The Role of Critical Ion Pairs in the Evolution of a Novel Enzyme in Tomato

by

Allan Castillo

Department of Biology
Duke University

Date:_______________________

Approved:

___________________________
Mark D. Rausher, Supervisor

___________________________
Storrs T. Mitchell-Olds

___________________________
Paul S. Manos

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Biology in the Graduate School of Duke University

2018
ABSTRACT

The Role of Critical Ion Pairs in the Evolution of a Novel Enzyme in Tomato

by

Allan Castillo

Department of Biology
Duke University

Date:_______________________

Approved:

___________________________
Mark D. Rausher, Supervisor

___________________________
Storrs T. Mitchell-Olds

___________________________
Paul S. Mannos

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Biology in the Graduate School of Duke University

2018
Abstract

Although there has been much research has been done on how new genes arise from gene duplications, little is still known on how these genes arose and by what mechanisms they evolved their new function. Here I address a case study on how the evolution of a novel defensive gene in the Solanaceae family arose by examining the novel defensive copy of threonine deaminase (TD2) in tomato (*Solanum lycopersicum*). This gene evolved by gene duplication from an ancestor involved in biochemical gene synthesis to become an exceptionally stable anti-nutritive defensive gene in insect guts. This stability is hypothesized to have evolved due to three critical ion pairs, ionically bonded amino acids that have been shown to stabilize proteins. Here, I test whether these three critical ion pairs stabilize the protein and what effects it has the activity of TD2. I found the second critical ion pair reduces the $K_{cat}$ of the enzyme, indicating that the evolution of it increased activity. Further, the first and third critical ion pairs increase activity during the temperature and pH stability assays. The results suggest that the interactions and effects of each critical ion pair is complex but are likely to be stabilizing without sacrificing much activity and in some cases it may increase activity as well.
Contents

Abstract ......................................................................................................................................... iv

List of Tables ............................................................................................................................... vii

List of Figures ............................................................................................................................ viii

Acknowledgements ..................................................................................................................... ix

1. Introduction ............................................................................................................................... 1

2. Study system .............................................................................................................................. 5

3. Methods ...................................................................................................................................... 9
   3.1 Protein structure and phylogenetic trees ...................................................................... 9
   3.2 Vector construction ........................................................................................................ 10
   3.3 Site-directed mutagenesis .............................................................................................. 11
   3.4 Expression and purification .......................................................................................... 12
   3.5 Kinetic assays .................................................................................................................. 13
   3.6 Stability assays ................................................................................................................ 14
      3.6.1 Thermostability assay ............................................................................................... 14
      3.6.2 pH assay ...................................................................................................................... 15
      3.6.3 Protease assays ........................................................................................................... 15
      3.6.4 Activity assays ........................................................................................................... 16

4. Results ....................................................................................................................................... 17
   4.1 Kinetic results .................................................................................................................. 17
   4.2 Stability ............................................................................................................................ 23
List of Tables

Table 1: The three critical ion pairs in tomato TD2 ................................................................. 2

Table 2: Buffer systems for pH assay .......................................................................................... 15

Table 3: A.) One-way analysis of variance (ANOVA) on the $K_{cat}$ comparison of the four enzymes. B.) Tukey’s Honest Significance Test to identify which enzymes were significant. ........................................................................................................................................ 19

Table 4: One-way analysis of variance (ANOVA) on the $K_m$ comparison of the four enzymes .................................................................................................................................................. 21

Table 5: One-way analysis of variance (ANOVA) on the $K_{cat}/K_m$ comparison of the four enzymes .......................................................................................................................................................... 23

Table 6: Two-way ANOVAs of activity by enzyme and temperature. .................................... 25

Table 7: Two-way ANOVA on activity by enzyme and pH. .................................................. 30

Table 8: Two-way ANOVA on activity by enzyme and proteolytic cleavage. ...................... 33
List of Figures

Figure 1: Maximum likelihood gene tree of threonine deaminase in the Solanaceae clade. ......................................................................................................................................................................................................................................................................................................................... 4

Figure 2: Crystal structure of tomato TD2 (3IAU)................................................................................................................................. 7

Figure 3: A. Schematic of TD2 structure. B. Vector used in site-directed mutagenesis. ..10

Figure 4: Phylogeny of TD genes in Solanaceae showing when each of the three critical ion pairs arose by substitution. ................................................................................................................................................................................................. 11

Figure 5: Menten-Michaelis $K_{cat}$ parameter for the four enzymes tested..........................18

