DOI: 10.1002/mbo3.1180

### ORIGINAL ARTICLE

Open Access WILEY

# The genome sequence of hairy root *Rhizobium rhizogenes* strain LBA9402: Bioinformatics analysis suggests the presence of a new opine system in the agropine Ri plasmid

Marjolein J. G. Hooykaas | Paul J. J. Hooykaas 💿

Revised: 24 February 2021

Institute of Biology, Leiden University, Leiden, The Netherlands

#### Correspondence

Paul J. J. Hooykaas, Institute of Biology, Leiden University, Sylviusweg 72, Leiden, The Netherlands. Email: p.j.j.hooykaas@biology.leidenuniv.nl

#### Abstract

We report here the complete genome sequence of the Rhizobium rhizogenes (formerly Agrobacterium rhizogenes) strain LBA9402 (NCPPB1855rifR), a pathogenic strain causing hairy root disease. To assemble a complete genome, we obtained short reads from Illumina sequencing and long reads from Oxford Nanopore Technology sequencing. The genome consists of a 3,958,212 bp chromosome, a 2,005,144 bp chromid (secondary chromosome) and a 252,168 bp Ri plasmid (pRi1855), respectively. The primary chromosome was very similar to that of the avirulent biocontrol strain K84, but the chromid showed a 724 kbp deletion accompanied by a large 1.8 Mbp inversion revealing the dynamic nature of these secondary chromosomes. The sequence of the agropine Ri plasmid was compared to other types of Ri and Ti plasmids. Thus, we identified the genes responsible for agropine catabolism, but also a unique segment adjacent to the TL region that has the signature of a new opine catabolic gene cluster including the three genes that encode the three subunits of an opine dehydrogenase. Our sequence analysis also revealed a novel gene at the very right end of the TL-DNA, which is unique for the agropine Ri plasmid. The protein encoded by this gene was most related to the succinamopine synthases of chrysopine and agropine Ti plasmids and thus may be involved in the synthesis of the unknown opine that can be degraded by the adjacent catabolic cluster. The available sequence will facilitate the use of R. rhizogenes and especially LBA9402 in both the laboratory and for biotechnological purposes.

#### KEYWORDS

Agrobacterium rhizogenes, chromid, opine, ornithine cyclodeaminase, Ri plasmid, succinamopine synthase

# 1 | INTRODUCTION

Hairy root, a neoplastic plant disease with a wide host range, is characterized by the formation of adventitious roots from infected wound sites (Riker, 1930). It was originally encountered as a problem in tree nurseries, but nowadays the disease is also increasingly causing problems in the greenhouse by inducing extensive root mats, thereby reducing the harvest of cucumbers and tomatoes (Weller et al., 2000). The causal agent, a bacterium that was for long called *Agrobacterium rhizogenes*, contains a large, about 200 kbp

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made. © 2021 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

root-inducing (Ri) plasmid, which contains the essential virulence determinants (White & Nester, 1980). The molecular mechanism by which hairy root is induced is similar to that which is used by the related bacterium Agrobacterium tumefaciens to induce crown gall tumors in plants. During infection, part of the Ri plasmid (the T region) is transferred to plant cells and integrated into the plant genome (Bevan & Chilton, 1982). Expression of rol-genes located in the transfer DNA (T-DNA) leads to the transformation of normal cells into tumor cells that develop into roots that can grow in in vitro culture in the absence of added plant growth regulators (Jouanin et al. 1987). In the cells of hairy roots, unusual compounds called opines are produced, which are specific condensates of amino acids and keto acids or sugars (Petit et al., 1983). These opines, which are formed by enzymes encoded by the T-DNA, support the growth of the pathogen, which contains the catabolic genes usually in a region adjacent to the T region on the Ri plasmid (Dessaux et al., 1993). Based on the specific opines formed and degraded, agropine, cucumopine, mannopine, and mikimopine Ri plasmids are nowadays distinguished.

Like A. *tumefaciens*, A. *rhizogenes* has been disarmed by deleting the T-DNA genes to convert this region into a vector useful for plant genome engineering (Collier et al., 2018). Also, the bacterium as such is used for biotechnological research and application; in research for instance for gene function and gene expression analysis in roots (Ron et al., 2014); and in the industry to obtain roots that can be grown in bioreactors for the production of secondary metabolites (Mehrotra et al., 2015).

It has become apparent over the years that in nature various bacteria of the Rhizobiaceae family may cause hairy root or crown gall, depending on whether they carry a Ri plasmid or a Ti plasmid. Based on their physiological properties, three different groups (biovars 1-3) were distinguished already long ago (Kerr & Panagopoulos, 1977). The species name Rhizobium rhizogenes is now commonly used for the bacteria belonging to biovar 2. R. rhizogenes strains have two megabase DNA circles, a chromosome and a plasmid-derived megacircle, sometimes called a chromid (Jumas-Bilak et al., 1998; Harrison et al. 2010). Draft genome sequences of several strains are available on NCBI, and one draft genome sequence has been published for R. rhizogenes strain ATCC15834 consisting of 43 scaffolds (Kajala et al., 2014). However, up to now, only one complete genome sequence is available for R. rhizogenes, that of the avirulent agrocinproducing biocontrol agent Kerr 84 (Slater et al., 2009). Here, we present a second complete genomic sequence of R. rhizogenes, that of the hairy root-inducing strain LBA9402. This strain is a rifampicinresistant derivative of wild-type strain NCPPB1855, which is one of the most widely used laboratory strains (Desmet et al., 2020). By comparing the chromosome and chromid of LBA9402 with those of strain K84, we found that the chromosome was very similar, but that the chromid showed large differences due to a large 724 kb deletion accompanied by a large inversion, underscoring the dynamic nature of the chromid. Analysis of the sequence of the agropine Ri plasmid of LBA9402 revealed that this had a few unique areas including one that we predict encodes a new opine catabolic cluster, including the three genes characteristic for defining the three subunits of an opine

dehydrogenase. A candidate gene for a novel opine synthase was identified at the very right end of the TL-DNA.

# 2 | MATERIALS AND METHODS

#### 2.1 | Organism

*Rhizobium rhizogenes* strain LBA9402 is a rifampicin-resistant derivative of wild-type strain NCPPB1855, which was originally isolated from *Rosa* spp. (Hooykaas, 1979). The bacterium was grown on TY medium (Difco tryptone 5 g/l, Difco yeast extract 3 g/l, CaCl<sub>2</sub>.6H<sub>2</sub>O 1.3 g/l). The bacterium was tested for virulence by puncturing the stems of *Kalanchoe daigremontiana* and *Kalanchoe tubiflora* with a sterile wooden toothpick that had been dipped into a colony of the bacterium.

#### 2.2 | Sequencing methods

Rhizobium rhizogenes strain LBA9402 was cultured in TY medium (Beringer, 1974), followed by genomic DNA isolation using QIAGEN gravity-flow columns (QIAGEN Genomic-tip 100/G kit Cat No./ ID: 10243). The genome of LBA9402 was sequenced using a combination of Illumina and Oxford Nanopore Technologies platforms. Nanopore sequencing was done in-house, but Illumina sequencing was performed at the Leiden Genome Technology Center (LGTC) of the Leiden University Medical Center (Leiden, The Netherlands), where TruSeq DNA Libraries were sequenced on an Illumina HiSeq 2000 machine. The Oxford Nanopore sequencing library was generated with 200 ng DNA using the SQK-RBK004 Rapid Barcoding Kit. The library was pooled with another library, followed by in-house sequencing on a MinION flow cell (version R9.4.1).

### 2.3 | Data processing methods

After base calling with Albacore (version 2.3.4), the MinION reads were demultiplexed (with Epi2me). The total yield for LBA9402 was 298,712 reads, totaling 1,027,720,149 bp, with a mean read length of 3,441 bp. Nanopore reads were end-trimmed and filtered on average quality (>Q10) and length (>5,000 bp) with NanoFilt (64-fold coverage after filtering). A total of 4,518,191 99-nucleotide paired-end Illumina reads were quality and adapter trimmed using Cutadapt (70fold coverage). Hybrid assembly was performed using Unicycler version 0.4.7. Besides three contigs representing the two chromosomes and the Ri plasmid, the fourth contig of 5,386 bp was identified. This represented the bacteriophage  $\Phi$ X174 genome sequence, which is spiked-in at low concentration during Illumina library preparation. This contig was therefore removed from the assembly. The assembly was annotated using NCBI Prokaryotic Genome Annotation Pipeline (PGAP). In addition, PHASTER was used to annotate prophage sequences (Arndt et al., 2016). For the functional characterization of



**FIGURE 1** Large rearrangements in the chromid of *R. rhizogenes*. The chromid/secondary chromosome of LBA9402 was aligned to that of K84 using progressiveMauve. The aligned blocks were visualized with the R package genoplotR. Red areas are regions with the same orientation, whereas blue areas align in inverse orientation



**FIGURE 2** Circular representation of pRi1855 and comparative analysis to other Ri and Ti plasmids. The inner rings show BLASTn comparisons between Ri1855 and various Ri and Ti plasmids. Color intensity indicates the degree of sequence similarity, as shown in the legend. The number of regions and genes is indicated on the outermost ring

the encoded proteins, eggNOG-Mapper was employed (Huerta-Cepas et al., 2017). Insertion elements (IS elements) were identified using ISEScan (Xie & Tang, 2017). In Figure 2 and Appendix Figure A5, only complete insertion sequences, that is, including inverted repeats, are shown. IslandViewer was used to predict genomic islands (Dhillon et al., 2015), and CGView was used to generate a circular

map of pRi1855 (Stothard & Wishart, 2005). Mauve (progressive-Mauve) (Darling et al., 2010) and MUMmer (NUCmer) (Kurtz et al., 2004) were used to align the LBA9402 and K84 genomes. Average nucleotide identity (ANI) values were calculated with fastANI (Jain et al., 2018), and Digital DNA-DNA hybridization (DDH) values were estimated with GGDC 2.1 (identities/HSP length) (Meier-Kolthoff et al., 2013). BRIG was used to compare pRi1855 with other Ri and Ti plasmids (with BLASTN, e-value cut-off 1e-10) and to visualize the hits in concentric rings (Alikhan et al., 2011). Percent identity values between specific regions of pRi1855 and pRiA4 were obtained with EMBOSS needle (Needleman-Wunsch pairwise global alignment). Single-nucleotide polymorphisms (SNPs) and small indels were detected with Snippy (https://github.com/tseemann/snippy, version 4.6.0, ran versus pRiA4 with the contigs option --ctgs). For the comparisons between erythritol catabolism regions and between pRi1855 and Rhizobium lusitanum strain 629, BLASTn was run locally with BLAST version 2.9.0+, and the similarities were visualized with the R package genoplotR (Guy et al., 2010). Protein alignments were performed with MAFFT version 7.471, L-INS-I method (Katoh, 2013), and visualized with Jalview version 2.11.1.2 (Waterhouse et al., 2009) and Adobe Illustrator. Percentage identities were calculated with the R package seginr.

### 3 | RESULTS

# 3.1 | The genomic sequence of *Rhizobium rhizogenes* strain LBA9402

As Illumina sequencing data alone were not sufficient to obtain a high-quality and complete genome sequence of strain LBA9402, we additionally obtained long reads by Nanopore sequencing. Unicycler was used to obtain a hybrid assembly. This resulted in three circular contigs of 3,958,212 bp, 2,005,144 bp and 252,168 bp, respectively. More than 99.9% of both MinION and Illumina reads align to the assembly, indicating that the assembly is complete. The G + C content of the genome was 60%. The assembly was annotated using NCBI Prokaryotic Genome Annotation Pipeline (PGAP). In total, 5,822 coding sequences, 9 rRNA genes (3 operons), and 53 tRNA genes were annotated. COG categories were assigned to predicted coding sequences with eggNOG-mapper (Appendix Figure A1).

# 3.2 | Comparison to the sequence of *R. rhizogenes* strain K84

The LBA9402 genome sequence was uploaded to the Microbial Genome Atlas webserver for taxonomic classification by TypeMat, to verify that it is indeed a *R. rhizogenes* strain. Average nucleotide identity (ANI) to the type strain NBRC 13257 was very high (98.9%). The digital DNA-DNA hybridization (DDH) value as calculated with Genome-to-Genome Distance Calculator 2.1 (GGDC 2.1) was 91.6% versus the type strain. These high values (>95% ANI, >70% DDH)

confirm that LBA9402 is a R. rhizogenes strain. The only other complete, high-quality genome of R. rhizogenes obtained thus far is that of the avirulent biocontrol strain K84 (Slater et al., 2009). This latter strain does not contain an Ri plasmid, but the two LBA9402 megacircles have high sequence similarity to those of K84 (both genomes have 99% ANI, 92.1% DDH). As can be seen in Appendix Figure A2, the largest contig of LBA9402 is largely collinear with the primary chromosome of K84 (96% of the sequence can be aligned to that of K84 at >95% sequence identity, 99% ANI). It contains the genes for replication such as for a DnaA replication protein, repair, and DNA recombination, for cell division, for transcription and translation including the 3 rRNA clusters of the bacterium and the 53 tRNA genes. We annotated an extra tRNA in one of the unique regions of the LBA9402 primary chromosome. The chromosome contains a putative genomic island of about 100 kbp with a set of genes for conjugative DNA transfer encoding not only a Type IV secretion system for mating pair formation (from position 3,324,071 to 3,314,501), but also the enzymes necessary for DNA transfer and replication (from position 3,302,269 to 3,307,564). The genomic island contains a gene for a putative integrase and is surrounded by a direct repeat of 15 bp, which may be (the remains of) two att sites. Larger differences between the chromosomes of LBA9402 and K84 are mainly due to the presence/absence of other mobile elements. For example, various proteins encoded in the unique segment of DNA from position 745,195 to 784,185 in LBA9402 have homology to phage proteins (Rhizobium phage vB\_RleM\_PPF1 and other tailed phages) as revealed by the phage search tool PHASTER (Arndt et al., 2016) and thus seems due to the insertion of a prophage (Appendix Figure A3).

The second-largest replicons are less similar (Appendix Figures A2, 1), but still, 85% of the LBA9402 sequence aligns to that of K84 (and 64% of K84 aligns to LBA9402, 99% ANI). The LBA9402 sequence is smaller, mainly due to a large approximately 724 kb deletion, which seems accompanied by a large 1.8 Mbp inversion (Figure 1). The large deletion did not affect any class of genes in particular as can be seen in Appendix Figure A1b, which shows a similar distribution of the predicted proteins in COG categories in both replicons.

This secondary megacircle has a plasmid-like RepABC replication system but has a similar GC content as the primary chromosome. Such secondary megacircles are considered (developing) secondary chromosomes that over evolutionary time exchange genes with the primary chromosome and have been coined "chromids" (Slater et al., 2009; Harrison et al. 2010). The chromid of LBA9402 contains many metabolic genes, but also genes for the production of cell wall polysaccharides and fimbriae/pili. We found in the chromid a set of genes homologous to the erythritol region in Sinorhizobium meliloti and Rhizobium leguminosarum (Appendix Figure A4) including a transport operon with genes eryEFG, a catabolic operon with genes eryABCD and a deoR-type regulator (also called eryR) followed by genes called eryH and eryI (Barbier et al., 2014; Geddes & Oresnik, 2012; Yost et al., 2006). The ability to catabolize erythritol is one of the key characteristics distinguishing biotype 1 and biotype 2 agrobacteria (Kerr & Panagopoulos, 1977). The presence of erythritol catabolic

genes was thus expected, but it was remarkable that they were present on the more dynamic chromid instead of the chromosome. Finally, the 252 kbp circle represents the agropine pRi1855 plasmid, which is very different from the large nopaline catabolic plasmid carried by strain K84 and will be described below.

# 3.3 | General properties of the Ri plasmids

The pRi1855 plasmid comprises 252,168 bp. It has an approximately 4% lower GC content than the rest of the genome. In total, 236 protein-coding sequences were identified with an average size of 898 bp (Appendix Figure A5). Recently, a draft of agropine Ri plasmid pRiA4 was published (Thompson et al., 2020). This plasmid is slightly smaller than pRi1855 with a size of 249,350 bp, but has a similar restriction profile (Jouanin, 1984) and is indeed very similar to pRi1855 (99% ANI). We compared the agropine Ri plasmid pRi1855 sequence to publicly available Ri and Ti plasmid sequences: octopine Ti plasmid pTiAch5 (CP007228; Henkel et al., 2014; Huang et al., 2015), nopaline Ti plasmid pTiC58 (AE007871; Goodner et al., 2001; Wood et al., 2001), succinamopine Ti plasmid pTiEU6 (KX388535; Shao et al., 2019), agropine Ti plasmid pTiBo542 (DQ058764; Oger et al., 2001), chrysopine Tiplasmid pTiChry5 (KX388536; Shao et al., 2018), mannopine Ri plasmid pRi8196 (Weisberg et al., 2020), cucumopine Ri plasmid pRi2659 (NZ CP019703.3; Valdes Franco et al., 2016; Tong et al., 2018), and mikimopine Ri plasmid pRi pRi1724 (NC\_002575; Moriguchi et al., 2001). The conservation of the different areas in these plasmids is visualized in Figure 2, and we shall discuss these in the following parts.

As can be seen in Figure 2, the replication (*repABC*) and conjugative transfer (*tra*, *trb*) genes are very similar to those of other Ti and Ri plasmids. This was previously already shown for the replication and conjugative transfer genes of the closely related agropine Ri plasmid pRiA4 (Nishiguchi et al., 1987; Wetzel et al., 2015). Indeed most of these genes are very similar in both agropine Ri plasmids. For example, pRi1855 *repABC* and *tral* are 100% identical to those of pRiA4. The *traAFBH* and *traCDG* genes are 95.3% identical, and the *trb* operon 88.8%. In contrast, *traR* and *trbK* are less conserved (62.7%, 47.5%), also compared to the other Ri plasmids.

The agropine Ri plasmid has two T regions, one of which, the TL region, contains the rol-genes that are necessary and sufficient for the formation of hairy roots (Offringa et al., 1986; White et al., 1985). Other Ri plasmids have only one T region with very similar genes (Figure 3; Otten, 2018). In the agropine Ri plasmid, however, a copy of IS630 is inserted between orf3 and orf8. At the very left end in the agropine Ri TL region and the mannopine Ri T region, a gene for agrocinopine synthase is present, but only remnants of these genes are still present in the cucumopine and mikimopine Ri plasmids. At the very right end of the T region one (in the cucumopine and mikimopine Ri plasmids) or two (in the mannopine Ri plasmids), non-conserved genes are present. These encode the cucumopine (Valdes Franco et al., 2016) and mikimopine synthases (Moriguchi et al., 2001), respectively, and the two genes necessary for mannopine synthesis in the mannopine Ri plasmid (Figure 3). The TL region of the agropine Ri plasmid pRiA4 was previously sequenced (Slightom et al., 1986), and this revealed at the right end the presence of orf15/rolD and three smaller orfs. In our pRi1855 sequence, we find besides orf15/rolD only one larger orf, hereinafter called orf16. The function of orf16 is unknown but may encode an unknown opine synthase as will be discussed in the next paragraph. The agropine Ri plasmid has besides the conserved T region (TL region) an additional T region (TR region) containing aux-genes involved in the biosynthesis of the auxin indole acetic acid (Offringa et al., 1986) and the genes mas1, mas2, and ags for agropine biosynthesis (Bouchez and Tourneur 1991).

The virulence region of pRi1855 responsible for the transfer of the T-DNA into plant cells contains the essential virulence genes virA, virB1-virB11, virG, virC1, virC2, and virD1-virD5 in the same order as in other Ri and Ti plasmids, but although virE3 is present close to virD5, the virE1 and virE2 genes are missing and replaced by a new orf with some similarity to nopaline pTi virF. The sequence of the cucumopine and mikimopine Ri plasmids in this area is almost identical



**FIGURE 3** Unique genes located at the very right end of the T regions of Ri plasmids. Schematic overview of genes in the T regions of different Ri plasmids (pRi1724: mikimopine type, pRi2659: cucumopine type, pRi8196: mannopine type, pRi1855: agropine type). Homologous genes have the same color

suggesting that the deletion of virE2 occurred once before the divergence of these Ri plasmids. The virE2 gene is functionally replaced by a gene called GALLS as reported before for pRiA4 (Hodges et al., 2004). The GALLS gene was previously also identified in the cucumopine and mikimopine Ri plasmids by Southern analysis but is not present in the mannopine Ri plasmid, which still carries the virE2 gene (Hodges et al., 2004). Together with a tzs gene, which encodes an enzyme that catalyzes the synthesis of the cytokinin zeatin riboside 5'-phosphate (Krall et al., 2002), GALLS is located outside the vir region, about 65 kbp clockwise from the virE3 gene, near opine catabolic genes in the cucumopine and mikimopine Ri plasmids. We could now identify and locate the GALLS gene in the pRi1855 sequence at a completely different location next to the *traG* gene almost 90 kbp counterclockwise from the virE3 gene. Besides the core set of vir genes mentioned above, the vir region of the agropine Ri plasmid contains next to the virA gene the tzs gene, virK, a second nopaline Ti-like virF gene, and finally virH1, virH2. It resembles in this respect the vir region of the nopaline Ti plasmid and like in the nopaline Ti plasmids virJ is absent. The virD3 gene, which is very variable, also shows the highest similarity to that of nopaline Tiplasmids. Other types of Ri plasmids have similar vir regions, but a virH2 gene is absent from the cucumopine and mikimopine Ri plasmids, and as mentioned above a tzs gene is present, but located in an entirely different area of the plasmid.

