGLVSELLPAEQVLPRAREIAAQIAEQASPLSVALTRQMLWRQLGSPHPMSANRLESQAFLSLGGSADTKEGIAAFKEKRKPHFTTAVPDDLPAFYPWFADEEFADKS

3-ketoacyl-(acyl-carrier-protein) reductase(EC:1.-)	MRLDGKVALVTGASRGIGLAIANCYAVNGAKVMLASRKADGLAAAAATLDGQVAVKVANTRNLDDIEGLVDAVLDRWGRLDVLVNNAATCPYVGPVVDVEPGQRDRNFEVNLRGSLFLIQKAWRAWMVSNGGSIINISSVGAFRTHRTMGFYDIGKAALAQLTRHLAGELGPRVRINAIVPGLVETRMAESIIKTRGVDIAARIPLQRNGIPEDIAGAALFPAGAAWSWMTLPRRHAKGAIEPDRLAVDVGVLEQFARHAAELLRSSQPSRVRDVGCEGRTHFVTESSEHRGIEDPWRDCVHPDALRRQIARQDECHSAHTGFRGRIGRLTNLTLERSGRRRHHHGAPVAIGRRWIARDSFSREPDDVVAADEIHSDDFFETVQRVCVSVTVEYTQPIAAPSRAMYDRPQRSCINRGVERLTQTLGISDIRRDKGCADVLGGSRAVRRREVGDNHAGATPRQRRRTRESQPRRSSGHQGSRVIQIHLGLLYTSRQWIHINVKHLLARVDGTKSDRIVAGRRGVVLADSNLWLNPTAGGRRGWIPRHELTGEAPEHAHA
putative acyl-CoA dehydrogenase(EC:1.3.99.-)	MGGWLSANAESLDEFRHHRAEAAEERWARDLRFIRVLYAAGWNRHGWPAEYGGVGGPAVLRNVLYEELENAGYRVPDCVIQLELQGEAFMKFAPEIAAARIPAALRGDEMWTQGFSEPEAGSDLASLRCKAVQDGDEWVITGQKTWSSYALSARWMGVLTRTGTPESRHRGITMFLVDLESPGVEVRPLRLANGVDECADVFFDNVRVSADNRIGEVGGGWDVAMYLLQFERANYGWLRQAHLARRLRELVSNIDAPDRHTANVLGNCWLANLALRLRCGDTVRRLFAGQRIGPEASIDKVLIAQAEQPLNDAFRDLFPAQFLFDPGEESEVWRSEWYYSRAASIYGGAGEIQRGIIADRLLGLPKEV
hypothetical protein	MVQTKSTELFEEDGRNSPVFVKHIFKANVTGHRSRVAAVSS

acyl-CoA dehydrogenase domain protein	MRGVDGDSVSSILLLAGEEAVLVPTSGLQCSPAKGFDIVDGWTRWRGSVSADDSDVLSTDGARHVRAVVRAALASELNGIASKVNELAVDHVTSRHQFGQALGSFQSVRHRLAETHVAIHAAMPVIELAWQATACGDAGSAELATAAKALAGRAFEIAAKNAHQVCGGMGLSWEHPLHRWVRRGTVLNSLIGSPDELAADLGGMLARGAALPVPNALCDDENQSTPA
Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	MRRRESVHPCMTRQPGVEVDPRHDGVTPLGENVDSPSHHYVTSLRHFAVKRPQDLAIVTPARSVTFAELSERANSLAGYLHSRGVAAGTPVALMSANSAELIEAFYAALVLGAPPANVNPRYHAHEVRYVLENCAAGAVLFDPTSATVVAEAVSAMEEPPFPLEFGSSRWRAAISSTTDYVRPQSVDDRLLIYTGGTTGLPRAVEWRVVDHFRMIWQMVKPTEPPPTPEALIDSGRRAPTASPCSPILHGVGLSLTLNTLNGGGTVVVSDRLTFDAADTFDLVRRYDVAVLGIVGDAFARPMLDELETGRWNGKLPSLRAISSAGASWTSSIRDRMANLLPGVKMVSNFGSTEALVARDISSDHSIDPGDGLVVIGENDRPAQPGEVGVVATAGYLPLGYLGDPQKTAETFRTINGRRYAVTGDEARVEADGRITLLGRGSAVINTGGEKVWPDEVEAVLRSVPGVLDAVVVGRPDERWGQRVAAVLRIAGTAEISDQRLSQGCRERLAPYKCPRQYVRVDEIPRTPVGKPDYRAIDSLLEES
hypothetical protein	MPKRYGTKEKEQESAAREERSTMMGRVAVTELVNRGGSETYRGRHARYRDLDTELVGGY