### 3.4 | Agropine and agrocinopine catabolism genes

Hairy roots formed by agropine strains contain agropine, agropinic acid, mannopinic acid, and mannopine (Petit et al., 1983). The agropine Ri plasmid enables host strains to degrade agropine. R. rhizogenes strains such as A4, but not NCPPB1855 contain a second, catabolic plasmid with genes for catabolism of the other three mannityl opines (Petit et al., 1983). We have now identified the genes for agropine transport and catabolism in pRi1855, which are located in a segment of the plasmid adjacent to the TR region (Figure 2). This region contains genes with high similarity to the genes described by Kim and Farrand (1996) on the octopine Ti plasmid involved in agropine uptake and degradation. These genes encode an agropine permease and also comprise agcA for the delactonase converting agropine into mannopine, mocC for oxidizing mannopine into deoxyfructosyl glutamine, and mocDE determining the deconjugase liberating an amino acid and a phosphorylated sugar. In the octopine Ti plasmid, the genes mocA and mocB encode enzymes with weak homology to glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydratase (Kim & Farrand, 1996). These are probably involved in further catabolism of the released phosphorylated sugar. However, while an intact homolog of mocA was present in pRi1855, as well as a homolog of mocC, in between only a truncated remnant of a gene homologous to mocB was present due to a deletion of more than 1 kbp. Regulators closely related to mocS and mocR are present in an identical position in front of mocA and between mocC and mocD. Our detection of a mocD gene in pRi1855 was remarkable as such gene was thought to be

absent from the agropine Ri plasmid (Baek et al., 2005). The agropine Ri plasmid has adjacent to the left end of the TL region in pRi1855 a set of *acc* genes for agrocinopine catabolism (Figure 2), which matches the presence of an *acs* gene for the biosynthesis of agrocinopine in the TL region. These genes are also present in the mannopine Ri plasmid but absent from the cucumopine and mikimopine Ri plasmids.

## 3.5 | Genes for a novel opine system in pRi1855

The pRi1855 plasmid has several regions with genes of unknown function. It shares a large region of about 65 kbp (the area 88-153 kbp on the map of Figure 2) with the other Ri plasmids. This region contains genes putatively involved in sugar transport, glycerol metabolism and encodes several transcription regulators and two chemoreceptors (Moriguchi et al., 2001). Besides, the pRi1855 plasmid has several unique areas with genes that are found in none of the other types of Ri and Ti plasmids described. These include transposable elements (Figure 2) and two larger areas of about 20 kbp (the area 20-40 kbp on the map adjacent to the right border of the TL region in Figure 2) and about 24 kbp (the area 64-88 kbp on the map adjacent to the agropine catabolic genes). The latter area contains mainly metabolic genes and may have been introduced into pRi1855 by transposition as it is surrounded by IS5-like insertion sequences. It may have originated from the chromosome of another Rhizobium species, as a very similar stretch of DNA was detected in the recently sequenced chromosome of Rhizobium lusitanum strain 629 (Appendix Figure A6). The 20 kbp segment adjacent to the right border of the TL region may be involved in the transport and catabolism of a new opine. In this area, we identified all the three characteristic genes that together code for the three subunits of a putative flavin-containing opine dehydrogenase (Appendix Figures A7-A9). Flavin-containing opine dehydrogenases such as octopine, nopaline, and succinamopine dehydrogenase consist of three subunits OdhABC that are encoded by noxABC-like/ooxABC-like genes arranged in tandem in the genome (Watanabe et al., 2015). The three genes in pRi1855 (F3X89 28345, F3X89 28350, F3X89 28355) encode closely related proteins in which the characteristic binding sites for the FAD and FMN co-factors and the Fe-S cluster have fully been conserved (Appendix Figures A7-A9). These three genes are surrounded on both sides by genes for a transport system and a LysR-type regulator. In Ti and Ri plasmids, genes encoding an opine dehydrogenase are often accompanied by genes encoding the permease required for uptake of a specific opine into the bacterial cell. Also, genes encoding enzymes necessary for the further catabolism of the products liberated by the activity of the opine dehydrogenase on the opine substrates are often present in the vicinity. In this area of pRi1855 genes encoding such metabolic proteins are also present, including genes encoding a putative saccharopine dehydrogenase and a putative aminoadipate semialdehyde dehydrogenase, which may form part of a catabolic pathway of the amino acid lysine (de Mello Serrano et al., 2012). A gene for an AsnC/Lrp regulator is located at the end of this DNA segment. The Lrp family of

-WILEY

transcriptional regulators is known to control amino acid metabolism in bacteria (Brinkman et al., 2003).

If these genes are involved in the catabolism of an opine, a gene for an unknown opine synthase should be present in the T region of pRi1855. Genes for agrocinopine synthase are located at the extreme left end of the T region in Ti and Ri plasmids, while genes for nopaline synthase, octopine synthase, and succinamopine synthase are located at the extreme right end of the T region in Ti plasmids. In cucumopine, mikimopine, and mannopine Ri plasmids, the genes for cucumopine, mikimopine, and mannopine synthesis are likewise located immediately next to the right border repeat. We find at the very right end of the TL region of pRi1855 two related genes (Figure 3): orf15/rolD and orf16, which share 55% identity (Figure 4, Table 1). These genes are not present in the T regions of any of the other types of Ri plasmids (Figure 3). The orf15 has been called rolD; the encoded RoID protein has weak sequence homology with ornithine cyclodeaminases and indeed can convert ornithine into proline (Trovato et al., 2001). The role of rolD in hairy root formation is marginal, but the gene can influence plant development by its metabolic activity (Trovato et al., 2018). Using BLASTP with the proteins encoded by orf15 and orf16 as a query, we picked up the succinamopine synthases encoded by the T region of chrysopine pTiChry5 (Shao et al., 2018) and the agropine pTiBo542 (Oger et al., 2001) as the most related proteins. The proteins encoded by orf15 and orf16 share 44-47% identity with the two succinamopine synthases, which themselves share 93% identity (Table 1). All these proteins (encoded by orf15/rolD, orf16, susL) are evolutionarily related to ornithine cyclodeaminases (encoded by ocd genes). They share, for instance, about 19-21% identity with the ornithine cyclodeaminase encoded by the nopaline Ti plasmid. That ornithine cyclodeaminase (ocd) genes can evolve novel biochemical functions during evolution is known for some time. For instance, its function has been reported to evolve into an alanine dehydrogenase in Archaeoglobus fulgidus and into a tauropine dehydrogenase in Halichondria japonica (Sharma et al., 2013; Watanabe et al., 2014). It can evolve also in an opine (succinamopine) synthase. Therefore, it would seem possible that either or both of the ocd-like genes at the right end of pRi1855 (orf15, orf16) similarly have evolved a novel opine synthase function, producing an unknown opine that can be degraded by the putative opine dehydrogenase encoded in the area with genes of unknown function located adjacent to the right border of the TL region.

# 4 | DISCUSSION

The second complete, high-quality genomic sequence of a Rhizobium rhizogenes strain and the first of a virulent strain enabled us to make a comparison with the sequence of the previously sequenced biocontrol strain K84. This revealed high conservation of the primary chromosome but showed large differences in the secondary megacircle, the chromid. It has been described that chromids have a plasmidlike RepABC replication system but have a similar GC content as the primary chromosome and this is also the case in LBA9402. It has been proposed that chromids are plasmids that evolve into secondary chromosomes and overtime exchange genes with the primary chromosome (Slater et al., 2009; Harrison et al. 2010). We found that the chromid of strain LBA9402 was much smaller than that of strain K84 due to the absence of a segment of 724 kbp that may have been deleted in LBA9402 or inserted in K84. Also, we found that this insertion/deletion was accompanied by a large inversion of a segment of 1.8 Mbp. The presence of such complex rearrangements is in line with their plasmid descent and the genes which they carry being mostly non-essential.

Our genomic sequence includes the complete sequence of the agropine pRi1855 plasmid. Over the years, sequences have already been published dealing with specific parts of the closely related agropine Ri plasmid pRiA4 and recently a draft of the completed pRiA4 sequence was published (Thompson et al., 2020). However, it still differs in numerous areas, both by base substitutions, small insertions/deletions (803 differences were detected with variant caller Snippy), and a few larger insertions/deletions encoding complete genes. Also, likely due to technical sequencing differences (the pRiA4-carrying bacterium was solely sequenced with Illumina "short read" technology whereas for LBA9402 we additionally obtained "long reads" with Nanopore sequencing), the repeat-containing GALLS gene sequence is shorter (presumably collapsed) in pRiA4.

Hairy roots formed by agropine strains contain agropine, agropinic acid, mannopinic acid, and mannopine (Petit et al., 1983). The agropine Ri plasmid, however, enables host strains to degrade agropine, but not the other mannityl opines. *R. rhizogenes* strains such as A4, but not NCPPB1855, contain a second, catabolic plasmid with genes for catabolism of the other three mannityl opines (Petit et al., 1983). We have now identified the genes for agropine catabolism in pRi1855 adjacent to the right border of the TR region. The region

TABLE 1 Percentage identities between the two rightmost proteins of the LBA9402 TL-DNA, succinamopine synthase (susL) of pTiBo542 and pTiChry5, and two bacterial ornithine cyclodeaminases.

	LBA9402 orf15	LBA9402 orf16	Bo542 susL	Chry5 susL	C58 ocd	Pseudomonas putida ocd
LBA9402 orf15	100%	55%	46%	47%	19%	19%
LBA9402 orf16	55%	100%	44%	44%	21%	21%
Bo542 susL	46%	44%	100%	93%	19%	18%
Chry5 susL	47%	44%	93%	100%	19%	16%
C58 ocd	19%	21%	19%	19%	100%	56%
Pseudomonas putida ocd	19%	21%	18%	16%	56%	100%

LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL C58_ocd P.putida_ocd	1 1 1 1 1	M AKQL CEVHWTESAL SNL DIQITPAF VDEA M ADQIHEQCWTEESLSAF GIYITPEL VADA M TTEISERHWSEQKLQEL	30 30 30 30 54 29
LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL C58_ocd P.putida_ocd	31 31 31 31 55 30	L K S YWE V F A A G K A G H V K A YL T F N K N V P G W T E G A L I G Y F G R Y S G T L K D Y W H L F A E G R L G H N K V F V P L D R E G C T E G A L V G F V G D Y S G V L Q D Y W D S Q F E K R Q V K K Y L Q F D S I P N G W R E G V L I G C F K D Y S G V L Q D Y WE S Q F D K R Q V K K Y L Q F D S I P N G W R E G V L I G C F K D Y S G V L Q D Y WE - S Q F D K R Q V K K Y L Q F D S I P N G W R E G V L I G C F K D Y S G V L Q D Y WE - S Q F D K R Q V K K Y L Q F D S I P N G W R E G V L I G C F K D Y S G V L Q D F K R WE S Q F D K R Q V K K Y L Q F D S I P N G W R E G V L I G C F K D Y S G V I E E D F R R WE S Q F D K R Q V K K Y L Q F D S I P N G W R E G V L I G C F K D Y S G V I E E D F R R WE S F D K I P R I A S H S R D G V I E L M P T S D G T L Y - G F R D D F K R WQ A F D K S A R V A S H S E V G V I E L M P V A D K S R Y - A F	74 72 72 93 68
LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL C58_ocd P.putida_ocd	75 73 73 73 94 69	K D I H F A P T N A A T G K P L Q H S D I F L R D K V S G T L L M S V E G VA I S N G R T G WF S L A C V N 1 K S I H S N P S N A R Y K E P T S H I D V V L R N R F T G R R L F S V D G VA I S E R T G WL A L A C L N 1 K N I H F S P S N E R N N L P L R H I D I L L R R K S T G E R L L S L E G KA I S F G R T A WF A P A C I D 1 K N I H F S P S N E Q K N L P L R H I D I L L R R K S T G E R L L S L E G KA I S F G R T A WF A P A C I D 1 K N I H F S P S N E Q K N L P L R H I D I L L R R K S T G E R L L S L E G KA I S F G R T A WF A P A C V D 1 K Y V N G H P K N T K S G R Q T V T A F G V L S D V D S G Y P L L S E M T I L T A L R T A A T S A I A A K 1 K Y V N G H P A N T A R N L H T V M A F G V L A D V D S G Y P V L L S E L T I A T A L R T A A T S L MA A Q 1	28 26 26 26 47 22
LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL C58_ocd P.putida_ocd	129 127 127 127 148 123	LLLQGREDIDVFLFGAGKVAEAVILSLNYGAAKRIRKVAVLSRGKKSNFELVEK 1 LLLQGRSQINVFLFGAGKVAEAIILALNSGASGRIQRIAILSRCNQSNHELVQR 1 LLLNAKDKIEIFLFGAGELGREIVRALNENTSCRIKKIWILSRG-SSNERLVRE 1 LLLKAKDKIEVFLFGAGELGREIVRALNENISSRIKKIWILSRG-TSNERLVRE 1 YLARKDSR-TMALIGNGAQSEFQALAFKALIGVDRIRLYDIDPEATARCSRN 1 ALARPNAR-KMALIGNGAQSEFQALAFHKHLGIEEIVAYDTDPLATAKLIAN 1	82 80 79 79 98 73
LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL C58_ocd P.putida_ocd	183 181 180 180 199 174	L R D C V T F SL E A V G D R S L L P Y SQ F V I MAT N Y G K P V F E A A E I A P NA V T L S L G I D D M 2 L Q P E V K I SL KA V N N R A Y L SK SK F I I T AT N S N K P V F E M K E I T D N A V T L T L G I D E L 2 L A D T V K I P L A A V WE T S A L S G A D L V I T A T - S S P P A F N P G D L A P K A V T L S L S M N D V 2 L A D T V K I P L A A V WE T S A L S G A D L V I T A T - S S P P A F N P G D L A P K A V T L S L S M N D V 2 L A D T V K I P L A A V WE T S A L S G A D L V I I A T - S S P P A F N P G D L A P K A V T L S L S M D D V 2 L Q R F - G F Q I E A C T S A E Q A V E G A D I I T T A T A D K H N A T I L S D N M I G P G V H I N G V 2 L K E Y S G L T I R R A S S V A E A V K G V D I I T T V T A D K A Y A T I I T P D M L E P G M H L N A V 2	36 34 32 32 49 25
LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL C58_ocd P.putida_ocd	237 235 233 233 250 226	PPD YIEHVLSADGLIVADDLVAMEARNVDAVALYYSRRGMKLTQHGKR 2 PAD YFDHVLKLSGIVVADDMDAMETRNIDSLALHYSRRHLKLTKHGRD 2 PQR YLESLIEAKATIICDDMLAMEERNVNALALLFSRKGQKLSEIGPS 2 PQR YLESLIEAKAPIICDDMLAMEERNVNALALLFSRKGQKLSEVGPS 2 GGDCPGKTEMHRDILLRSDIFVEFPPQTRIEGEIQQLAPDHPVTELWRVMTGQD 3 GGDCPGKTELHADVLRNARVFVEYEPQTRIEGEIQQLPADFPVVDLWRVLRGET 2	84 82 80 80 03 79
LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL C58_ocd P.putida_ocd	285 283 281 281 304 280	DG I K NYTE I L D D E ALMND L KAWK GPANF SPVGLAS I D VAVAAHVYE T L L NK L AG 3 NG I R NYSEV L R D PVLMQRL I SWNGPANF SAVGLAS YD L AVAVHLYEK L AK T V VQ 3 VG VK HYSAI RAD K S L F D A I ST R GG P A HF T T VG L A SMD I AVAA H V Y E K M I A R P N - 3 VG VK HYSAI RAD K S L F D A I SAR GG P A HF T T VG L A SMD I AVAA H V Y E K M I A L P N - 3 VG R K SDKQ I T L F D S VG F A I E D F SAL R Y V R D R V E G SS H S 3 E G R Q S D S Q V T V F D S VG F A L E D Y T V L R Y V L Q Q A E K R G MG 3	38 36 33 33 41 17
LBA9402_orf15 LBA9402_orf16 Bo542_susL Chry5_susL	339 337	PMEHGH       3         KYVDHAASVLLA       3	44
C58_ocd P.putida_ocd	342 318	SPLDG 3 TK <mark>ID</mark> LV <mark>P</mark> WVEDD <mark>P</mark> KDLFSHTR <mark>G</mark> RA <mark>G</mark> KRRIRRVA 3	71 50

**FIGURE** 4 The proteins encoded by *orf15* and *orf16* are evolutionarily related to succinamopine synthase and more distantly to ornithine cyclodeaminase. Multiple sequence alignment of proteins encoded by *orf15* and *orf16* of pRi1855, *susL* from A. *tumefaciens* strains Chry5 and Bo542, and *ocd* from A. *tumefaciens* C58 and *Pseudomonasputida*. Colors are according to the Clustal X color scheme

\_MicrobiologyOpen

WILEY

embraces an *agcA* gene for the delactonase converting agropine into mannopine, *mocC* for oxidizing mannopine, and *mocD* and *mocE* together determining the enzymes that can release the amino acid and a phosphorylated sugar from the conjugate. This pathway would allow the bacterium to degrade both agropine and mannopine. However, it is known that the bacteria carrying pRi1855 cannot degrade mannopine. This may be because mannopine cannot induce the catabolic genes or because the bacterium cannot import mannopine into the cell. Indeed, the pRi1855 plasmid contains genes for an agropine permease, but not for a mannopine transport system.

Agrobacteria induce neoplasias in which opines are formed that serve as a nutritional source of the bacteria. All the different types of Ti and Ri plasmids described so far have a gene coding for an opine synthase at the very right end of the T region. Our sequence now shows that also the agropine Ri plasmid has one larger orf (orf16) at the very right end of the TL region, which shares 55% identity with the neighboring rolD gene. This gene encodes a protein that is evolutionary related to ornithine cyclodeaminase (ocd) and which still has ornithine cyclodeaminase activity (Trovato et al., 2001). Here, we discovered that both rolD and orf16 have a significant identity of 44%-47% with the susL genes encoding succinamopine synthase in the agropine and chrysopine Ti plasmids. Ornithine cyclodeaminase encoding (ocd) genes have also been reported to have evolved into genes encoding new enzymatic activities such as alanine dehydrogenase activity in Archaeoglobus fulgidus and tauropine dehydrogenase activity in Halichondria japonica (Sharma et al., 2013; Watanabe et al., 2014). Previously, it was found that enzymes involved in the biosynthesis of mannopine are related to and thus may have been evolved from the enzymes required for degradation (Kim & Farrand, 1996). In this case, the enzymatic activity of the enzyme remained the same. but acted in the other direction, synthesis instead of degradation. We now find that an ocd-like gene may have evolved into a gene encoding a succinamopine synthase, which catalyzes a very different enzymatic step. Similarly, we hypothesize that an ocd-like gene may have evolved in the agropine Ri plasmid into a gene for a new opine synthase. Opine catabolic genes are often located close to the synthase gene, but on the other side of the right border, an arrangement seen in many different Ti and Ri plasmids. Indeed, three of the genes residing here together have the signature of the trios of genes such as ooxABC and noxABC (Appendix Figures A7-A9) that are known to encode the three subunits of octopine and nopaline dehydrogenase (Watanabe et al., 2015) and thus may encode the opine dehydrogenase needed for catabolism of the novel unknown opine.

Given its frequent application, the available sequence will facilitate the use of *R. rhizogenes* and especially LBA9402 in both the laboratory and for biotechnological purposes.

#### ACKNOWLEDGEMENTS

We are grateful to Amke den Dulk-Ras and Shuai Shao for help with DNA isolation, and Dr. Marcel van Verk and Dr. Chris Henkel for contributions to the initial analyses of the Illumina data. P.J.J.H. was supported by the fellowship associated with the appointment as Academy Professor by the Royal Netherlands Academy of Sciences.

## CONFLICT OF INTEREST

None declared.

#### AUTHOR CONTRIBUTIONS

Marjolein J. G. Hooykaas: Conceptualization (equal); Data curation (equal); Formal analysis (lead); Funding acquisition (supporting); Investigation (lead); Methodology (lead); Project administration (equal); Resources (supporting); Software (lead); Supervision (supporting); Validation (lead); Visualization (lead); Writing-original draft (equal); Writing-review & editing (equal). Paul J. J Hooykaas: Conceptualization (equal); Data curation (equal); Formal analysis (supporting); Funding acquisition (lead); Investigation (supporting); Methodology (supporting); Project administration (equal); Resources (lead); Software (supporting); Supervision (lead); Validation (supporting); Visualization (supporting); Writing-original draft (equal); Writing-review & editing (equal).

#### ETHICS STATEMENT

None required.

#### DATA AVAILABILITY STATEMENT

The complete genome sequence of *R. rhizogenes* LBA9402 is available in GenBank under accession numbers CP044122, CP044123, and CP044124. The raw reads are deposited in the Sequence Read Archive under accession numbers SRR10177303 and SRR10177304. BioProject PRJNA566100: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA566100

#### ORCID

Paul J. J. Hooykaas 🕩 https://orcid.org/0000-0002-9736-6927

#### REFERENCES

- Alikhan, N. F., Petty, N. K., Zakour, N. L. B., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics, 12, 402.
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, 44(W1), W16-W21.
- Baek, C., Farrand, S. K., Park, D.-K., Lee, K.-E., Hwang, W., & Kim, K.-S. (2005). Genes for utilization of deoxyfructosyl glutamine (DFG), an amadori compound, are widely dispersed in the family Rhizobiaceae. *FEMS Microbiology Ecology*, 53, 221–233.
- Barbier, T., Collard, F., Zúñiga-Ripa, A., Moriyón, I., Godard, T., Becker, J., Wittmann, C., Schaftingen, E. V., & Letesson, J.-J. (2014). Erythritol feeds the pentose phosphate pathway via three new isomerases leading to D-erythrose-4-phosphate in Brucella. Proceedings of the National Academy of Sciences of the United States of America, 111, 17815–17820.
- Beringer, J. E. (1974). R factor transfer in Rhizobium leguminosarum. Journal of General Microbiology, 84, 188–198.
- Bevan, M. W., & Chilton, M. D. (1982). T-DNA of the Agrobacterium Ti and Ri plasmids. Annual Review of Genetics, 16, 357–384.

- Bouchez, D., & Tourneur, J. (1991). Organization of the agropine synthesis region of the T-DNA of the Ri plasmid from Agrobacterium rhizogenes. *Plasmid*, 25, 27–39.
- Brinkman, A. B., Ettema, T. J., de Vos, W. M., & van der Oost, J. (2003). The Lrp family of transcriptional regulators. *Molecular Microbiology*, 48, 287–294.
- Collier, R., Thomson, J., & Thilmony, R. (2018). A versatile and robust Agrobacterium-based gene stacking system generates high quality transgenic Arabidopsis plants. *The Plant Journal*, *95*, 573–583.
- Darling, A. E., Mau, B., & Perna, N. T. (2010). progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*, 5, e11147.
- de Mello Serrano, G. C., e Silva Figueira, T. R., Kiyota, E., Zanata, N., & Arruda, P. (2012). Lysine degradation through the saccharopine pathway in bacteria: LKR and SDH in bacteria and its relationship to the plant and animal enzymes. *FEBS Letters*, 586, 905–911.
- Desmet, S., Dhooghe, E., De Keyser, E., Van Huylenbroeck, J., Müller, R., Geelen, D., & Lütken, H. (2020). Rhizogenic agrobacteria as an innovative tool for plant breeding: current achievements and limitations. Applied Microbiology and Biotechnology, 104, 2435–2451.
- Dessaux, Y., Petit, A., & Tempe, J. (1993). Chemistry and biochemistry of opines, chemical mediators of parasitism. *Phytochemistry*, 34, 31–38.
- Dhillon, B. K., Laird, M. R., Shay, J. A., Winsor, G. L., Lo, R., Nizam, F., Pereira, S. K., Waglechner, N., McArthur, A. G., Langille, M. G. I., & Brinkman, F. S. L. (2015). IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Research*, 43, W104–W108.
- Franco, J. A. V., Collier, R., Wang, Y., Huo, N., Gu, Y., Thilmony, R., & Thomson, J. G. (2016). Draft genome sequence of Agrobacterium rhizogenes strain NCPPB2659. *Genome Announcements*, 4, e00746-16.
- Geddes, B. A., & Oresnik, I. J. (2012). Genetic characterization of a complex locus necessary for the transport and catabolism of erythritol, adonitol and l-arabitol in *S.meliloti*. *Microbiology*, 158, 2180–2191.
- Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Goldman, B. S., Cao, Y., Askenazi, M., Halling, C., Mullin, L., Houmiel, K., Gordon, J., Vaudin, M., Iartchouk, O., Epp, A., Liu, F., Wollam, C., Allinger, M., ... Flanagan, C. (2001). Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science*, 294, 2323–2328.
- Guy, L., Kultima, J. R., & Andersson, S. G. (2010). genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* 26: 2334–2335.
- Harrison, P. W., Lower, R. P. J., Kim, N. K. D., & Young, J. P. W. (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiology*, 18, 141–148.
- Henkel, C. V., den Dulk-Ras, A., Zhang, X., & Hooykaas, P. J. (2014). Genome sequence of the octopine-type Agrobacterium tumefaciens strain Ach5. Genome Announcements, 2, e00225-14.
- Hodges, L. D., Cuperus, J., & Ream, W. (2004). Agrobacterium rhizogenes GALLS protein substitutes for Agrobacterium tumefaciens singlestranded DNA-binding protein VirE2. *Journal of Bacteriology*, 186, 3065–3077.
- Hooykaas, P. J. J. (1979). The role of plasmid determined functions in the interactions of Rhizobiaceae with plant cells. A genetic approach. PhD Thesis. Leiden University, The Netherlands.
- Huang, Y. Y., Cho, S.-T., Lo, W.-S., Wang, Y.-C., Lai, E.-M., & Kuo, C.-H. (2015). Complete genome sequence of Agrobacterium tumefaciens Ach5. Genome Announcements, 3, e00570-15.
- Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., & Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution*, 34, 2115–2122.
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic

genomes reveals clear species boundaries. *Nature Communications*, *9*, 5114.

- Jouanin, L. (1984). Restriction map of an agropine-type Ri plasmid and its homologies with Ti plasmids. *Plasmid*, 12, 91–102.
- Jouanin, L., Vilaine, F., Tourneur, J., Tourneur, C., Pautot, V., Muller, J. F., & Caboche, M. (1987). Transfer of a 4.3 kb fragment of the TL-DNA of *Agrobacterium rhizogenes* strain A4 confers the pRi transformed phenotype to regenerated tobacco. *Plant Science*, *53*, 53–63.
- Jumas-Bilak, E., Michaux-Charachon, S., Bourg, G., Ramuz, M., & Allardet-Servent, A. (1998). Unconventional genomic organization in the alpha subgroup of the Proteobacteria. *Journal of Bacteriology*, 180, 2749–2755.
- Kajala, K., Coil, D. A., & Brady, S. M. (2014). Draft genome sequence of Rhizobium rhizogenes strain ATCC15834. Genome Announcements, 2, e01108-14.
- Katoh, S. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology* and Evolution, 30, 772–780.
- Kerr, A., & Panagopoulos, C. G. (1977). Biotypes of Agrobacterium radiobacter var. tumefaciens and their biological control. Journal of Phytopathology, 90, 172-179.
- Kim, K. S., & Farrand, S. K. (1996). Ti plasmid-encoded genes responsible for catabolism of the crown gall opine mannopine by Agrobacterium tumefaciens are homologs of the T-region genes responsible for synthesis of this opine by the plant tumor. *Journal of Bacteriology*, 178, 3275–3284.
- Krall, L., Raschke, M., Zenk, M. H., & Baron, C. (2002). The Tzs protein from Agrobacterium tumefaciens C58 produces zeatin riboside 5'-phosphate from 4-hydroxy-3-methyl-2-(E)-butenyl diphosphate and AMP. FEBS Letters, 527, 315–318.
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*, 5, R12.
- Mehrotra, S., Srivastav, V., Rahman, L. U., & Kukreja, A. K. (2015). Hairy root biotechnology-indicative timeline to understand missing links and future outlook. *Protoplasma*, 252, 1189–1201.
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., & Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, 14, 60.
- Moriguchi, K., Maeda, Y., Satou, M., Hardayani, N. S., Kataoka, M., Tanaka, N., & Yoshida, K. (2001). The complete nucleotide sequence of a plant root-inducing (Ri) plasmid indicates its chimeric structure and evolutionary relationship between tumor-inducing (Ti) and symbiotic (Sym) plasmids in Rhizobiaceae. *Journal of Molecular Biology*, 307, 771–784.
- Nishiguchi, R., Takanami, M., & Oka, A. (1987). Characterization and sequence determination of the replicator region in the hairy-rootinducing plasmid pRiA4b. *Molecular and General Genetics MGG*, 206, 1–8.
- Offringa, I. A., Melchers, L. S., Regensburg-Tuink, A. J. G., Costantino, P., Schilperoort, R. A., & Hooykaas, P. J. J. (1986). Complementation of Agrobacterium tumefaciens tumor-inducing aux mutants by genes from the TR-region of the Ri plasmid of Agrobacterium rhizogenes. Proceedings of the National Academy of Sciences of the United States of America, 83, 6935–6939.
- Oger, P. M., Reich, C., Olsen, G. J., & Farrand, S. K. (2001). Complete nucleotide sequence and analysis of pTiBo542: what genomics tells us about structure and evolution of plasmids in the family *Rhizobiaceae. Plasmid*, 45, 169–170.
- Otten, L. (2018). The Agrobacterium phenotypic plasticity (*Plast*) genes. Current Top Microbiology Immunology, 418, 375–420.
- Petit, A., David, C., Dahl, G. A., Ellis, J. G., Guyon, P., Casse-Delbart, F., & Tempé, J. (1983). Further extension of the opine concept: Plasmids in Agrobacterium rhizogenes cooperate for opine degradation. Molecular and General Genetics MGG, 190, 204–214.

VILEY

- Riker, A. J. (1930). Studies on infectious hairy root of nursery apple trees. Journal of Agricultural Research, 41, 507–540.
- Ron, M., Kajala, K., Pauluzzi, G., Wang, D., Reynoso, M. A., Zumstein, K., Garcha, J., Winte, S., Masson, H., Inagaki, S., Federici, F., Sinha, N., Deal, R. B., Bailey-Serres, J., & Brady, S. M. (2014). Hairy root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene expression and function using tomato as a model. *Plant Physiology*, 166, 455–469.
- Shao, S., van Heusden, G. P. H., & Hooykaas, P. J. J. (2019). Complete sequence of succinamopine Ti-plasmid pTiEU6 reveals its evolutionary relatedness with nopaline-type Ti-plasmids. *Genome Biology* and Evolution, 11, 2480–2491.
- Shao, S., Zhang, X., van Heusden, G. P. H., & Hooykaas, P. J. J. (2018). Complete sequence of the tumor-inducing plasmid pTiChry5 from the hypervirulent Agrobacterium tumefaciens strain Chry5. *Plasmid*, 96–97, 1–6.
- Sharma, S., Shinde, S., & Verslues, P. E. (2013). Functional characterization of an ornithine cyclodeaminase-like protein of Arabidopsis thaliana. BMC Plant Biology, 13, 182.
- Slater, S. C., Goldman, B. S., Goodner, B., Setubal, J. C., Farrand, S. K., Nester, E. W., Burr, T. J., Banta, L., Dickerman, A. W., Paulsen, I., Otten, L., Suen, G., Welch, R., Almeida, N. F., Arnold, F., Burton, O. T., & Du, Z. (2009). Genome sequences of three Agrobacterium biovars help elucidate the evolution of multichromosome genomes in bacteria. Journal of Bacteriology, 191, 2501–2511.
- Slightom, J. L., Durand-Tardif, M., Jouanin, L., & Tepfer, D. (1986). Nucleotide sequence analysis of TL-DNA of Agrobacterium rhizogenes agropine type plasmid. Identification of open reading frames. Journal of Biological Chemistry, 261(1), 108–121.
- Stothard, P., & Wishart, D. S. (2005). Circular genome visualization and exploration using CGView. Bioinformatics, 21, 537–539.
- Thompson, M. G., Cruz-Morales, P., Moore, W. M., Pearson, A. N., Keasling, J. D., Scheller, H. V., & Shih, P. M. (2020). Draft genome sequence of Agrobacterium fabrum ARqua1. Microbiology Resource Announcements, 9, e00506-20.
- Tong, X., Li, Y., Xiang, T., Chen, Z., Wang, L., Zhang, C., Feng, M., & Wu, L. (2018). The complete genome sequence of cucumopine-type Agrobacterium rhizogenes strain K599 (NCPPB2659), a nature's genetic engineer inducing hairy roots. International Journal of Agriculture and Biology, 20, 1167–1174.
- Trovato, M., Maras, B., Linhares, F., & Costantino, P. (2001). The plant oncogene rolD encodes a functional ornithine cyclodeaminase. Proceedings of the National Academy of Sciences of the United States of America, 98, 13449–13453.
- Trovato, M., Mattioli, R., & Costantino, P. (2018). From A. *rhizogenes* RolD to plant P5CS: exploiting proline to control plant development. *Plants*, 7, 108.
- Watanabe, S., Sueda, R., Fukumori, F., & Watanabe, Y. (2015). Characterization of flavin-containing opine dehydrogenase from bacteria. *PLoS One*, 10, e0138434.

- Watanabe, S., Tozawa, Y., & Watanabe, Y. (2014). Ornithine cyclodeaminase/μ-crystallin homolog from the hyperthermophilic archaeon *Thermococcus litoralis* functions as a novel Δ1-pyrroline-2-carboxylate reductase involved in putative *trans*-3-hydroxy-lproline metabolism. *FEBS Open Bio*, 4, 2211–5463.
- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25, 1189–1191.
- Weisberg, A. J., Davis, E. W., Tabima, J., Belcher, M. S., Miller, M., Kuo, C.-H., Loper, J. E., Grünwald, N. J., Putnam, M. L., & Chang, J. H. (2020). Unexpected conservation and global transmission of agrobacterial virulence plasmids. *Science*, *368*, eaba5256.
- Weller, S. A., Stead, D. E., O'Neill, T. M., Hargreaves, D., & McPherson, G. M. (2000). Rhizogenic Agrobacterium biovar 1 and cucumber root mat in the UK. Plant Pathology, 49, 43–50.
- Wetzel, M. E., Olsen, G. J., Chakravartty, V., & Farrand, S. K. (2015). The repABC plasmids with quorum-regulated transfer systems in members of the Rhizobiales divide into two structurally and separately evolving groups. Genome Biology and Evolution, 7, 3337–3357.
- White, F. F., & Nester, E. W. (1980). Hairy root: plasmid encodes virulence traits in Agrobacterium rhizogenes. Journal of Bacteriology, 141, 1134–1141.
- White, F. F., Taylor, B. H., Huffman, G. A., Gordon, M. P., & Nester, E. W. (1985). Molecular and genetic analysis of the transferred DNA regions of the root-inducing plasmid of Agrobacterium rhizogenes. *Journal of Bacteriology*, 164, 33–44.
- Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K., Zhou, Y., Chen, L., Wood, G. E., Almeida, N. F. Jr, Woo, L., Chen, Y., Paulsen, I. T., Eisen, J. A., Karp, P. D., Bovee, D. Sr, Chapman, P., Clendenning, J., Deatherage, G., ... Nester, E. W. (2001). The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science*, *294*, 2317–2323.
- Xie, Z., & Tang, H. (2017). ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics*, 33, 3340–3347.
- Yost, C. K., Rath, A. M., Noel, T. C., & Hynes, M. F. (2006). Characterization of genes involved in erythritol catabolism in *Rhizobium leguminosarum* bv. Viciae. Microbiology, 152, 2061–2074.

**How to cite this article:** Hooykaas MJ, Hooykaas PJ. The genome sequence of hairy root *Rhizobium rhizogenes* strain LBA9402: Bioinformatics analysis suggests the presence of a new opine system in the agropine Ri plasmid.

MicrobiologyOpen. 2021;10:e1180. <u>https://doi.org/10.1002/</u> mbo3.1180

# APPENDIX





COG functional category

Information storage and processing	-	······································			
	А	RNA processing and modification			
	к	Transcription			
	L	Replication, recombination and repair			
	в	Chromatin structure and dynamics			
	D	Cell cycle control, cell division, chromosome partitioning			
	Y	Nuclear structure			
	V	Defense mechanisms			
	Т	Signal transduction mechanisms			
Cellular processes	М	Cell wall/membrane/envelope biogenesis			
and signaling	N	Cell motility			
	Z	Cytoskeleton			
	W	Extracellular structures			
	U	Intracellular trafficking, secretion, and vesicular transport			
	0	Posttranslational modification, protein turnover, chaperones			
	С	Energy production and conversion			
	G	Carbohydrate transport and metabolism			
	Е	Amino acid transport and metabolism			
	F	Nucleotide transport and metabolism			
Metabolism	н	Coenzyme transport and metabolism			
	I.	Lipid transport and metabolism			
	Р	Inorganic ion transport and metabolism			
	Q	Secondary metabolites biosynthesis, transport and catabolism			
Poorly	R	General function prediction only			
characterized	s	Function unknown			
L					

Translation, ribosomal structure and biogenesis

COG functional categories

J.