Butyryl-CoA dehydrogenase (EC 1.3.99.2)	MRRTLFDQEHEEFRVLVRDFIAKRVVPHFDEWVDEGIVPRQLYRELGAIGAMGFAIPEQYGGGGVDDYRYNAVLQEEAARANVTLGTMRTHLDVVLPYLLKYANDEQRERWFPGLVSGDLFLAIAMTEPGTGSDLAGVKTTAKRDGSDFIVNGAKTFITGAAHADLIIVLARTGEPIDGNRRSGLSLLVVEDGVAGFTKGRRLPKLGLRVQDTVELSFEDVRVPAINLLGEVDQAFSYLGHNLAQERLAIAVGSVAQSRAAMEMTIDYVKNRNLFGTTVSAMQNTKFEIAAVATEIEAAQAMCDNGVRDLVDGELSGADAAKIKLFCTEMQARTVDRCLQLHGGYGYILEYPISRLYADARVSRIYGGTSEVLKSIISKSLGL
Enoyl-CoA hydratase (EC 4.2.1.17)	MSDVLIVRDEAGVRQITMNRPERRNAMDIDTRRRFADLIVESHDNPDVRSIVITGAGGHFCAGADVTRMARSVDEEAAYERVQTAQEIAQVMAEGAKPQIAAVTGSAVGLGMGIALACDYIVAGPDARFAGGFVRVGLCADNGVLFTLPQRVGPARARTMMLLSETVPADEALRIGLIDKLTASDDQVLPAAIEVAATLAAGPPLALATISRAFTALPMTFMDALVREAELQAPLLCSDDHQEAATAFREKRRPVFTGR
hypothetical protein	MARWSDEFVFGQVWTDSDIPWEQQILVAITALATLGHHTQLRNYLHGALQGGVSEPALRHALSMLTVYAGFPVAIQALNVLSEVVAREARTTTTHPENAPGSASGR
Acyl-CoA thioesterase II (EC 3.1.2.-)	MSDRLQERLQLDDVGNDRFSSSAPGAPVNIYGGELAAHALVAANMTVTDDRLPHSLHCTYLAAGDPAHGLEHQVTRIRDGRSFSVRRVDVFNGGRLSVSATVSYQTKSAGVEHRRTAPRVPCPDDLPTFHGASNAAWIPWAEENPELEMRVAPHDPGDSLGRRTFWLKIRHDPAIGSADDLLEAAYAAYASDFTMIASIRLPHEEPDVKTHVMTTLTHSLYFHKPFAASQWHLVDHWSPAAAGGRGLSIAHAYNASGDLTMTAVQESLVRSASKKAG

probable oxidoreductase/Short-chain dehydrogenase	MKIIITGASSGIGKATATALAARGHQVVIACRTLKKGEEAAAQMSGDVRVQHLDLADLASVRRFADAVDTVDVLVNNAGVLGLPYTRTVDGFEAHMGTNHLGHFALTCLLADRTTDRVVSLSSSAYKYATPHLDDLNWERRKYSKSQAYGESKLATMLFINELARRGVRAYAADPGIVFSDITRDGGAILRWSGKYMHPIIGQTTGNGARSTLLAIETTAPSGTCFAPGGLMHQWGKPKVVIPHRKARNPRTALRLWELSAELTGCDAPYEPQASE
putative cytochrome P450 hydroxylase	MTETAVGKPDVRYDPYDFEIDNDPYPTFRRLRDEAPLYYNDEYDFYALSRFHDVKQASVEWETYPSGRGSVLELIKSGAAIPPGFILWEDPPRHDVHRGLLARVFTMRRIAQIEDKVRTVCTATLDKFVDTGRFDFVADLGSEIPMRTIGMMLGIPDSDQVSHRDRTNEGIKLADGQAPVDGSALLTLTDTQRIGDYINFRRRHPADDLMTDLLTATCTDSGGHERQLQDSTIQTYVGLLAAAGNETTVRLFSWAGKVLAEHPDQRRELANDPSLIPAAIEELLRFEAPSPVQARYVARDVEHHGRTMPQGSVVLLLTASANRDERAFEDPDRFDIHRKTKNHVSFGFGIHHCLGAHLARLEARVGLEEVFKRFPEWDVDWDNAVQSHTSTVRGWEKLPVVASPSLRDR
hypothetical protein	MAFVTGAARGQGRAHRFPTSSPEEVANLAAGEARYITGQQIRIDGGALIEFPNGPTRI
hypothetical protein	MKLPVEPELESLRAAIAARAKSTVRPLAAQVDREQKFSWELWAAVRELGLSRIHFPEDHGGDGGTFRAYTIASTEPAEYCAVASLYPGTSIQVAMALLQHGNPEHSRRFFAFWAAKRCGLGLYRAVDRCGFSAQRVAFIPGYMSARQAAAVDPEAPTRPTSRWLRSSSSLSWNFPAQTRRKIPARQLRPPMLWPSRR