FIGURE A1 LBA9402 and K84 proteins were annotated with eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups (OGs)) mapper. Shown in the figure are the frequencies at which the COG functional categories were assigned to proteins (normalized to the total number of proteins). In cases where multiple categories were assigned to a single protein, each category was counted once. On the right, a description of each COG category is shown. (a) The relative frequencies are shown per LBA9402 replicon. (b) Functional categories are shown for the chromids (secondary chromosomes) of K84 and LBA9402

#### Rhizobium rhizogenes K84



FIGURE A2 The LBA9402 and K84 genomes were aligned using progressive Mauve. Colored boxes are local collinear blocks. Blocks below the center line are aligned in reverse complementary orientation compared to the reference (top) sequence, and blocks above the center line are in the forward orientation. Similarity profiles are shown inside the boxes



FIGURE A3 Multiple bacteriophage sequences were detected in LBA9402 chromosome 1. Phage search tool PHASTER was used to find prophage sequences. The schematic of the particular genomic region was generated by visualizing the raw output from PHASTER with genoplotR. The default PHASTER legend was slightly adapted by changing the color of protein categories not present in the displayed region to black. Genes annotated as "Hypothetical phage proteins" had a hit in the PHASTER phages database, but no gene function was known, whereas "Hypothetical proteins" did not have a hit in the PHASTER database



FIGURE A4 Comparison between the erythritol transport and catabolism regions of *Sinorhizobiummeliloti* 1021, *Rhizobiumleguminosarum* plasmid pRL12, and LBA9402. The red (same orientation) and blue (reverse orientation) ribbons show BLASTn hits, with darker shading indicating higher similarity. The data were visualized with genoplotR



**FIGURE A5** Map of pRi1855. From center to border, the rings represent GC skew, GC content, locations of genes predicted by PGAP (the colors of the arrows represent the COG functional categories predicted for the proteins; a description of the COG letters is provided in Appendix Figure A1), and finally the location of the number of features with their annotation



FIGURE A6 The region between two IS5 family insertion sequences shows similarity to a region in the *Rhizobiumlusitanum* strain 629 chromosome. A megablast search, with the sequence between both insertions sequences as a query, against the NCBI RefSeq genomes database (restricted to Rhizobiales sequences), and excluding Ri plasmid sequences, yielded the chromosome of *Rhizobiumlusitanum* strain 629 as best hit (67% query coverage, 94.50% sequence identity). Depicted are the areas with a high similarity between both sequences (plotted with genoplotR). Black arrows represent coding sequences, green arrows insertion sequences, blue arrows rRNAs, and pink arrows tRNAs

LBA9402_F3X89_28350 Chrv5_LecG	1	M I	6 1
Bo542_LecG	1	M	1
EU6_SacG	1	M	1
Pp_OpnDH_A	1	M	1
Ach5_OoxA	1	MTL	3
Bj_OpnDH_A Ph_LProDH1_A	1	MIVAP MLMR <mark>P</mark> LDLTEKR <mark>G</mark> KKVTIYFE <mark>G</mark> KELEAYE <mark>G</mark> EKL <mark>P</mark> VALLANEIYWLTTSNE <mark>G</mark> RKR <mark>G</mark> AFTF <mark>GP</mark> V <mark>P</mark> MTVN <mark>G</mark> VK	5 70
1 BADAD2 E3780 28350	7		18
Chry5_LecG	2	SRQVDLLIVGAGPAGMSAATELATAGCRVTVDMM	36
Bo542_LecG	2	SRHVDLLIVGAGPAGMTAAVAARRHGLNVLVVDEQ	36
C58 NoxA	2	······································	34 38
Pp_OpnDH_A	2	SRTFDVVIVGAGPAGMSAAVVLSEQGLQVLVVDEQ	36
Ach5_OoxA Bi OpnDH A	4	·····REVSVATD·····SVATD······SVATD······SVASUBASUBASUBASUBASUSUBASUSUSUSUSUSUSUSUS	47
Ph_LProDH1_A	71	GLEARRIKVKD <mark>G</mark> MKIERQ <mark>G</mark> YYDFHEE <mark>P</mark> VVE <mark>PG</mark> EIERVVVDVAII <mark>GGGPAG</mark> IGAALELQQ <mark>Y</mark> -LTVALIEER	139
LBA9402_F3X89_28350	49	PSPGGQIF-RAVEANHAARPNTESLLAALGPAYAAGFELVERFRSTPGIDYRPETTVWELRA	109
Chry5_LecG	37	PTPGGQIW-RNIEAVSGTPRMDILGKAYAEGLDLVRQFRAS-SARYEPGTQVWQVEA	91
B0542_LecG EU6_SacG	37	PSPGGQIF-RGVERNVKSPVFEALGTDYQKGQALTAAFRAS-GANYEPGTQVWQVED	91 89
C58_NoxA	39	SQPGGQ I W - RNAG RNAT S P I I N I L GAE Y RRG V RQ VEAF L AC - GAD Y I P DAQ V S RL S Q	93
Pp_OpnDH_A Ach5_OoxA	37 48	PAPGGQIW-RAVETIAHIATGDILGAEYKSGAELVQRFRAC-GAKYEPNTQMWKIEP PRPGGQIY-REITRNSPDRRTYLGPDYWKGQPLAEAFCLS-NVDYASRATVWSLETRDKTAGQA	91 109
Bj_OpnDH_A	42	I GPGGQVF - RAISST PVT DRNQLGADYWVGADL VQSLRAS - NAEV I QRAMVWSLD R	95
Ph_LProDH1_A	140	<u>G</u> WLGGDMWLKGIKQEGFNKDSRKVVEELVGKLNEN-TKIYLETSALGVFDKGE	191
LBA9402_F3X89_28350	110	DGTVGWLRDMSAGYLRASCVVLANGAMERPVPFPGWTLPGVLTAGAVQTLLKAGRLKPEGRIVLAGAGPL	179
Unry5_LecG Bo542 LecG	92 92	GHRATVTRUGAASSTEANTLLLATGAQERPAPFPGWILPGVMIIGAAQIVLKTSDQIPSEPVWIAGSGPL GHRAYVTRAGSASSIDATYLLLATGAQERPAPFPGWALPGVMTVGAAQIVLKTSDQIPSEPVWIAGSGPL	161 161
EU6_SacG	90	SWSLFL <mark>T</mark> SE <mark>G</mark> KARRITARAVLL <mark>S</mark> NGAQE <mark>R</mark> PVPFE <mark>G</mark> WTLP <mark>G</mark> VMTVGAAQILL <mark>KSGG</mark> MLPEGKVWIAGAGPL	159
C58_NoxA Pp_OppDH_A	94 92	GWT VEY VWGGE I RSVRGRHLLLATGAQERPVPFTGWTLPGVMTVGAAQILLKTAGQLPRGPVAVVGSGPL	163 161
Ach5_OoxA	110	RNVVGVTVAGSARMVETNAVVLATGAQERPMPVPGWTLPGVMTAGAAQIALKAAGAMPAGPVVLIGCGPL	179
Bj_OpnDH_A Rh   ProDH1_A	96 102	NLDIAV SVGGASAFVKAKRVILATGALERPFPIPGWTLPGVMTAGAAQTMLKSSALVPDGRTVIAGQGPL	165 258
	152		200
LBA9402_F3X89_28350	180 162	IMLLADOLRRLGVRPVLIARTDTFGDKFNALSKLRIA ALPALGKGLGWIASLKMAG IPMLTGISNLR	246
Bo542_LecG	162	SLLYAAQLLQAGGQIAGFLDTSRAGQISGQLSGVIPDLFAAFRNAPMDIVKGIGWLQSVKRR-VQYIKHVTEIE	230
EU6_SacG	160	PLLYATOFLNLGGRIAGYLDTAASPKLSAVSRLPRAWR - DFGGLLKGLRWLRDIKRSGM-MVRGFSDLR	226
C58_NOXA Pp_OpnDH_A	164	PLLYMQUMRLAGARPVAHLDTTPPRGLTSROLROPRGALQ-EPGQTLAGTAWLPGFSGVRHVRNVVRTS PLLYMAQLTRAGGKVAGWLDTTPPGGWRRALPWASSMVA-EFKEVSKGLAWLHETRRSGARRTKGVKELR	230
Ach5_OoxA	180	LYLLAS <mark>Q</mark> LVDA <mark>GVP</mark> DLTVLDTAQSPFR <mark>G</mark> AVLRHM <mark>P</mark> EFLL-SP-YVLKGIGLLLKV <mark>RR</mark> H-AQVVYGVRSIA	246
Bj_OpnDH_A Ph_LProDH1_A	166 259	LWLLAAQILRLGGRIDRILDTTERGNYFAALPHAFAFLT-SP-YFAKGLSMMREVKAK-VQVVTGVTELT KADEVIQELERWGIDYVHIPNVKR	232 282
	247		245
Chry5_LecG	231	ATGKESLERLRYVTASGQSATVDAKLLLVHEGIVPTIHPTLALGCRHVWNADQDSFAPELDSWGETSEAN	300
Bo542_LecG	231	ATGKESLEHLRYVTAS <mark>GQSGT</mark> VDAKLLLVHEGVIPTIHPTLALGCRHVWNGDQDSFAPELDSWGETSEAN	300
C58_NoxA	231	A GOOGLE TETW - EARGERRER VEADVLLVHEGTVPRTHETLALNODHHWNEEUGTLAARLDRWGETGREG AKGSGRLETVRFETSSGRSDRLEVKSLLVHEGLVPSHQLAVSAGAQLMWDVGQSAFRLERDEWMNAGPDG	295 300
Pp_OpnDH_A	231	AL <mark>GDG</mark> RLEQLQFRLKS <mark>G</mark> ELCTVPASVLLSHEGVIPSIHITQSLGCKHSWSAQQRCLVPDLDEWGETDQSG	300
Ach5_OoxA Bj OpnDH A	247	ASGDGQLANVSY - AVGQGERSTPAKSVLLHQGVTPSTSLSNAAGCELQWNDEURAFQPVTDHDGRTTKAG ASGDGQLANVSY - VAGGKRETLPVDLLLLHQGVVPNVNLAMAAGVEHRWDELQLCWSPVLDANGSSSVAG	315 301
Ph_LProDH1_A	283	YE <mark>GNEKVE</mark> R VIDMNNHEYKVDALIFAD <mark>G</mark> RRPDINPITQA <mark>GG</mark> KLRFR R <mark>G</mark> YYSPVLDEYHRI-KD <mark>G</mark>	345
LBA9402_F3X89_28350	316	GPALTSGPCRIFV <mark>SGD</mark> AR <mark>GIGGADAAMAHG</mark> RHVAASIIRDLGK <mark>S</mark> ISTKHGSAAKALSIAMA-G <mark>R</mark> PF	380
Chry5_LecG	301	······································	359
EU6_SacG	296		359
C58_NoxA	301	······································	359
Pp_OpnDH_A Ach5 OoxA	301 316	IYVAGDGAGIGGAKAACVRGELAAIGMAKRAGRMAADEWAKNIDALSRRLQILLR-LRPM	359 374
Bj_OpnDH_A	302	IAIA <mark>GDG</mark> AGIGGANAAVVR <mark>G</mark> RIAARAAVEALA <mark>P</mark> AAAAKLA-SMATLRADLAKAER-G <mark>R</mark> AF	359
Ph_LProDH1_A	346	······································	403
LBA9402_F3X89_28350	381	I DAAF PPGLAGRLPE - DDT I V <mark>CRCEEL DAGTLR</mark> KA I REGA - RDMNL V <mark>RG I LR</mark> CGMGPCQG <mark>R</mark> HCS I TLARL	448
Unry5_LecG Bo542 LecG	360 360	L DKL TKM KM S I F V M F - DHT I AUKUEE V SVASVKAQL I SG R - PGMNQ I KAFTRAGMGPCQGRQCGYT I ARV L DK F Y R P R P A V F V P P - DHT I ACRCEE V T V ASV R AQL SSG K - PGPNQ I KAFTRAGMGPCQGRQCGYT I ARV	427 427
EU6_SacG	355	LDAMYPPPRNRL D - DDVVVCRCEEVTAGAIRAAARNSP - ADPSAVKAATRCGMGPCQGRQCGYTVQAL	420
C58_NoxA Pp OpnDH A	360 360	L DGAYPPNIARIPPD - DATIVCRCEEL SAGVL HEAAVRGACRGPNQL KSFTRAGMGPCQGRQCGYPVHEL L DSMYPPRESIYTPS - DETVVCRCEEL TAGDIRKACAIAQ - PGVNQL KAFTRAGMGPCQGRQCGYTLASI	428 427
Ach5_OoxA	375	L DALYTPRQSFLAPSAPETIVCRCEEITVRKLREAIALGP-PGPNQLKTFVRCGMGQCQGRLCAATVTEI	443
ы_OpnDH_A Ph_LProDH1_A	360 404	LUILFRMAPQFRIPS-GUTIVCHCEEVTAKDVLDSVAIGA-TGPNQLKAYRTGMGPCQGRLCGLTVTEL LDKFNLEDVQICGC-DVSLKKVDEVIRKGI-TDLQIIKRLTHLAMGFCQGRYCLFNGAVV	427 461
	140		506
Chry5_LecG	428	LADAEGRAAPDVGLYRVRPPLKPVTLSELAALDSQERSS	466
Bo542_LecG	428	LADAEGRSAPDVGLYRVRPPLKPVTLSELAALDSQDETR	466
EUD_SacG C58_NoxA	421 429	VKSVA <mark>G</mark> LTAGEVGLFNPRPFVPLTVSMLAAQQAEEVVSASTHE	459 472
Pp_OpnDH_A	428	VAAEQQKAVEDVGFYRIRPPLKPITLGELASLDVEEK	464
Ach5_OoxA Bj OpnDH A	444 428	MAEERK VSPADVGLYRLRSPVKPVRLAELAHLPHIARALKAVTGRDPVDHDTTETGHIL MAQARGKSPQEIGYYRLRAPVKPITLAELAAVPKTEDDVKAVVRG	502 472
Ph_LProDH1_A	462	VSQRTGKKLSEIDLPVARSPIKNVKMGILARR	493

FIGURE A7 Multiple sequence alignment of  $\alpha$  subunits of various opine dehydrogenases. Indicated with orange triangles is the ADP binding motif Gly-X-Gly-X-X-Gly and with pink triangles the [2Fe-2S] iron-sulfur cluster binding motif (Cys-X-Cys-X<sub>31-33</sub>-Cys-X<sub>4</sub>-Cys). The residues putatively interacting with FMN are indicated with blue triangles. This analysis was based on the crystal structure of *Pyrococcushorikoshi* (Ph) L-Proline dehydrogenase (Ph\_L-ProDH) (Tsuge et al. 2005. Crystal structure of a novel FAD-, FMN-, and ATP-containing L-proline dehydrogenase complex from *Pyrococcushorikoshii*. J Biol Chem 280: 31045–31049). Shown are the amino acid sequences of opine dehydrogenase subunits encoded by the *Agrobacterium* Ti plasmids carried by strains Chry5, Bo542, EU6, C58, and Ach5, which function as succinamopine, nopaline, and octopine dehydrogenases, respectively. In addition, sequences from *Pseudomonasputida* (Pp) and *Bradyrhizobiumjaponicum* (diazoefficiens) USDA110 opine dehydrogenases (which function as nopaline and octopine dehydrogenases, respectively). Finally, the sequence of the L-Proline dehydrogenase of *Pyrococcushorikoshi* (Ph) is shown. The background color of the amino acids is according to the Clustal X color scheme.

WILEY

1 8 40 400 50000 00055			
LBA9402 F3X89 28355	1 MNEKVRDEV	TI <mark>GGGLQG</mark> CSTAFFLA <mark>R</mark> AG <b>RR</b> VTIVEKSTPGRHASGVNAGGLRLLL <mark>R</mark> DAR	REYPLSERAMEM 70
Chrv5 lecH	1 MSGNAL DT I \	VIGGGLHGLSAAMNLARAGKTVAILERSWVGRHSSGATAAGVRTLNRDLA	EIPISLEAMDM 70
5			
Bo542 LecH	1 MNAKAIDIIV	VI <mark>GGGLHGLSAALNLARAGKR</mark> VVILERSWVGRHSSGATAAGVRILN <mark>R</mark> DLA	ELPISLEAMDM 70
EU6 SacH	1 MN ADVL \	VIGGGLHGISAALQAAQRGVSVHLVERHELGSHASGATAAGVRTLGRDVA	ELPLSIEAAEA 67
<u> </u>			
EU6_Sace		V GGA V V GGALAYGI ANL GKKVI V L DGQ DQ DMRAARAN FGL VWI QGKGANAPA	TQQLIGESSDL /0
C58 NoxB	1 MDHE - FETAI	IV <mark>GGGVVG</mark> AAIGYGLA <mark>K</mark> LG <b>RR</b> VIIIDGADDALRASRTNFGLIWVQG <mark>K</mark> GDEYPA	YHRITROSAEL 72
Рр_ОДН_В		TIGGGMVGAATGYGLAGGKKRTLMLDGADTDYRAAKANFGLVWAHGKGMRDPN	IYQRL <mark>S</mark> LNAAKA 72
Ach5 OoxB/	1 M T	I I GGGL V GASI AWGLARSG TKPL V L DGADL DL RASRANFAL VWVQGKGL HAPH	YALWSDASARR 66
			V T VINT VOA DOA 70
BJ_OpnDH_B1	IMSGE-YDVA	VVGGGLLGSATAWGLGK GKKVAVLDEGDTTKRASRANFALVWVQSKGLGMPA	YIVWIVQASQA 72
Bi OppDH B2	1 MTGN - VDAIN	VIGGGIHGCSTTIHICIAGI KPVIJEKDYAGRHASGVNAGGVROJARHVP	FIPISIRSMGI 69
PII_LPIODH I_B		VIGGIVGVITAHELANRGEEVIVIERRFIGSGSTFRCGIGIRQQFNDEA	
1040400 50000 00055			
LBA9402_F3X89_28355	71 WENLDQFVGC	GPAAAAAEVCLGISQIALAVDREEMQWAESRSKDMIRRGIVIE	LVDAAELH <mark>R</mark> LL 134
Chrv5 lecH	71 WHNIERIVG	D SCGFHA YGQMSIAEKPEHLTVLEKRLAKIRSAG YDHEE	LIDRAELLRLV 129
Po542 LocH			
D0J42_Lech			
EU6_SacH	68 WHGMEALVG	D <u>-</u> - <u>D</u> CGFVA <u>C</u> GQLQVAEDEQSLNE I EARIKRLEAMG HHHEK	LVGPEEMQRLL 126
ELIG SacE		ESACVELDYSR PEGEVYCL SETEYENR ETL NKRTHNYL SASDL R	
LOO_OUCE			
C58_NoxB	73 WPGFASELKE	ELTGVDVEYRR NGGLLFCASEAQAEEE <u>-</u> KAIGEKMAAIHPDYRFE	MLERSRIE <mark>R</mark> ML 136
Pn ODH B	73 WPSEASOLE	EESGUSUDYEOKGGUNECUSDDEEDARAKAMDAWNROTPEUAPSAH	IMI DRAFI I RRF 138
, p_00, _D			
Ach5_OoxB/	67 WPIMANILL	DDSGIDVGLQQDGAFIFALSEEELEANRQDMESIELEINGRAPQFE	VLDRQQILDRV 132
Bj OpnDH B1	73 WGRLAFELKO	QQTGLDVSLQQ NGGFHLTLGEDEFGQRTELVKRMHNQVGAA DYKME	MLPASEVKKAL 138
ы_орноп_вг		ETTTE AV CRARVAELNAL OTTE HEE	LIDAIELKKLV 128
Ph LProDH1 B	71 WKKYSEEYG	FSFKQTGYLFLLYDDEEVKTFKRNIEIQNKFGVPTK	LITPEEAKEIV 126
=			
LBA9402_F3X89 28355	35 PGL T - NSA	ALGGLISRRDGHANPASAASAFRLAAEGAGVEIWEGCNMRGMSP-TSSGWR-T	ETDRGALESDV 202
Chrv5 lecH		CMGALLARN DGAADPHRTI KAEROATEAACYTIXEGACYLATEU AGADWR	
Chiry5_lech	30 <b>F</b> 5 <b>I</b> 5 - <b>F I</b> (	CMGALTARNDGAADFHRTLRAFRQATEAAGVTTTEGAGVTATEH-AGADWA-V	RIEIMEFIAFI 19/
Bo542 LecH	30 PHI S - PHC	CVGALIARNDGAADPHRTLRAFRHATEAVGVTIYEGAGVIAIEH-AGADWR-V	KTETMEFIAPT 197
		CLOCAMVSSDCSADDHDAVDAEDDAALKACVKLAEECLVLSMLD, SENGWD, V	
LOU_Sach			
EU6 SacE	35 PAVAFG - PEN	VVGASYCKADGHVNPLQLLAAMQRGVTAHGGKILFRNTVQSITP-GADGFT-V	TAALGSYKAPC 204
C58 NovB		VCSASESHMDGDVNPLLVL RAMLDGEVRLGGTLMTGEHVDSLAP, TLGGER, L	ATDGITIEAER 206
Pp_ODH_B	39 <u>PTM</u> KI <mark>G</mark> - KAN	V V <mark>G</mark> A S F <mark>G</mark> E F <mark>D G H A N P</mark> L R L L A A MQ K A F M A R <mark>G G</mark> Q L Q N N R P V T D I R R L <mark>S G G G</mark> F <b>E</b> - V	Q <mark>SG</mark> EYRAQCER 209
Ach5 OoxB/	33 LGI G - PEN	VVGSIYCAADGHVNALRLFHALHAAMERQGATYRPNHPVQSIEP-TTGGFI-L	KGEAFSILSRR 200
Bi OppDH B1		VECCENCEL DE HVINEL DE EDAE HE CEMAECUDYL DE DRVEALEK SCOEFE	TTRKCEL RAAK 200
ы_орпон_вт	39 FLI 0 - FLI	VSGSTCPLDGHVNSLKTPKAPHTGPMAPGTDTLPERPVSATSK-SGGEPK-L	I PROELRAAR 200
Bj OpnDH B2	29 PAV A - ET C	CPGGVVSRRDGAANPAQTTTAFRRKAEQLGATVREGVAASNIRY-SDGLWY-V	DV <mark>G</mark> SDT <mark>Y</mark> AAPV 196
Ph I ProDH1 B		VIAASWNPTDGKADPEEATTAEAVKAKEYGAKLI EYTEVKGELL.ENNELKGV	
LBA9402 F3X89 28355	203 VV <mark>N</mark> CA <mark>G</mark> AW <mark>G</mark> S	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG	SLVIGGGISGK 271
LBA9402_F3X89_28355		SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG	SLVIGGGISGK 271
LBA9402_F3X89_28355 Chry5_lecH	203 VV <mark>NCAG</mark> AW <mark>G</mark> 98 VV <mark>NAAG</mark> AWAN	SKIAA <mark>S</mark> VD EEL <mark>P</mark> VNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - F <mark>K</mark> QSAVG VKIAAMV <mark>G</mark> DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - F <mark>K</mark> QSDQG	SLVIGGGISGK 271 TLVIGGGLQGR 266
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH	203 VVNCAGAWG 98 VVNAAGAWA 98 VVNAAGAWA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG	SLVIGGGISGK 271 TLVIGGGLQGR 266 TLVIGGGLQGR 266
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH	203 VVNCAGAWG 98 VVNAAGAWA 98 VVNAAGAWA 98 VVNAAGAWA	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSAIAG DDIQIGHKASMVIVTERIAP - LLKPVVSSEGKKIS - LKOTSOG	SLVIGGGISGK 271 TLVIGGGLQGR 266 TLVIGGGLQGR 266
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH	203 VVNCAGAWG 98 VVNAAGAWA 98 VVNAAGAWA 95 VVNAAGAWA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG	SLVIGGGISGK 271 TLVIGGGLQGR 266 TLVIGGGLQGR 266 TLLIGGGSQGR 263
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE	203 VVNCAGAWG 198 VVNAAGAWA 198 VVNAAGAWA 195 VVNAAGAWA 195 VVNAAGAWA 205 VIVAAGHAT	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG	SLVIGGGISGK 271 TLVIGGGLOGR 266 TLVIGGGLOGR 266 TLLIGGGSOGR 263 TVMIGATKENV 267
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58 NoxB	203 VVNCAGAWG 98 VVNAAGAWA 98 VVNAAGAWA 95 VVNAAGAWA 205 VIVAAGAWA 205 VIVAAGAWA 207 VIIAAGNDS	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERIAP - LLKPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG	SLVIGGGISGK271 TLVIGGGLOGR266 TLVIGGGLOGR266 TLIGGSCGR263 TVMIGATKENV267 TFOLGVTNERV269
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB De_CDU_P	203 VVNCAGAWG 98 VVNAAGAWA 98 VVNAAGAWA 95 VVNAAGAWA 205 VVNAAGAWA 205 VIVAAGHAT 207 VIVAAGHAT	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLL V PEKGQLLITERVAP - VFPYAASG IRQTMTG	SLVIGGGISGK271 TLVIGGGLQGR266 TLVIGGGLQGR266 TLLIGGGSQGR263 TVMIGATKENV267 TFQLGVTNERV269
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B	203 VVNCAGAWA 98 VVNAAGAWA 995 VVNAAGAWA 95 VVNAAGAWA 95 VVNAAGAWA 205 VIVAAGHAT 207 VIIAAGNDS 210 IVIAAGLGC	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPVPASG LRQTSDG MAFAKSLD LPLLVPEKGQLLITERVAP - VFPYAASG LRQTAGG	SLVIGGGISGK271 TLVIGGGLQGR266 TLVIGGGLQGR266 TLLIGGGSQGR263 TVMIGATKENV267 TFQLGVTNERV269 TVMIGVTQEEV272
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5 OoxB/	203 VVNCAGAWG 198 VVNAAGAWA 198 VVNAAGAWA 195 VVNAAGAWA 195 VVNAAGAWA 105 VIVAAGAWA 107 VIVAAGHAT 101 VVIAAGLOCA	SKIAASVD - EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSQVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAEG KRLAPMVG - LSCPLKRSKGQILVTEKIQT - ALPCLSAG MRQADEG	SLVIGGGISGK271 TLVIGGGLQGR266 TLVIGGGLQGR266 TLIGGGSQGR263 TVMIGATKENV267 TFQLGVTNERV269 TVMIGVTQEEV272 GIMIGDSEETD263
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bi OppDH B1	203         VVNCAGAWA           98         VVNAAGAWA           99         VVNAAGAWA           205         VIVAAGAWA           207         VIAAGAA           207         VIAAGA           208         VVNAAGA           209         VVNAAGA           201         VIAAGAA           201         VIAAGAAA           202         VVAAAGAAA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVSSFGKKLS - LKQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTSDG AALGPMVG LNVPIRPQRGQLLVTERLAP - LLPIPGSG LRQTAEG KRLAPMVG LSCPLKRSKGQILVTERLAP - LLPIPGSG LRQTAEG	SLVIGGGISGK271 TLVIGGGLQGR266 TLLIGGGSQGR266 TLLIGGSQGR266 TVMIGATKENV267 TVMIGTKENV269 TVMIGVTQEEV272 GIMIGDSKEDD228
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1	203 VVNCAGAWG 98 VVNAAGAWA 998 VVNAAGAWA 95 VVNAAGAWA 105 VVNAAGAWA 105 VINAAGAWA 107 VIIAAGNDS 101 VVIAAGLDNF 201 IVLAAGLDNF 207 VVLAAGNAN	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERIAP - LLKPVVSVFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLLVTERLAP - LLPIPGSG LRQTAEG KRLAPMVG LSCPLKRSKGQILVTERTMP - FLPHPLTT IRQTDEG	SLVIGGGISGK271           TLVIGGGLOGR266           TLVIGGGLOGR266           TLVIGGGLOGR266           TLVIGGCSCGR263           TVMIGATKENV269           TGUSTKEV269           TOMIGATKENV269           TVMIGATKENV269           TVMIGATKEN269           TVMIGATKEN269           TVMIGATKEN269           TVMIGATKEN269           TVMIGDSKED262
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2	203         V V N CAGAWA           98         V V NAAGAWA           95         V NAAGAWA           95         V NAAGAWA           205         V I V AAGAWA           207         V I AAG L GC           201         I V L AAG L DNF           201         I V L AAG L DNF           207         V L AAG NANC           207         V L AAG AWA	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLVTERIAP - KISPVVSSFGKKLS - LKQTSDG MAFAKSLD LPLLV PEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG AALGPMVG - LSCPLKRSKGQILVTERLAP - LLPIPGSG IRQTAEG CTLAPMVG - LHAPMGPTRGQIVVTERIMP - FLPPLTT IRQTDEG GKIAAALG EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG	SLVIGGGISGK271           TLVIGGGLOGR266           TLVIGGGLOGR266           TLLIGGSSOGR263           TVMIGATKEN/267           TFQLGVTNER/269           TVMIGVTOEE/272           GIMIGDSKED266           TVMIGDSKED268           TVMIGDSKED268           TVMIGDSKED268           TVMIGDSKED268           TVMIGDSKED268           TVMIGDSKED268           TVMIGDSKED288           TVMIGDSKED288           TVMIGDSKED288           TVLIGGGHAT285
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B	203         VVNCAGAWA           98         VVNAAGAWA           98         VVNAAGAWA           95         VVNAAGAWA           95         VVNAAGAWA           95         VVNAAGAWA           95         VVNAAGAWA           96         VVNAAGAWA           97         VIAAGANA           98         VVNAAGAWA           99         VVNAAGAWA           90         VVAAGANA           91         VAAGAA           92         VVAAGAAA           93         VVAAGAWA           94         VVAAGAWA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAEG KRLAPMVG - LNVPIRPQRGQLLVTERLAP - LLPIPGSG LRQTAEG KRLAPMVG - LSCPLKRSKGQILVTERTMP - FLPHPLTT IRQTDEG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG	SLVIGGGISGK271 TLVIGGGLOGR266 TLVIGGGLOGR266 TLIGGSOGR266 TVMIGATKENV267 TFQLGVTNERV269 TVMIGVTQEEV272 GIMIGDSEETD263 TVMIGDSEETD263 TVLIGGGHLAT265 GIIGGICYEIG269
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B	203         V V N CAG AWG           198         V V NAAG AWA           105         V I V AAG AWA           107         V I AAG L DNE           101         I V L AAG L DNE           107         V L AAG L DNE           107         V L AAG L ANAG           197         L V NAAG AWA           197         V NAAG AWA           197         V NAAG AWA	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDLQIGHKASMIVTERSVP - KISPVVSVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTSDG KRLAPMVG - LNVPIRPQRGQLLVTERIAP - ULPIPGSG	SLVIGGGISGK271           TLVIGGGLOGR266           TLVIGGGLOGR266           TLVIGGGSOGR263           TVMIGATKENV267           TFQLGVTNERV269           TVMIGVTCEV272           GIMIGDSETD268           TVMIGDSEED269           TVMIGDSEED269           TVLIGGSACGR263           GIMIGDSEED269           TVLIGGSHLAT265           GIIGGIGYEIG289
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B	203         VVNCAGAWA           208         VVNAAGAWA           98         VVNAAGAWA           95         VVNAAGAWA           95         VVNAAGAWA           95         VVNAAGAWA           95         VVNAAGAWA           96         VVNAAGAWA           97         VIAAGADA           97         VVAAGAAAA           97         VNAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DKVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG - FKIPLKADRGQLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAGG AALGPMVG - LNVPIRPQRGQLLVTERLAP - LLPIPGSG LRQTAGG CTLAPMVG - LSCPLKRSKGQILVTERTAP - LLPIPGSG IRQTDG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG	S         L         V         I         GG         I         GG         X         271           T         L         V         GG         L         GG         Z         266           T         L         I         GG         L         GG         Z         266           T         L         I         GG         L         GG         Z         266           T         L         I         GG         S         GG         R         263           T         V         I         G         X         K         N         267           T         F         Q         V         T         R         269         272           G         I         I         G         S         K         T         263           T         V         I         G         S         K         T         263           T         V         I         G         G         H         1         265           G         I         G         G         G         Y         I         G         269
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355	203         VVNCAGAWG           198         VVNAAGAWA           98         VVNAAGAWA           95         VVNAAGAWA           97         VIAAGNA           97         VVLAAGAWA           97         VVNATNAWA           772         PCL         DDTSF	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTSDG AALGPMVG - LNVPIRPQRGQLLVTERLAP - LLPIPGSG LRQTAEG KRLAPMVG - LSCPLKRSKGQILVTERLAP - LLPIPGSG IRQTDEG QTLAPMVG - LHAPMGPTRGQIVVTERTMP - FLPHPLTT IRQTDEG GKIAAALG - EPVPETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG	SLVIGGGISGK271           TLVIGGGLOGR266           TLVIGGGLOGR266           TLVIGGGLOGR266           TVMIGATKENV267           TVMIGATKENV267           TGUMIGATKEN266           TVMIGDSEETD263           TVMIGDSKED260           TVMIGDSKED280           TVMIGDSKED280           TVMIGO384
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_locH	203         VVNCAGAWA           98         VVNAAGAWA           95         VVNAAGAWA           205         VIVAAGAWA           207         VIAAGADA           201         IVIAAGLGCA           201         IVIAAGLGCA           207         VVLAAGNA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVAAGAWA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG - FKIPLKADRGQLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAGG ALGPMVG - LNVPIRPQRGQLVTERLAP - LLPIPGSG LRQTAGG QTLAPMVG - LSCPLKRSKGQILVTERTQF - FLPHLTT	SLVIGGGISGK271 TLVIGGGLQGR266 TLLIGGSQGCR266 TLLIGGSQGR266 TVMIGATKENV267 TVMIGVTQEEV272 GIMIGDSETD263 TVMIGVTQEEV272 GIMIGDSKED266 TVMIGDSKE226 GIIGGIGYEIG269 VVIGCLGGIGYEIG269
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH	203         V V N CAG AWA           198         V V N AAG AWA           198         V V N AAG AWA           198         V N AAG AWA           198         V N AAG AWA           195         V V N AAG AWA           195         V V N AAG AWA           105         V I V AAG AWA           107         V I AAG ND SN           210         I V I AAG L G C           201         I V AAG L G C           201         I V AAG L G NANG           207         V L AAG L DNN           207         V L AAG L ANNO           207         V V AAG AWA           207         V L AAG L ANNO           207         P C L D D D T S F           207         A D L D AQ R S F	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTSDG AALGPMVG - LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAEG CRLAPMVG - LSCPLKRSKGQILVTERIAP - LLPIPGSG IRQTDEG QTLAPMVG - LHAPMGPTRGQIVVTERTMP - FLPHPLTT IRQTDEG GKIAAALG - EPVPETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG	SLVIGGGISGK271           TLVIGGGLOGR266           TLVIGGGLOGR266           TLVIGGLOGR266           TVMIGATKENV267           TFQLGVTNERV269           TVMIGDSEETD263           TVL16666
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH	203         VVNCAGAWA           98         VVNAAGAWA           95         VVNAAGAWA           205         VIVAAGAWA           205         VIVAAGAWA           207         VIAAGLONE           201         IVLAAGLONE           207         VVAAGAWA           207         PCL           200         DAGRSE           202         PCL         DDT SF           203         ADL         DAGRSE           204         AQRSE         N	SK IAAS VD EEL PVNVMALSMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDLQIGHKASMMIVTERSVP - KISPVVSVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG AALGPMVG - LSCPLKRSKGQIVTERIMP - LPPLTT IRQTMEG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAAT VAAFPALANISVVRSWTGLEGLTPDGVPYISPSARHPGLW	SLVIGGGISGK271           TLVIGGGLQGR266           TLVIGGGLQGR266           TLUIGGGSQGR263           TVMIGATKEN/267           TFQLGVTNER/269           TVMIGVTQEV272           GIMIGDSETD260           TVMIGQTQE222           TMIGDSETD260           TVMIGVTQE222           GIMIGDSEED260           TVMIGQTQE222           GIMIGDSEED260           TVMIGQTQE222           GIMIGDSE2202           TVMIGQTQE202           GIIGQIQE202           TVMIGQTQE202           TVMIGQTQE202           TVMIGQTQE202           TVMIGQTQE202           TVMIGQTQE202           TVMIGQTQE202           TVMIGQTQE202           TVMIGQTQE202           TVMIGQTQE202 <td< td=""></td<>
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH	203         V V N CAG AWG           198         V V N AAG AWA           198         V V N AAG AWA           198         V V N AAG AWA           195         V V N AAG AWA           195         V V N AAG AWA           195         V V N AAG AWA           105         V I V AAG AWA           107         V I AAG LOSA           210         I V I AAG LOGA           101         V AAG AWA           207         V I AAG LOGA           201         I V AAG LOSA           207         V L AAG LOSA           207         V L AAG LOSA           207         V V NAT NAWA           207         P CL D D D T S F           207         P CL D D Q R S F           207         P CL D D Q Q R S F           207         A D L D A Q R S F           207         A D L D A Q R S F	SKIAASVD EELPVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSQG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTSQG ALGPMVG - LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAEG KRLAPMVG - LSCPLKRSKGQILVTERIAP - LLPIPGSG LRQTAEG QTLAPMVG - LSCPLKRSKGQILVTERIAP - LLPIPGSG IRQTDEG QTLAPMVG - LSCPLKRSKGQILVTERTMP - FLPHPLTT IRQTDEG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPWVISFKYGHAYLTQTFHG VNFKELSKGARAATVAAFPALANISVVRSWTGLEGLTPDGVPYISPSARHPGLW VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASCVF	SLVIGGGISGK271 TLVIGGGLQGR266 TLVIGGGLQGR266 TLVIGGGLQGR266 TVMIGATKENV267 TVMIGVTQEV272 GIMIGDSEETD263 TVMIGDSEETD263 TVMIGDSEETD263 TVMIGDSEETD263 TVMIGDSEETD263 TVMIGDSEETD263 TVGGGLGGHLAT265 GIIGGIGYEIG269 TVLIGGGGGALAT265 GIIGGIGYEIG269 TVLIGGGGGALAT265 GIIGGIGYEIG269 TVLIGGGGGALAT265 GIIGGIGYEIG269 TVLIGGGGGALAT265 GIIGGIGYEIG269 TVLIGGGGGALAT265 GIIGGIGYEIG269 TVLIGGGGGALAT265 GIIGGIGYEIG269 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGGGGALAT265 TVLIGGALAT265 TVLIGGGALAT265 TVLIGGALAT265
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH	203         VVNCAGAWA           208         VVNAAGAWA           98         VVNAAGAWA           95         VVNAAGAWA           205         VIVAAGAWA           207         VIAAGAL           201         IVLAAGLONE           207         VVAAGAWA           97         VVAAGAWA           97         VVAAGAWA           97         VVAAGAWA           97         VVAAGAWA           97         VVAATNAWA           272         PCL DDDT SF           267         ADL DAQRSF           267         ADL DAQRSF           267         ADL DAQRSA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDLQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG AALGPMVG - LSCPLKRSKGQILVTERLAP - LLPIPGSG IRQTMTG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGSSNAAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRPVIGLSKKVEG I	SLVIGGGISGK 271 TLVIGGGLQGR 266 TLVIGGGLQGR 266 TLLIGGSQGR 263 TVMIGATKENV 267 TFQLGVTNERV 269 TVMIGDSKEDE 299 TVMIGDSKEDE 269 TVLIGGGHLAT 265 GIIGGIGYEIG 269 VVVFGFCGHGFQ 344 HSFGYSGHGFQ 339 HVFGFSGHGFQ 339
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE	203         VVNCAGAWA           98         VVNAAGAWA           99         VVNAAGAWA           205         VIVAAGAWA           205         VIVAAGAWA           207         VIAAGAUA           207         VVAAGAAA           208         VVAAGAAA           209         VVAAGAAA           201         VAAGAAA           202         VVAAGAAA           203         VVAAGAAA           204         AAARSEA           205         AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLVTERIAP - LLKPVVSSVGRPLS - FKQSDQG MAFAKSLD LPLLVPEKGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLVPEKGQLLVTERIAP - VFPYAASG LRQTAEG KRLAPMVG LSCPLKRSKGQILVTERTQT - ALPCLSAG LRQTAEG GTLAPMVG LAPMGPTRGQILVTERTMP - FLPHPLTT IRQTDEG GKIAAALG EPVPVETVAPMINITSRVPH -F IDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VNFKELSKGARAATDLFPIVGQLRIVTEKTQTAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVTRWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVTRWAGMEAMTSDHRPVIGLSKKVEGLI	SLVIGGGISGK271           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGCVTNERV269           TVMIGVTQEV272           GIMIGVTQEV272           GIMIGDSETD26           TVMIGVTQEV272           GIMIGDSETD26           TVMIGVTQEV272           GIMIGDSETD263           TVMIGVTQEV272           GIMIGDSETD263           TVMIGVTQE202           GIGGIGYEIC           TVLIGGGHLAT265           GIIGGIGYEIG299           HVFGFCGHGFQ344           HSFGYSGHGFQ339           HVFGFSGHGFQ338           HVFGFSGHGFQ338           HVFGSGGGLG70333           HVFGSGHGFQ338           HVFGSGHGFQ338           HVFGSGHGFQ338           HVFGSGHGFQ338           HVFGSGHGFQ338           HVFGSGHGFQ338           AACHS-SVT3333
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58 NoxB	203         V V N CAG AWG           198         V V NAAG AWA           197         V NAAG AWA           107         V I AAG LONE           101         I V L AAG LONE           107         V L AAG LONE           108         V NAAG AWA           1097         V L AAG LONE           101         V L AAG LONE           102         P C L D D D T S F           103         A D L D AQ R S F           104         LAA D ROSAN           105         A D L D AQ R S F           106         L AA D ROSAN           107         R S T D Y TA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DDLQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP -VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERTAP - FLPHPLTT IRQTMTG GKIAAALG EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRPVIGLSKKVEG I SATKLA LRATRIIPALAPARLVRQWGGLRVLTPDGVPYADSSRYPGIF	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L L I G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G V T Q E V 272 G I M I G D S K E D E 269 T V I I G G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 T V G F S G H G F Q 339 H S F G Y S G H G F Q 339 H S F G Y S G H G F Q 339 H S F G Y S G H G F Q 339 H S F G F S G H G F Q 333 A A A C H S - G V T 336 A A A C H S - G V T 336
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pn_OH_P	203         VVNCAGAWA           98         VVNAAGAWA           95         VVNAAGAWA           205         VIVAAGAWA           207         VIAAGAWA           207         VIAAGANAGAWA           201         IVLAAGLOCA           201         IVLAAGAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVLAAGAWA           207         VVNATNAWA           207         PCLDDDTSF           208         AUDARSF           207         PCLDDARSF           208         T           209         CENTUTA           200         R           201         AUROSA	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLVTERIAP - LLKPVVSVVGRPLS - FKQSDQG MAFAKSLD - LPLLLVPEKGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLVVPEKGQLLVTERIAP - VFPYAASG LRQTAGG QTLAPMVG LSCPLKRSKGQIVVTERLAP - LLPIPGSG LRQTAGG QTLAPMVG LHAPMGPTRGQIVVTERTMP - FLPHPLT IRQTDEG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRPVIGLSKKVGGLI ASATKLA LRATRIIPALAPARLVRQWSGFRVLPPDGTPIYAQSEEYPGLF	S L V I G G G L G G R 261 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G V G R 263 T V M I G A T K E NV 267 T V I G A T K E NV 267 T V I G S K E D 263 T V M I G D S E T D 263 T V M I G D S K E D 263 T V I G G H L A T 265 G I I G G I G Y E I G 269 V V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 336 A A A C H S G V T 333 V A A C H S G V T 333 V A C H S G V T 333 V A C H S G V T 333 V A C H S G V T 333
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B	203       V V N C AG AWG S         198       V V N AAG AWA N         98       V V N AAG AWA N         98       V V N AAG AWA N         98       V V N AAG AWA N         99       V V N AAG AWA N         99       V V N AAG AWA N         99       V V N AAG AWA N         205       V I V AAG AWA N         207       V I AAG NA NG NANG         201       I V L AAG L D N         207       V L AAG NA NG         207       P C L D D D T S F         207       A D L D AQ R S F         208       T D RG V T V A         209       N D RG V T V A         200       G R - S T D V T A A         201       G R - S L A T T	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLVTERIAP - LLPIPGSG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERIAP - LLPIPGSG IRQTMTG GKLAPMVG - LSCPLKRSKGQILVTERIAP - LLPIPGSG IRQTDEG QTLAPMVG - LSCPLKRSKGQILVTERTMP - FLPHPLTT IRQTDEG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAATVAAFPALANISVVRSWTGLEGLTPDGVPYISPSARHPGLW VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRPVIGLSKKVEGII SATKLA LRATRIPALAPARLVRQWSGFRVLPPDGTPIYAQSEYPGIH EGAARIT RKALKILPDLAQAKLVRQWSCLRIMTPDGNPVYARSLTHPGAE	SLVIGGGISGK271           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGGSQGR263           TVMIGATKENV267           TFQLGVTNERV269           TVMIGDSEE           TVMIGDSEE           TVGGGLQGR263           GIMIGDSEE           GIMIGDSEE           TVLIGGGLQGR263           GIMIGDSEE           GIMIGDSEE           TVLIGGGHLAT           GIGIGGIGYEIG           TVESCHGFQ38           HYFGFCGHGFQ38           HYFGFSGHGFQ38           HYFGFSGHGF338
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/	203       VVNCAGAWA         98       VVNAAGAWA         99       VVNAAGAWA         205       VIVAAGAWA         205       VIVAAGAWA         207       VLAAGLOND         201       IVLAAGLOND         207       VLAAGAAA         207       VLAAGAWA         207       VLAAGAWA         207       VLAAGAAA         207       VLAAGAAAA         207       PCL       DDT         207       PCL       DDT         207       PCL       DACRSF         208       I - DRC       VIVA         209       VVNAT       NAWA         201       DACRSF       NAWA         202       PCL       DACRSF         203       GF - DLAT       T         204       NT       NT         205       GR - T       NT         206       NT - RISSEN       SP	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERLAP - LLPIPGSG LRQTAEG QTLAPMVG LSCPLKRSKGQIVTERTMP - FLPHPLTT IRQTMEG GKIAAALG - EPVPVET VAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAATDLFPIVGQLRIVTERTMQAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVTRWAGMEAMTADHLPIIGSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVTRWAGMEAMTSDHRPVIGLSKVEGI ASATKLA LRATRIIPALAPARLVRQWSGFRVLPPDGTPIYAQSEEYPGLF GARHIA NRAINVMPDVGALRVIRQWGCLRIVTPDGNPVYARSLTHPGAE	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G L Q G R 266 T L U I G G S Q G R 263 T VM I G A T K E NV 267 T VM I G A T K E NV 267 T VM I G V T Q E E V 272 G I M G D S E T D 263 T VM I G D S K E D E 269 T VM I G G G H L A T 265 G I I G G I G Y E I G 269 T V I G G G H L A T 265 G I I G G I G Y E I G 269 T V G F S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 333 V V A C H S - G V T 333 V A C H S - G V T 339 I V A C H S - G V T 330
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00xB/ Bj_OnDH_B1	203         VVNCAGAWAS           198         VVNAAGAWAS           98         VVNAAGAWAS           98         VVNAAGAWAS           98         VVNAAGAWAS           98         VVNAAGAWAS           98         VVNAAGAWAS           99         VVNAAGAWAS           99         VVNAAGAWAS           905         VVNAAGAWAS           205         VIVAAGAWAS           207         VIAAGAMAS           201         IVAAGLOCK           202         PCLDDDTSEL           203         IVAAGLOCK           204         IVAAGLOCK           205         AULDACKS           206         AULDACKS           207         GR-STDVTA           208         F-DLATTE           204         TRSPE           205         F.TESPE	SKIAASVD EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTMTG AALGPMVG - LNVPIRPQRGQLVTERIAP - LLPIPGSG LRQTAG GKLAPMVG - LNVPIRPQRGQLVTERIAP - LLPIPGSG IRQTDG QTLAPMVG - LNVPIRPQRGQLVTERTAP - FLPHPLTT IRQTDG GKIAAALG - EPVPETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQXITOPIKRGTINPMVISFKYGHAYLTQTFHG VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRPVIGLSKKVEGLI SASATKLA RRATRIPALAPISQUVVIRTWAGMEAMTSDHRPVIGLSKKVEGLI GARHIA NRAINVMPDVGALRVIRQWSGFRVLPPDGTPIYAQSEEYPGLF GARHIA NRAINVMPDVGALRVIRQWSCLRIMTPDGNVYARSLTHPGAE	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L L I G G S Q G R 263 T VM I G A T K E N V 267 T F Q L G V T N E R V 269 T VM I G T C E V 272 G I I G G I G Y E I C 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G G H C T 333 I V F G F S G H G F Q 336 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 I V A C H S - G V T 333 I V A C H S - G V T 339 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 340 L V A C H S - G V T 34
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 OC D_D_S	203       V V N C A G A WA         98       V V N A A G A WA         98       V V N A A G A WA         98       V V N A A G A WA         98       V V N A A G A WA         98       V V N A A G A WA         98       V V N A A G A WA         99       V V N A A G A WA         905       V I V A A G A WA         205       V I V A A G A WA         206       V I V A A G A WA         207       V L A A G L D N H         207       V L A A G A WA         207       P C L D D D T S F H         206       A D L D A Q R S F         207       G R - S T D V T A         208       T - D R C V T A         2070       G R - S T D V T A         2070       G R - S T D V T A         2070       D - D R T L K H	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG AALGPMVG - LSCPLKRSKGQIVTERLAP - LLPIPGSG IRQTMTG AALAGPMVG - LSCPLKRSKGQIVTERTMP - FLPHPLT IRQTMTG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASCVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGSPNASCVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTSDHRPVIGLSKKVEGII ASATKLA RKATRIPALAPARLVRQWSGFRVLPPDGTPIYAQSEEYPGLF GGARHIA NRAINVPDQALRIVRTWAGMEAMTSDHRPVIGLSKKVEGII AGARHIA NRAINVPDVGALRVIRQWSCLRIMTPDGNPVYARSLTHPGAE DISAVLA SRALRIFPALSDLNVRSWTGFRVKTADGVPIYDHSERYPGAF	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T V M I G A T K E NV 267 T F Q L G V T N E RV 269 T V M I G V T Q E E V 272 G I M G D S K E D E 269 T V I G G G H L AT 265 G I I G G I G Y E I G 269 T V I G G G H L AT 265 G I I G G I G Y E I G 269 H V F G F C B H G F Q 339 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G F S G H G F Q 339 I V G C F S - G V T 330 V A C H S - G V T 330 V A C H S - G V T 330 V A C C H S - G V T 330 V A C C H S - G V T 330
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00xB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00xB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2	203       V V N C AG AWG S         198       V V N AAG AWA N         198       V N AAG AWA N         195       V N AAG AWA N         195       V N AAG AWA N         195       V N AAG AWA N         107       V I AAG LOS N         101       V I AAG LOS N         101       V I AAG LOS N         107       V I AAG LOS N         108       I - D RG V T V         1097       V N AAG AWAT         1007       V I AAG LOS N         101       D AQ R S F         102       D AQ R S F         103       A A D ROS AN         104       A D ROS AN         105       G F - D R S T D T A         106       A T - R I S S P         107       L - D R T L KHS         104       P Q D R N E T V	SK I A A S V D E E L P V N V MAL S MM V T T R V A P - F V K P V V I G I D R P L S - F K Q S A V G VK I A A M V Q D D I Q I G H K A S M I V T E R I A P - L L K P V V S V V G R P L S - F K Q S D Q G D H V S A A G - D P L Q H A I R T S M V V T E R I A P - L L K P V V S V V G R P L S - F K Q S D Q G P S L T D P L G F K I P L K A D R G Q L L V T E R I A P - L L P Y P A S G L R Q T S D G M A F A K S L D - L P L L L V P E K G Q L L V T E R I E P - L L P Y P A S G L R Q T S D G A A L G P M V G - L N P L R L V P E K G Q L L V T E R L A P - V L P I P G S G	SLVIGGGISGK271           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGGLQGR266           TVMIGVTQEV272           GINGST           TVMIGVTQEV272           GINGST           TVMIGVTQEV272           GINGSKED260           TVMIGVTQEV272           GINGSKED260           TVLIGGGHLAT265           GIIGGIGYEIG269           HVFGFCGHGFQ330           HVFGFSGHGFQ330           HVFGFSGHGFQ333           VACHS-GVT333           VACHS-GVT333           VACHS-GVT333           VACHS-GVT333           VACHS-GVT333           VACHS-GVT333           VACHS-GVT333           VACHS-GVT333
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacA EU6_SacA EU6_SacB C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B	203       V V N CAGAWA         208       V V NAAGAWA         98       V V NAAGAWA         99       V V NAAGAWA         205       V I V AAGAWA         207       V I AAG NANC         207       V I AAG AGAWA         97       V NAAGAWA         272       P C L D D D T S F         266       A D L D AQ R S F         270       G R - S T D V T AA         270       G R - S T D V T AA         270       G R - S T D V T AA         270       G R - S T D V T AA         270       G R - S T D V T AA         273       G F - D L A T T T E         264       N T - R I S S S P         270       D - D R T L K H         266       P Q D R N E T V	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAGG MAFAKSLD - LPLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERIAP - LLPIPGSG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERTAP - FLPHPLTT IRQTAGG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHPVIGLSKKVEGLI SATKLA NRAINVMPDVGALRVIRQWSGFRVLTPDGTPIYAQSEYPGIF EGAARIT RKALKILPDLAQAKLVRQWSCLRIMTPDGNPVYARSLTHPGAE DISAVLA SRALRIFPALSDLNVVRSWIGFRVKTADGVPIYDSSRHKGLT SISAVMT DRAQRMFPHLSRLNVIRSWAGIRVMPQDFPIYDGSSRHKGLY LVMKSLAISARTVFELFPWMRSATIVRAWGGIEAKKTDSNPAVGSFRVKTADGVPIYDSSRHKGLT SISAVMT DRAQRMFPHLSRLNVIRSWAGIRVMPQDSSRHKGLTPSSRHKGLY SSRKKZARTT RKALKILPDLAQAKLVRQWSCLRIMTPDSNPAJGRIFFSSRHKGLY SSRKKZARTT SRALRIFPALSDLNVVRSWIGFRVKTADGVPIYDSSRHKGLY SSRKKZARTDLFPALSDLNVVRSWIGFRVKTADGVPIYDSSRHKGLY LDWKSLAISARTVFELFPWMRSATIVRAWGIEAKKTDDSNPAJGRIFFSSRHKGLY SSRHKZARTDLFPALSDLNVVRSWIGFRVKTADGVPIYDSSRHKGLY SSRHKZARTDLFPALSDLNVVRSWIGFRVKTADGVPIYDRSSRYSAR SISAVMT DRAQRMFPHLSRLNVIRSWAGIRVMFONGNPVYARSLTHPGAF	SLVIGGGISGK271           TLVIGGGLQGR266           TLVIGGGLQGR266           TLVIGGCLQGR266           TLVIGGCLQGR263           TVMIGATKENV267           TFQLGVTNERV269           TVMIGDSKEDE299           TVLIGGCHLAT265           GIMIGDSKEDE299           TVLIGGCHLAT265           GIIGGIGYEIGE2033           HSFGYSGHGFQ339           HSFGYSGHGFQ339           HVFGFSGHGFQ339           HVFGFSGHGFQ336           VXGCHS-GVT336           IATCHS-GVT330           VACCHS-GVT336           HQFGFSCHGFQ338           IAFCHS-GVT330           VACCHS-GVT336           IAAGFSGHGFQ338           IAAGFSGHGFN339
LBA9402_F3X89_28355 Chry5_lecH B0542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH B0542_LecH EU6_SacH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B	203       VVNCAGAWA         98       VVNAAGAWA         99       VVNAAGAWA         205       VIVAAGAWA         205       VIVAAGAWA         207       VLAAGLODA         201       IVLAAGLONE         207       VVAAGAWA         207       VVAAGAWA         207       VVAAGAWA         207       VVAAGAWA         207       VVAAGAWA         207       PCL       DDTSF         207       PCL       DDTSF         208       LAADROSA       A         209       GR - STUVA       A         200       GR - STUVA       A         201       D - DRTLKH       A         202       Q - DRTLKH       A         203       GF - DLATTF       A         204       NT - RISSSP       A         205       L D - DRTLKH       A         206       PYQDRN	SK IAAS VD EEL PVNVMALSMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG PSLTDPLG FKIPLKADRGQLVTERIAP - LLKPVVSVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD - LPLLVVPKGQLLITERVAP - VFPYAASG LRQTAGG QTLAPMVG - LSCPLKRSKGQIVTEKTQT - ALPCLSAG	SLVIGGGISGK271           TLVIGGGLQGR266           TLVIGGGLQGR266           TLUIGGGLQGR266           TLUIGGGV07           TFQLGV17           TVMIGV16           TVMIGV16           TVMIGV16           TVMIGV16           TVMIGV16           TVMIGV16           TVGGG10           TVGG10           TVG10           TUG0           TVG10           TUG0           TUG0           TUG0           TUG0           TUG0           TUG0
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B	203       VVNCAGAWA         198       VVNAAGAWA         197       VNAAGAWA         107       VIAAGNANC         107       VIAAGLOH         107       VIAAGLOH         107       VIAAGLAGA         107       VVAAGLAGANAC         107       VVAAGLAGANAC         107       VVAAGLAGAAA         107       VVAAGLAGAAAA         107       VVAAGLAGAAAA         107       VAAGAAAA         107       VAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG IROTMTG AALGPMVG - LNVPIRPQRGQLLVTERIAP - LLPIPGSG IROTMTG GKLAPMVG - LSCPLKRSKGQILVTERIAP - LLPIPGSG IROTMTG AALGPMVG - LNVPIVERSVGQLLVTERIAP - LLPIPGSG IROTMTG AALGPMVG - LSCPLKRSKGQILVTERTAP - FIDPVVILRGRKLS - FKQFSNG GKLAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRVVIGLSKKVEGIF SATKLA RATRIPALAPARLVRQWSGFRVLTPDGTPIYAQSEYPGJF GAARIT RKALKILPDLAQAKLVRQWSCLRIMTPDGNVYARSLTHPGAE DISAVLA SRALRIFPALSDLNVVRSWTGFRVKTADGVPIYDRSRYPGJF SISAVMT DRAQRMFPHLSRLVRZWGFRVKTADGVPIYDRSRYPGJF SISAVMT DRAQRMFPHLSRLVRZWGFRVKTADGVPIYDGSETHPGAF SISAVMT RATRIPALSDLNVVRSWTGFRVKTADGVPIYDGSETHPGAF SISAVMT RAARVEFPLSRLNVIRSWAGIEAMKDDIPVFGPSSHKGLY DIYALAASVQAAVRVFPLSRLNVIRSWAGIEAMKDDIPVFGPSSHKGLY SISAVMT RAARVFFPLSRLNVIRSWAGIENVLTPDGNVYARSLTHPGAF SISAVMT RAARVFFPLSRLNVIRSWAGIEAMKDDIPVFGPSSHKGLY DISAVLA SRALRIFPALSDLNVVRSWTGFRVKTADGVPIYDRSRYPGAF SISAVMT RAARVFFPLSRLNVIRSWAGIEAMKDDIPVFGPSSHKGLY EFLREVS YYFTKIIPALKNLIL	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L L I G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G V T Q E V 272 G I M I G D S K E D E 269 T VM I G D S K E D E 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 333 I A T C H S - G V T 336 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_	203       VVNCAGAWA         98       VVNAAGAWA         95       VNAAGAWA         95       VVNAAGAWA         205       VIVAAGAWA         205       VIVAAGAWA         207       VLAAGLDNF         201       IVLAAGLDNF         207       VLAAGLANA         207       VLAAGAWA         207       VLAAGLANA         207       VLAAGAWA         207       PCLDDDTSF         206       ALADROSA         2070       GR-STDVTA         2070       GR-STDVTA         2070       LD-DRTLKHS         2070       LD-DRTLKHS         2070       LD-DRTLKHS         2070       LD-DRTLKHS         2070       LD-DRTLKHS         2070       LSPAVGETV	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQ SAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQ SDQG DHVSALAG - DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVVGRPLS - FKQSDQG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAEG (KRLAPMVG - LNVPIRPQRGQLVTERLAP - LLPIPGSG LRQTAEG (KRLAPMVG - LSCPLKRSKGQIVTERIMP - LPPLT IRQTMEG (KRLAPMVG - LSCPLKRSKGQIVTERIMP - FLPHPLT IRQTAEG (KRLAPMVG - LSCPLKRSKGQIVTERIMP - FLPHPLT IRQTAEG (KRLAPMVG - LSCPLKRSKGQIVTERIMP - FLPHPLT	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G L Q G R 266 T L U I G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S E T D 263 T VM I G D S E T D 263 T VM I G D S K E D E 269 T VM I G G K E A Z 255 G I I G G I G Y E I G 269 T V I G G G H L A T 265 G I I G G I G Y E I G 269 T V G F S G H G F Q 334 H S F G Y S G H G F Q 334 H S F G Y S G H G F Q 338 H V F G F S G H G F Q 333 V A C H S G V T 333 V A C H S G V T 336 L A T C H S G V T 336 H Q F G F S L H G F Q 338 I A A G F S C H G F M 339 386
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH	203       VVNCAGAWAS         198       VVNAAGAWAS         198       VVNAAGAWAS         198       VVNAAGAWAS         198       VVNAAGAWAS         205       VIVAAGAWAS         205       VIVAAGAWAS         205       VIVAAGAWAS         205       VIVAAGAWAS         205       VIVAAGAWAS         207       VIAAGNOS         201       IVAAGLOS         202       PCL       DDTSF         203       IVAAGNAS         204       IAADROSA         205       ADLDAQRSF         206       ADLDAQRSF         2070       GR-STDVTA         2070       GR-STDVTA         2070       GR-STDVTA         2070       PTYD-TTTT         2070       PTYD-TTTT         2070       PTYD-TTTT	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAG ALGPMVG LNVPIRPQRGQLLVTERIEP - LLPIPGSG LRQTAG GKLAPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG IRQTMTG AALGPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG IRQTMTG GKLAPMVG LNVPIRPQRGQLLVTERTAP - FLPHPLTT IRQTDG QTLAPMVG LNVPIRPQRGQLLVTERTAP - FLPHPLTT IRQTDG GKIAAALG EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG SVNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRPVIGLSKKVEGLI ASATKLA RATRIIPALAPARLVRQWSGFRVLTPDGVPYJRSRXPGIH GARRIA NRAINVMPDVGALRVIRQWGGLRVLTPDGVPYJRSRXPGIH SISAVMA SRALRIFPALSDLNVVRSWTGFRVKTADGVPYVDRSRTYPGIH EGAARIT RKALKILPDLAQAKLVRQWSCLRIMTPDGNPVYARSLTHPGAE JISAVKA RARGIFPALSDLNVVRSWTGFRVKTADGVPIYDSSRKVGGI SISAVKA RARGIFPALSDLNVVRSWTGFRVKTADGVPIYDSSRKYGFH SISAVKA RARGIFPALSDLNVVRSWTGFRVKTADGVPIYDSSRKGLY EFLREVS YYFTKIIPALKNLLILRTWAGYYAKTPDSNPAIGRIEELNDYY ARSLTTGSLDPRL - APFAADRFAKDTTRAKEALR - TDLIVHGGTNRVI - EPFSAKGLIRKFVRHVAH	S L V I G G G I S G K 271 T L V I G G L Q G R 266 T L V I G G L Q G R 266 T L U I G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G V T Q E V 272 G I M I G D S K E D E 269 T VM I G D S K E D E 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 I V A C H S - G V T 333 I A T C H S - G V T 330 V A C C H S - G V T 330 I A A G F S G H G F M 339 386 370
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH EU6_Sa	203       V V N CAGAWA         208       V V NAAGAWA         98       V V NAAGAWA         99       V V NAAGAWA         205       V I V AAGAWA         205       V I V AAGAWA         207       V I AAG L D NE         201       I V L AAG L D NE         207       V L AAG AWA         207       V L AAG L D NE         207       V L AAG L D AC R S F         208       T - D R C V T VE         2070       G R - ST D V T AF         2070       G R - ST D V T AF         2070       C R - ST D V T AF         2070       C R - ST D V T AF         208       T - R I S S P I         209       L T R I S S P I         200       D - D R T L K HE         201       D - D R T L K HE         202       L S P AVG E T VE	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQ SAVG VKIAAMVG DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQ SDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG IRQTMTG AALGPMVG LNVPIRPQRGQLLVTERLAP - LLPIPGSG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG AALAPMVG LSCPLKRSKGQIVTERLAP - LPIPGSG IRQTMTG AALGPMVG LHAPMGPTRGQIVVTERIMP - FLPHPLTT IRQTMEG GKIAAALG EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASCVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGSPNASCVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTSDHRPVIGLSKKVEGLI ASATKLA RRATRIPALAPARLVRQWSGFRVLPPDTPIYQQSETYPGLF GGARHIA NRAINVPDQALRVIRQWSGFRVLPPDGNPYYARSLTHPGAE DISAVLA SRALRIFPALSDLNVVRSWTGFRVKTADGVPYDRSRRYPGIH GGARHIA NRAINFPALSDLNVRSWTGFRVKTADGVPYDRSRRYPGAF SISAVMT DRAQRMFPHLSRLNVIRSWAGIRVMPQDGFPIYDASETHPGAE DISAVLA SRALRIFPALSDLNVRSWTGFRVKTADGVPIYDRSRRYPGAF SISAVMT ORAQRMFPHLSRLNVIRSWAGIRVMPQDGFPIYDQSETHPGAF LDWKSLAISARTVFELFPVMRSATIVRAWAGYAKTPDSNPAIGRIEELNDYY ARSLTGSLDFRL - APFAADRFAKDTRAKEALR - TDLIVNGGTNRVI - EPFSAKELIRKEYRNVAH	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G K S Q R 263 T V M I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S E T D 263 T VM I G D S E T D 263 T VM I G D S E T D 263 T VM I G G G H L AT 265 G I I G G I G Y E I G 269 T V G F S G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 I V A C H S - G V T 336 V A C H S - G V T 336 V A C C H S - G V T 336 I A C F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 386 379
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH	203       VVNCAGAWG         198       VVNAAGAWA         98       VVNAAGAWA         99       VVNAAGAWA         95       VVAAGAWA         905       VIAAGAWA         905       VIAAGAWA         907       VIAAGAWA         907       VIAAGAWA         907       VIAAGAWA         907       VVAAGLAGUA         907       VAAGLAGUA         907       VAAGUA         907       VAAGUA         907       VAAGUA         907	SK IAAS VD EE L P VNVMAL SMMV TT RVAP - FVK PVV I G I DRPLS - FKQ SAVG VK IAAMVG DD I Q I GHKASMM I VTER I AP - LL KPVVS VVG RPLS - FKQ S D Q G DHV SALAG DP L QHA I RT SMMVVTER SVP - KI SPVVS SFGKKLS - LKQ TS Q G PS LT DP L G FK I PL KADR Q LL VTER I EP - LL P YPASG L RQ T S D G MAFAK SL D L PL LL V PEK GQ LL VTER I EP - LL P YPASG L RQ T M TG AAL GPMVG LNVP I RPQ RGQ LL VTER LAP - LL P I PGSG L RQ T M TG AAL GPMVG LNVP I RPQ RGQ LL VTER LAP - LL P I PGSG I RQ TM TG AAL GPMVG LNVP I RPQ RGQ LL VTER LAP - LL P I PGSG MRQ AD EG QT LAPMVG L NP VET VAPMLMI TS RVPH - FI D PVVI L RG RKLS - FKQ FS NG NL I NAMAG I KT KI P I EP Y KHQ AVI T QP I KRGT I NPMV I SF KYGHAYLT QT FHG VNFKEL SKG ARAAT DL FP I V GQ L R I VRTWAGMEAMT AD HL P I I GPS PNASG VF VNFKEL SKG ARAAT DL FP I V GQ L R I VRTWAGMEAMT AD HL P I I GPS PNASG VF VNFKEL SKG ARAAT DL FP I V GQ L R I VRTWAGMEAMT PD HL P I I GA SANAKG I F VD VVAL AASVQAAV RVF PEL QNVQ I V RTWAGMEAMT PD HL P I I GA SANAKG I F VD VVAL AASVQAAV RVF PEL QNVQ I V RTWAGMEAMT PD HL P I I GA SANAKG I F QA RH I A NRA I NVMP DV G AL RVI RQWSG F RVL PP DGT PI YAQ SE EY PG LF AG ARH I A NRA I NVMP DV G AL RVI RQWSG I RVL T PD GV PVYDRS RRY PG I H EG AARI T RKALKI L PD L AQ AKL V RQWSCL RI MT PD GN PVYARSL T HPG AE DI S AVLA SRALR I F PAL SD L NVVR SWTG F RVKT ADG VP Y DRS RRY PG I H EG AARI T RKALKI L PD L AQ AKL V RQWSCL RI MT PD GN PVYARSL T HPG AE DI S AVLA SRALR I F PAL SD L NVVR SWTG F RVKT ADG VP Y DRS RRY PG I H EG AARI T RKALKI L PD L AQ AKL V RQWSCL RI MT PD GN PVYARSL T HPG AE DI S AVLA SRALR I F PAL SD L NVVR SWTG F RVKT ADG VP Y DRS SR KKGL Y E F L RE VS YY F T K I I PAL KNLL I L RTWAGY YAKT PD SNPA I G R I EEL NDYY ARSL T T GSL D PRL - A PF AA D RF AKDT T RAKEAL R - TDL I V H GG T NRVI - EPF SAK SL I RKEVR HVAH TDL I V H G E T NRVI - EPF SAK SL I RKEVR HVAH TDL I V H G E T NRVI - EPF SAK SL I RKEVR HVAH	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L L I G G S Q G R 263 T VM I G A T K E N V 267 T F Q L G V T N E R V 269 T VM I G D S K E D E V 272 G I I G G S Q G R 263 T VM I G D S K E D E 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 339 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 386 379 372
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacH EU6_SacH Bj_OpDH_B Ach5_0ox8/ Bj_OpDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH	203       VVNCAGAWA         98       VVNAAGAWA         95       VNAAGAWA         95       VNAAGAWA         97       VIAAGLONE         97       VIAAGLONE         97       VNAAGAWA         97       VNAAGLONE         97       VNAAGAWA         97       VNAAGLONE         97       VNAAGAWA         97       VNAAGASA <t< td=""><td>SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP -VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP -LLPIPGSG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP -LPIPLSAG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERTMP -FLPHPLTT</td><td>S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S K E D E 299 T VM I G D S K E D E 299 T VL I G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 338 I A T C H S - G V T 336 I A T C H S - G V T 336 I A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 366 379 373</td></t<>	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP -VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP -LLPIPGSG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP -LPIPLSAG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERTMP -FLPHPLTT	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S K E D E 299 T VM I G D S K E D E 299 T VL I G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 338 I A T C H S - G V T 336 I A T C H S - G V T 336 I A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 366 379 373
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00xB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00xB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacF	203       VVNCAGAWG         198       VVNAAGAWA         198       VVNAAGAWA         198       VVNAAGAWA         198       VVNAAGAWA         198       VVNAAGAWA         195       VVNAAGAWA         195       VVNAAGAWA         195       VVNAAGAWA         107       VIAAGAMA         107       VIAAGLOCK         101       VAAGLOCK         207       VLAAGLOCK         201       VVAAGLOCK         201       VLAAGLOCK         202       PCL       DDTSF         203       VVNATNAWA         204       ADCOSA         205       ADLDACRSF         206       ADCOSA         206       F.OLATTE         206       P.QODRNETVIA         2070       D.D.RTLKB         206       P.QODRNETVIA         2070       D.D.RTLKB         204       VVNATRINA </td <td>SK IAAS VD EE L P VNVMAL SMMV TT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMIV TERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAG AALGPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAG GKLAPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAG GKLAPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG</td> <td>S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L L I G G S Q G R 263 T V M I G A T K E NV 267 T F Q L G V T N E RV 269 T V M I G D S E T D 263 T V M I G D S K E D E 269 T V I I G G H L AT 265 G I I G G I G Y E I G 269 V V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 338 H V F G F S G H G F Q 338 I A G F S G H G F M 339 V A C C H S - G V T 336 I A A G F S G H G F M 339 366 379 372 373 370</td>	SK IAAS VD EE L P VNVMAL SMMV TT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMIV TERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAG AALGPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAG GKLAPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAG GKLAPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L L I G G S Q G R 263 T V M I G A T K E NV 267 T F Q L G V T N E RV 269 T V M I G D S E T D 263 T V M I G D S K E D E 269 T V I I G G H L AT 265 G I I G G I G Y E I G 269 V V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 338 H V F G F S G H G F Q 338 I A G F S G H G F M 339 V A C C H S - G V T 336 I A A G F S G H G F M 339 366 379 372 373 370
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacH EU6_SacH Bj_OpDH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH	203       V V N CAG AWG         208       V V N AAG AWA         98       V V N AAG AWA         99       V V N AAG AWA         205       V I V AAG AWA         205       V I V AAG AWA         207       V I AAG AG AWA         207       V I AAG L D H         207       V V AAG AWA         207       P C L D D D T S F         206       A D L D AQ R S F         207       G F - D L AT T T         206       P Q D R N E T V         207       G F - D L AT T T         206       P Q D R N E T V         207       P T Y D L T P T Y E         206       P Q Q R N E T V         207       P T Y D L T P T Y E         2045       L SP AVG E T V         207       P T Y D L T P T Y E         2040       L V P V L G A IM         2037       L Y P S Q R V I	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQSAVG VKIAAMVG - DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG - DPLQHAIRTSMMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG - FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG MAFAKSLD - LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERIEP - LLPIPGSG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERIAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERIAP - VFPYAASG IRQTMTG AALGPMVG - LNVPIRPQRGQLVTERTMP - FLPHPLTT	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S K E D E 269 T VM I G D S K E D E 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G G H L A T 265 G I I G G I G Y E I G 269 T V G F S G H G F Q 339 H S F G Y S G H G F Q 339 H S F G Y S G H G F Q 339 H S F G Y S G H G F Q 339 I S F G Y S G H G F Q 349 I S F G Y S G H G F Q 349 I S F G Y S G H G F Q 349 I S
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OOXB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacH Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacE C58_NoxB	203       VVNCAGAWG         198       VVNAAGAWA         198       VVNAAGAWA         195       VVNAAGAWA         195       VVNAAGAWA         195       VVNAAGAWA         195       VVNAAGAWA         195       VVNAAGAWA         195       VVNAAGAWA         107       VIAAGNA         101       VIAAGLOCA         101       VVAACA         102       VVAACA         103       VVAACA         104       VVAACA         105       F         106       F         107       PTYDLTPTY         108       F         104       VYPVLGAIM         105       LYPSVGAIM	SK IAAS VD EE L P VNVMAL SMMVTT RVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQ SSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSQ MAFAKSLD LPLLLVPEKGQLLITERVAP -VFPYAASG LRQTMTG AALGPMVG LNVPIRPQRGQLLVTERIAP - LLPIPGSG LRQTAG GKLAPMVG LSCPLKRSKGQILVTERTAP - VFPYAASG IRQTMTG GKLAAALG EPVPVETVAPMLMITSRVPH - FLPHPLTT IRQTDEG GKIAAALG EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITQPIKRGTINPWVISFKYGHAYLTQTFHG VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF GARHIA NRAINVPPVEEQUVIVERTWAGMEAMTADHLPIIGASANAKGIF GARHIA NRAINVPDVGALRVIRQWSGFRVLPPDGTPIYAQSEEYPGJF AGARHIA NRAINVPDVGALRVIRQWAGLRVLTPDGVPYARSLTHPGAF JSAVKA SRALRIFPALSDLNVVRSWTGFRKKTADGVPYARSLTHPGAF GARHIA NRAINVPDVGALRVIRQWAGIEAKMKDDIPVFARSLTHPGAF GARHIA NRAINVPDVGALRVIRQWAGIRVKTADGVPYARSLTHPGAF GARHIA NRAINVPDVGALRVIRQWAGIRVKTADGVPYARSLTHPGAF GARHIA NRAINVPDVGALRVIRQWAGIRVKTADGVPYARSLTHPGAF JSAVKA SRALRIFPALSDLNVVRSWTGFRVKTADGVPYARSLTHPGAF LDWKSLAISARTVFELFPVMRSATIVRAWAGIEAKMKDDIPVFGPSSRHKGLY EFLREVS YYFTKIIPALKNLLIL RTWAGYAKTPDSNPAIGRIEELNDY ARSLTTGSLDPRL - APFAADSFAKDTTRAKEALR - TDLIVHQGTNNVI - EPFSAKRLIRE	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G L Q G R 266 T L U I G G G V M G R 269 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S E T D 263 T VM I G D S E T D 263 T VM I G D S K E D E 269 T VM I G G K L A T 265 G I I G G I G Y E I G 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 T VL G G F S G H G F Q 338 H V F G F S G H G F Q 338 H V F G F S G H G F Q 338 V A C H S - G V T 333 V A C H S - G V T 333 V A C H S - G V T 336 H Q F G F S C H G F Q 338 I A A G F S C H G F M 339 386 379 372 373 373 379 372 373
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE Chry5_lecH Bo542_LecH EU6_SacE Chry5_lecH Bo542_LecH EU6_SacE Chry5_lecH Bo542_LecH EU6_SacE Chry5_lecH Bo542_LecH EU6_SacE Chry5_lecH Bo542_LecH EU6_SacE Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Po_ODH_B	203       V V N C AG AWG S         198       V V N AAG AWA N         98       V V N AAG AWA N         99       V V N AAG AWA N         99       V V N AAG AWA N         205       V I V AAG AWA N         207       V I AAG NA NG LOHN         207       V L AAG LOHN         207       L D DDT SF         207       AL AA DROSA         208       L - D RG VT VA         209       G R - ST DV TAA         200       G R - ST DV TAA         201       D - D RT ST DV TAA         202       G R - ST DV TAA         203       G R - ST DV TAA         204       N - RISS P L         207       P T Y D L T P T Y         2045       L SP AVG E T VA         207       P T Y D L T P T Y	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG LRQTAG ALGPMVG LNVPLVPEKGQLLVTERIEP - LLPYPASG LRQTAG GKLAPMVG LNVPLVPEKGQLLVTERIAP - LLPIPGSG IROTMTG AALGPMVG LNVPLVPEKGQLLVTERIAP - LLPIPGSG IROTMTG GKLAAALG EPVVETVAPMGQILVTERTAP -FLPHPLTT IROTMTG GKIAAALG EPVVETVAPMLMITSRVPH -FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG VVNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF SVNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTSDHRVVIGLSKVVGGI SATKLA RATRIPALAPARLVRQWSGFRVLPPDGTPIYAQSEYPGJF GAARIT RKALKILPDLAQAKLVRQWSCLRIMTPDNVJSFKYGHYGIH EGAARIT RKALKILPDLAQAKLVRQWSCLRIMTPDGNVYARSLTHPGAE JISAVKA SRALRIFPALSDLNVVRSWTGFRVKTADGVPIYDSSTHFGAF SISAVKA SRALRIFPALSDLNVVRSWTGFRVKTADGVPIYDGSSTHFGAF LDWKSLAISARTVFELFPVWRSATIVRAWAGIEAKKDDIPVFGSSTHKGLY EFLREVS YYFTKIIPALKNLLILRTWAGWSGIRVKTADGVPIYDHSSTHFGAF LDWKSLAISARTVFELFPVMSSATIVARWAGIEAKKDDIPVFGSSTHFGAF SISAVMT DRAQRMFPHLSRLNVIRSWTGFRVKTADGVPIYDRSRYPGAF LDWKSLAISARTVFELFPVMSSATIVARWAGIEAKKDDIPVFGSSTHFGAF LDWKSLAISARTVFELFPVMSSATIVARAWAGIEAKKDDIPVFGSSTHFGAF ADLICHGQTNSNL - APFAADRFAKDTRAKEALR - TDLIVHGETNRVI - EPFSAKRLIRE	S L V I G G G I S G K 271 T L V I G G G L OG R 266 T L V I G G G L OG R 266 T L U I G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G V T Q E V 272 G I M I G D S K E D E 269 T V I I G G G H L A T 265 G I I G G I G Y E I G 269 T V L I G G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 333 I A T C H S - G V T 336 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 S A A C H S - G V T 336 386 370 372 373 369 372
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B2 LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacH EU6_SacH Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B	203       V V N CAGAWA         203       V V N AAGAWA         208       V V N AAGAWA         205       V I V AAGAWA         207       V I AAG L D NE         201       I V L AAG L D NE         207       V AAGAWA         207       P C L D D D T S F I         206       P Q D A Q R S F I         206       P Y Q D R NE T V I         206       P Y Q D R NE T V I         207       L O - D R T L K HS         206       P Y Q D R NE T V I         207       L S P AVG E T V I         204       L Y P V L G A IM         207       L S P AVG E T V I         208       L S P AVG E T V I <t< td=""><td>SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQ SAVG VKIAAMVG DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQ SDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP - LLPIPGSG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP - LLPIPGSG</td><td>S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G V G R 263 T V M I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S K E D E 269 T VM I G D S K E D E 269 T V I G G H L A T 265 G I I G G I G Y E I G 269 T V I G G H L A T 265 G I I G G I G Y E I G 269 T V G F S G H G F Q 334 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 334 H S F G Y S G H G F Q 338 I A G H S - G V T 336 V A C H S - G V T 336 U A C C H S - G V T 336 I A G F S G H G F Q 338 I A A G F S G H G F M 339 366 379 372 373 372</td></t<>	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVVIG IDRPLS - FKQ SAVG VKIAAMVG DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQ SDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLKPVVSVSFGKKLS - LKQTSQG MAFAKSLD LPLLLVPEKGQLLITERVAP - VFPYAASG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP - LLPIPGSG IRQTMTG AALGPMVG LNVPIRPQRGQLVTERIAP - LLPIPGSG	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L U I G G G V G R 263 T V M I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S K E D E 269 T VM I G D S K E D E 269 T V I G G H L A T 265 G I I G G I G Y E I G 269 T V I G G H L A T 265 G I I G G I G Y E I G 269 T V G F S G H G F Q 334 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 334 H S F G Y S G H G F Q 338 I A G H S - G V T 336 V A C H S - G V T 336 U A C C H S - G V T 336 I A G F S G H G F Q 338 I A A G F S G H G F M 339 366 379 372 373 372
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/	203       V V N CAG AWG         198       V V N AAG AWA N         98       V V N AAG AWA N         99       V V N AAG AWA N         99       V V N AAG AWA N         99       V V N AAG AWA N         905       V I V AAG AWA N         907       V I AAG NA NG L         101       I V I AAG L G L         101       I V AAG L AG NANC         102       P C L D D D T S F         103       G R - D L A C R S F         104       L A A D RO S A         105       G R - S T D V T A         104       L O P T Y D L T P T Y         104       L Y P V L G A IM         104       L Y P V L G A IM         1034       L AAA H A A I V         104 </td <td>SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP -VFPYAASG LRQTMTG AALGPMVG LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG GKLAPMVG LSCPLKRSKGQILVTERLAP - LLPIPGSG IRQTMTG GKIAAALG EPVPVETVAPMGPTRGQIVVTERTMP - FLPHPLTT IRQTDEG GKIAAALG EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG VVNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF AGARHIA LRATRIPELQNVQIVRTWAGMEAMTSDHRPVIGLSKYCGLI SATKLA LRATRIPELQNVQIVRTWAGMEAMTSDHRPVIGLSKYCGLI GARHIA SRALRIFPALANISVVRSWTGLRVHTPDGVPYJSPSARHPGLW VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTSDHRPVIGLSKYCGLI GARHIA SRALRIFPALAPARLVPQWSCLRIMTPDGNVYARSLTHPGAE DISAVLA SRALRIFPALSDLNVVRSWTGFRVKTADGVPYJDSSRRYPGIH EGAARIT RKALKILPDLAQAKLVPQWSCLRIMTPDGNVYARSLTHPGAE DISAVLA SRALRIFPALSDLNVVRSWTGFRVKTADGVPIYDGSTHKGAF SISAVMT DRAQRMFPHLSRLNVIRSWAGIEAKMKDDIPVFGPSSRHKGLY EFLREVS YYFTKIIPALKNLLILRTWAGYYAKTPDSNPAIGRIEELNDYY ARSLTTGSLDFRL -APFAADRFAKDTTRAKEALR - TDLIVHGTNRVI - EPFSAKRLIRE</td> <td>S L V I G G G I S G K 271 T L V I G G L Q G R 266 T L V I G G L Q G R 266 T L U I G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G V T Q E V 272 G I M I G D S K E D E 269 T VM I G D S K E D E 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S C H G F Q 333 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 3866 379 372 373 369 372 375 366</td>	SK IAAS VD EEL PVNVMAL SMMVTT RVAP - FVKPVVIG IDRPLS - FKQSAVG VKIAAMVG DDIQIGHKASMIVTERIAP - LLKPVVSVVGRPLS - FKQSDQG DHVSALAG DPLQHAIRTSMVVTERSVP - KISPVVSSFGKLS - LKQTSQG PSLTDPLG FKIPLKADRGQLLVTERIAP - LLPYPASG LRQTSDG MAFAKSLD LPLLLVPEKGQLLITERVAP -VFPYAASG LRQTMTG AALGPMVG LNVPIRPQRGQLVTERLAP - LLPIPGSG IRQTMTG GKLAPMVG LSCPLKRSKGQILVTERLAP - LLPIPGSG IRQTMTG GKIAAALG EPVPVETVAPMGPTRGQIVVTERTMP - FLPHPLTT IRQTDEG GKIAAALG EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKQFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG VVNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF AGARHIA LRATRIPELQNVQIVRTWAGMEAMTSDHRPVIGLSKYCGLI SATKLA LRATRIPELQNVQIVRTWAGMEAMTSDHRPVIGLSKYCGLI GARHIA SRALRIFPALANISVVRSWTGLRVHTPDGVPYJSPSARHPGLW VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTSDHRPVIGLSKYCGLI GARHIA SRALRIFPALAPARLVPQWSCLRIMTPDGNVYARSLTHPGAE DISAVLA SRALRIFPALSDLNVVRSWTGFRVKTADGVPYJDSSRRYPGIH EGAARIT RKALKILPDLAQAKLVPQWSCLRIMTPDGNVYARSLTHPGAE DISAVLA SRALRIFPALSDLNVVRSWTGFRVKTADGVPIYDGSTHKGAF SISAVMT DRAQRMFPHLSRLNVIRSWAGIEAKMKDDIPVFGPSSRHKGLY EFLREVS YYFTKIIPALKNLLILRTWAGYYAKTPDSNPAIGRIEELNDYY ARSLTTGSLDFRL -APFAADRFAKDTTRAKEALR - TDLIVHGTNRVI - EPFSAKRLIRE	S L V I G G G I S G K 271 T L V I G G L Q G R 266 T L V I G G L Q G R 266 T L U I G G S Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G V T Q E V 272 G I M I G D S K E D E 269 T VM I G D S K E D E 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S C H G F Q 333 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 3866 379 372 373 369 372 375 366
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacH EU6_SacH EU6_SacH EU6_SacH EU6_SacH Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1 BJ_OpnDH_B1 BJ_OPDH_B3 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_OoxB/ Bj_OpnDH_B1	203       V V N A G A WA S         203       V V N A G A WA S         208       V V N A G A WA S         208       V V N A G A WA S         208       V V N A G A WA S         208       V V N A G A WA S         205       V I V A A G A WA S         205       V I V A A G A WA S         205       V I V A A G A WA S         205       V I V A A G A WA S         207       V I A A G L D N F         207       V A A G A WA S         207       V A A G A WA S         207       V N A A G A WA S         207       V N A A G A WA S         207       V N A A G A WA S         207       V N A A G A WA S         207       V N A A G A WA S         207       V N A A G A WA S         207       P C L D D D T S F         208       I - D R O L A T T T         207       G F - D L A T T T         208       F - D R T L K HS         209       P Y D D T T T T         201       P T Y D L T P T Y E         202       P S V G R V I A         203       L A A H A A I V A         203       L A A H A A I A A         204       L A A H A A I A A	SK I AAS VD EEL PVNVMAL SMMVTT RVAP - FVK PVV I G I DR PLS - FKOS AVG VK I AAMVG - DDIQI GHKASMMI VTER I AP - LLK PVVS VVG RPLS - FKOS D OG DHVSALAG - DPLQHAI RTSMMVVTER I AP - LLK PVVS VVG RPLS - FKOS D OG PSLT DPLG - FK I PLKADRGQLLVTER I EP - LL PVVS SFGKKLS - LKOTS OG MAFAKSLD - LPLLLVPEKGQLL I TERVAP - VFPYASG	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L V I G G G L Q G R 266 T L V I G G G L Q G R 263 T VM I G A T K E NV 267 T F Q L G V T N E RV 269 T VM I G D S E T D 263 T VM I G D S E T D 263 T VM I G D S E T D 263 T VM I G D S E T D 263 T V L I G G H L A T 265 G I I G G I G Y E I G 269 T VL I G G H L A T 265 G I I G G I G Y E I G 269 H V F G F C G H G F Q 344 H S F G Y S G H G F Q 339 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 338 I A G H S - G V T 336 I A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 366 379 372 373 369 372 375 366
LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00x8/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00x8/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH Bo542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_00x8/ Bj_OpnDH_B1 EU6_SacE C58_NoxB Pp_ODH_B Ach5_00x8/ Bj_OpnDH_B1 EU6_SacE C58_NoxB Pp_ODH_B Ach5_00x8/ Bj_OpnDH_B1 EU6_SacB C58_NoxB Pp_ODH_B Ach5_00x8/ Bj_OpnDH_B1 B1 B1 B1 B1 B1 B1 B1 B1 B1	203       VVNCAGAWAS         198       VVNAAGAWAS         195       VVNAAGAWAS         197       VIAAGAMAS         101       VLAAGLOCK         101       VLAAGLOCK         101       VLAAGLOCK         107       VLAAGLOCK         108       IADCOK         1097       VVNATNAWANS         1007       VLAAGLOCK         1008       IADCOK         1010       IADCOK         1020       INT         1020       INT         1020       INT         1020       INT         1020	SKIAASVD - EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKOSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVVGRPLS - FKOSDQG DVSALAG - DPLQHAIRTSMMVVTERIAP - LLKPVVSVVGRPLS - FKOSDQG PSLTDPLG - FKIPLKADRGQLLVTERIAP - LLKPVVSVGRPLS - FKOSDQG RSLTDPLG - FKIPLKADRGQLLVTERIEP - LLPYPASG LRQTSDG AAFAKSLD - LPLLVPEKGQLLVTERIEP - LLPYPASG LRQTMTG AALGPMVG - LNVPIRPQRGQLVTERLAP - VFPYAASG RQTAEG GTLAPMVG - LNPPLSCQLVTERIAP - FLPHPLTT IRQTDEG QTLAPMVG - LHAPMGPTRGQIVVTERTMP - FLPHPLTT IRQTDEG GKIAAALG - EPVPVETVAPMLMITSRVPH - FIDPVVILRGRKLS - FKOFSNG NLINAMAGIKTKIPIEPYKHQAVITOPIKRGTINPMVISFKYGHAYLTQTFHG VVFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTSDHRPVIGLSKKVEGI GARHIARATRIIPALAPARLVROWSGFRVLPPDGTPIYAQSEEYPGLF GARHIARATRIIPALAPARLVROWSGFRVLPPDGTPIYAQSEEYPGLF GARRHARATRIIPALSDLNVIERTMQGIRVMPQDGFPIYDSSRRYPGIH EGAARITRALKILPLAQAKLVROWSGFRVLPPDGTPIYAQSEEYPGLF ASATKLARATRIIPALSDLNVIRSWTGFRVKTADGVPYARSLTHPGAE DISAVLASRALKIPPLSRLNVIRSWTGFRVKTADGVPIYDSSRRYPGIH EGAARITRALKILPLSQKLIVRTWAGMEAKTDDIPVYARSLTHPGAE DISAVLASRALKIPPLSRLNVIRSWTGFRVKTADGVPIYDSSRRYPGIH EGAARITRALKILPLSQKLIVRTWAGMEAKTDDIPVGSFRVEPDGFPIYDOSETHPGAF IJSAVMTDRAQRMFPHLSRLNVIRSWTGFRVKTADGVPIYDSSRRYPGIH ASATKLARATRIIPALSDLNVVRSWTGFRVKTADGVFIYDSSRRYPGIH EFLREVSYYFTKIIPALSKLIILTRVAGWSGIRVMPQDGFPIYDOSETHPGAF ADUSLAISARTVFELFPVMRSATIVRAWAGIRAARDDIPVGSSRHKGLY EFLREVSYYFTKIIPALSRLNVISSWTGFRVKTADGVFINDSSRAIGRIELNDYY ARSLTGSLDPRL - APFAADRFAKDTTRAKEALR - TDLIVHGTNSNL DAFCPGRVSIERAAA AAMAGGREPSY DAFRGSSFKMGIAA ACMAAGGREPSY DAFRGSSFKMGIAA ACMAAGGREPSY DAFRGSSFKMGIAA ARMVAQGALEPELVGAFSASRFGAAAANNSGY ARMVAQGALEPELVGAFSASRFGAAAANNSGY ARMVAQGALEPELVGAFSASRFGAAAANNSGY ARMVAQGALEPELVGAFSASRFGAAAANNSGY ARMVAQGALEPENSY AN FFGAASASAGAANNSGY ARTWAQGALEPENSY AN FFGAASASAGAANNSGY ARTWAQGALEPENSY AN FFGAASASAGAANNSGY ARTWAQGALEPENSY AN FFGAASASAGAANNSGY ARTWAQGALEPENS	S L V I G G G I S G K 271 T L V I G G G L Q G R 266 T L U I G G S Q G R 263 T V M I G A T K E N V 267 T F Q L G V T N E R V 269 T V M I G A T K E N V 267 T F Q L G V T N E R V 269 T V M I G D S K E D E 299 T V I G D S K E D E 299 T V I I G G G H L A T 265 G I I G G I G Y E I G 269 T V I G S C G H G F Q 344 H S F G Y S G H G F Q 339 H V F G F S G H G F Q 339 H V F G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F Q 338 I A A G F S G H G F M 339 366 379 372 373 369 372 375 366 378
LBA9402_F3X89_28355 Chry5_lecH B0542_LecH EU6_SacH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH B0542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B2 Ph_LProDH1_B LBA9402_F3X89_28355 Chry5_lecH B0542_LecH EU6_SacE C58_NoxB Pp_ODH_B LBA9402_F3X89_28355 Chry5_lecH B0542_LecH EU6_SacE C58_NoxB Pp_ODH_B Ach5_0oxB/ Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1 Bj_OpnDH_B1	203       V V N C AG AWA S         203       V V N AAG AWA S         208       V V N AAG AWA S         208       V V N AAG AWA S         208       V V N AAG AWA S         205       V I V AAG AWA S         207       V I AAG NANC         201       I V L AAG L D H         207       V V AAG AWA S         207       P C L D D D T S F         206       A D L D AQ R S F         207       G R - S T D V T A         208       T - R S S S F         207       G R - S T D V T A         207       G R - S T D V T A         207       G R - S T D V T A         207       G R - S T D V T A         208       T - R S S S F         204       L P V V L G A IM         207       P T Y D L T P T Y         208       L AA A H A I V A         207	SKIAASVD - EEL PVNVMALSMMVTTRVAP - FVKPVVIGIDRPLS - FKOSAVG VKIAAMVG - DDIQIGHKASMMIVTERIAP - LLKPVVSVGRPLS - FKOSDOG DHVSALAG - DPLQHAIRTSMMVVTERIAP - LLKPVVSVGRPLS - FKOSDOG DHVSALAG - DPLQHAIRTSMMVVTERIAP - LKPVVSVGRPLS - FKOSDOG ASIDPLG - FKIPLKADRGQLLVTERIEP - LLPYPASG LROTSDG MAFAKSLD - LPLLLVPEKGQLLVTERIEP - LLPYPASG LROTSDG MAFAKSLD - LPLLLVPEKGQLLVTERIAP - VFPYAASG LROTMIG GKIAAALG - EPVPVETVAPMLVTERIAP - LLPIPGSG LROTMIG GKIAAALG - EPVPVETVAPMLMITSVVFERIMP - FLPHPLTT IROTDEG GKIAAALG - EPVPVETVAPMLMITSVVFERIMP - FLPHPLTT IROTDEG GKIAAALG - EPVPVETVAPMLMITSVVFERIMP - FLPHPLTT IROTDEG VV MARASSAAATVAAFPALANISVVRSWTGLEGLTPDGVPYISPSARHPGLW VNFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGPSPNASGVF VVFKELSKGARAATDLFPIVGQLRIVRTWAGMEAMTADHLPIIGASANAKGIF VDVVALAASVQAAVRVFPELQNVQIVRTWAGMEAMTSDHRPVIGLSKKVEGLI ASATKLALRATRIIPALAPARLVRQWSGFRVLPPDGTPIYAQSEEYPGAF SISAVMTDRAQRMFPHLSRLVRQWGCLRIMTPDGVPYJRSRSRTPGIF SISAVMTDRAQRMFPHLSRLVRQWSGFRVLPPDGTPIYAQSEEYPGAF SISAVMTDRAQRMFPHLSRLVRQWSGIRVMPQDGFPIYDHSERYPGAF SISAVAASRALRIFPALSDNVVRSWTGFRVKTADGVYJRSRTPGAF AGARITANRALNIPDVGALRVIRQWGGIRVMPQDGFPIYDHSERYPGAF SISAVATDRAQRMFPHLSRLNVIRSWAGIRVMPQDGFPIYDHSERYPGAF AGARITSKKALKILPAARLVRQWSCTRIMTPDGVPYARSLTHPGAE DISAVAASRALRIFPALSDNVVRSVTGFRVKTADGVYARSQTHPGAF AGARITSKKALKILPDLAQKLVRQWSCTRIMTPDSNPAIGRIEELNDYY ARSLTGSLDPRL - APFAADFAKDTRAKEALR - TDLIVHGGTNRVI - EPFSAKLIRKEVRHVAH ADLICHGQTNSNL DAFCPGRVSIERAAA APAMIEGRLSTDL SAFSNGRFDVQKCA ARWVAGGALEPELVGAFSASRFMCGIAA ARWVAGGALEPELVGAFSASRFDQAKVGIAA ARWVAGGALEPELVGAFSASRFGGASAANNSGY - ARMVAQGALEPELVGAFSASRFGGASAANNSGY - AELIVNGTTRTI - SDLGIDRFHPST AFUTYFKTTON - EPFSAKRFDVAGGFGGASAANNSGY - AELIVNGTTRTI - SDLGIDRFHPST	SLVIGGGISGK271           TLVIGGGLOGR266           TLVIGGGLOGR266           TLVIGGGLOGR263           TVMIGATKENV267           TFQLGVTNERV269           TVMIGDSKEDE209           TVLIGGCHLAT265           GIMIGDSKEDE209           TVLIGGFCGHGFQ339           HSFGYSGHGFQ339           HSFGYSGHGFQ339           HSFGYSGHGFQ339           HVFGFSGHGFQ339           HVFGFSGHGFQ339           HVFGFSGHGFQ339           HVFGFSGHGFQ339           HVFGFSGHGFQ339           HSFGYSGHGFQ339           HVFGFSGHGFQ339           HVFGFSGHGFQ338           HVFGFSGHGFQ338           HVFGFSGHGFQ338           HVFGFSGHGFQ338           HVFGFSGHGFQ338           HVFGFSGHGFQ338           HVFGFSGHGFQ338           IATCHS-GVT330           VACHS-GVT330           YACHS-GVSGHGFQ338           IAAGFSGHGFM333           386           372           373           369           372           373           372           372           373           374           SGFSGHGFN339

**FIGURE A8** Multiple sequence alignment of  $\beta$  subunits of various opine dehydrogenases. Indicated with orange triangles is the ADP binding motif Gly-X-Gly-X-Gly. The arginine and tryptophan residues (putatively) interacting with FMN (as based on *Pyrococcushorikoshi* (Ph) L-Proline dehydrogenase (Ph\_L-ProDH) crystal structure data) are indicated with blue triangles. Abbreviations of species are described in the legend of Appendix Figure A7.

	icrobiologyOpen	HOOYKAAS AND HOOYKAAS
	(Open Access)	
LBA9402 F3X89 28345	1 M F T <mark>P I G</mark> S - S E A T I <b>T</b> I F F <b>E</b> N	IE <mark>PIRARQGETVAACLLR</mark> AGVRYF <b>R</b> TTPV <b>SG</b> S 49
Chrv5 LecF		EAVAAEANESVAAVLLRQDPVWSRTTPVSET 50
Bo542 LecF		EAVTAETNESVAAVLLRQDPVWSRTTPVSEA 50
EU6 SacF	1 M LERPDK TADMTILVNG	VETIAWEGETVASVLLRVP-EAGRRSPVSGK 47
C58 NoxC	1 M T S E I L F P S G Q A V E I T V D G	RLVTARDGEPLAAVFLRLDDIHTRTSFPSAQ 50
Pp OpnDH C	1 MDVSIFK <mark>K</mark> LPELHAITITIYVDG	VPVMAEQ <mark>GETVAAVLL</mark> RQSTAVN <mark>RTSP</mark> VKGS 54
Ach5 OoxC	1 M L N <mark>R</mark> L YR - I E H P I Q F S F D G	TVIQAER <mark>GD</mark> MLATALLAAGVDRV <mark>R</mark> QSIVSGS 49
Bj_OpnDH_C	1 M F R <mark>R</mark> SEQDKR <mark>P</mark> QVQ I F V DG	VAVAARQ <mark>GDTVS</mark> AALLAS <mark>G</mark> QDAR <mark>R</mark> STAVSGA 50
	<b>• • •</b>	▼
LBA9402_F3X89_28345	50 K <mark>RLPYCMIGHCFDCLVEIDGIGS</mark>	RQA <mark>CLSQVQDGMK</mark> LRV <mark>Q</mark> SGAAAID <mark>G</mark> NSDQ 101
Chry5_LecF	51 K <mark>RAPYCMMGVCFECLAVVDG</mark> -AS	VQTCLVSVRDGMRIERQLGRRRIAL 97
Bo542_LecF	51 R <mark>R</mark> APY <mark>CMMGVCFEC</mark> LALVDG-AS	VQTCLVAVQDGMRIERQLGRRRIAR 97
EU6_SacF	48 PRLPYCQMGVCFECLAIVDGIAS	T <mark>QGCLVP</mark> VRH <mark>GMRIE</mark> SQR <mark>GPR</mark> EISR 95
C58_NoxC	51 PRAPYCMMGVCFECVVHVEDIGT	I <mark>R</mark> SCQQTVHAGMRVERHDHPKRLK 97
Pp_OpnDH_C	55 PRAPYCMMGVCFDCLAVVDGNAS	TQTCLVTVKDGMRVERQYGRRSVVS 102
Ach5_OoxC	50 PRAPYCLMGVCFECLVTVDGVQN	IRQACLTEVENGMTVLSQKGATTIPGLARSDE 103
Bj_OpnDH_C	51 P <mark>R</mark> LPY <mark>CMMGVCFDCLVTIDGVG</mark> N	IR <mark>QGCIVPVAEGMQIEIQ</mark> K <mark>G</mark> K <mark>R</mark> EI <mark>G</mark> R 98
LBA9402_F3X89_28345		
Chry5_LecF		
Bo542_LecF		
EU6_SacF		
C58_NoxC		
Pp_OpnDH_C		
Ach5_OoxC	104 GTAL	107
Bj_OpnDH_C		

FIGURE A9 Multiple sequence alignment of  $\Upsilon$  subunits of various opine dehydrogenases. Indicated with pink triangles is the [2Fe-2S] iron-sulfur cluster binding motif (Cys-X<sub>4</sub>-Cys-X<sub>2</sub>-Cys-X<sub>11-12</sub>-Cys). Abbreviations of species are described in the legend of Appendix Figure A7.