Figure 6: Menten-Michaelis $K_m$ parameter for the four enzymes tested. .........................20

Figure 7: Menten-Michaelis $K_{cat}/K_m$ ratio for the four enzymes tested............................22

Figure 8: Activity vs Temperature for each enzyme tested from 30 to 90 °C. .....................24

Figure 9: Activity vs. pH assay for all four enzymes across pH of 3 to 13.........................27

Figure 10: Protease resistance assay using chymotrypsin to cleave protein followed by an enzyme activity assay. ......................................................................................................................................................................................................................................................................................................................... 31
Acknowledgements

First and foremost, I would like to thank Katharine Korunes for inspiration and for making graduate school an amazing experience. Thanks to my family for their support on this endeavor that moved me over 2000 miles away from them. Thanks to my advisor Mark Rausher, it was a pleasure to work with and learn from such a gifted scientist. To my undergraduate research assistants, Yilin Yang and Jenny Bai, who did countless assays and helped generate a lot of the data used in this thesis. Thanks to my lab-mates, Irene, Joanna, Rong-Chien, Kimmy, Kate, and Peng for fun times and great conversations about science and life. Thanks to the BioGrads, BioCoRE, and SACNAS for their support throughout my graduate studies.
1. Introduction

Understanding the forces and mechanisms that drive evolutionary novelty is a major goal of evolutionary biology. Ernst Mayr (1963) described evolutionary novelty “as any newly acquired structure or property that permits the performance of a new function”. Despite extensive study on the origins of evolutionary innovation, much remains unknown about the mechanisms and processes of novel adaptations (Zhang 2003; Lynch and Katju 2004; Kaessmann 2010). The advent of genomic sequencing has revealed that novel genes are abundant in all eukaryotic genomes and are a major source of variation (Lynch and Conery 2000) and by far most novel genes arise through the process of gene duplication (Zhang 2003; Kaessmann 2010; Tautz and Domazet-Lošo 2011; Kondrashov 2012). Thus one of the leading questions in the field is by what mechanisms do novel genes evolve from following a gene duplication event?

Many processes are suspected to affect how a novel gene may arise in a gene duplicate. Epistasis is thought to play a major role in the evolution of genes, particularly the role of intramolecular epistasis which can manifest in the form of historical contingency (DePristo et al. 2005; Harms and Thornton 2013; Starr and Thornton 2016). In fact, some researchers go as far as stating that “epistasis is the primary factor in molecular evolution” (Breen et al. 2012). While much of the field of molecular evolution and protein evolution has focused on epistasis as the main factor affecting novel...
evolutionary trajectories, other mechanisms may play a large role in affecting evolution of novel genes. Those include codon composition, regulatory changes, and active site amino acid changes amongst others. One mechanism that has for the most part been overlooked is that of ionic pairing by amino acids. Ion pairs are ionically bonded amino acids that can be used by proteins to become more structurally sound. This study will look more closely into the role that ion pairs play in the evolution of a novel enzyme.

### Table 1: The three critical ion pairs in tomato TD2

<table>
<thead>
<tr>
<th>(+) Amino Acid</th>
<th>(-) Amino Acid</th>
<th>Abbr./Position¹</th>
<th>Mutation²</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glutamic Acid (E)</td>
<td>Arginine (R)</td>
<td>E116 – R133</td>
<td>E116 -&gt; Q116</td>
<td>Mut1</td>
</tr>
<tr>
<td>Aspartic Acid (D)</td>
<td>Lysine (K)</td>
<td>D100 – K245</td>
<td>K245 -&gt; I245</td>
<td>Mut2</td>
</tr>
<tr>
<td>Glutamic Acid (E)</td>
<td>Lysine (K)</td>
<td>E93 – K335</td>
<td>K335 -&gt; L335</td>
<td>Mut3</td>
</tr>
</tbody>
</table>

¹Single letter amino acid abbreviation with amino acid position based on pet30aTD2 vector
²The mutation to disrupt the critical ion pair by site-directed mutagenesis.

Ion pairs, also known as salt bridges, occur when two amino acids of opposite charge interact with each other in an ionic fashion (Table 1). Studies have shown that this interaction improves the stability of proteins, particularly in extreme conditions. For example, in enzymes of thermophilic and psychrophilic organisms an increased number of ion pairs leads to higher stability in conditions that would normally denature homologous enzymes with a lower number of ion pairs (Bae and Phillips 2004). Further, critical ion pairs—ion pairs with a stronger interaction leading to a distance between the amino acids of ≤ 4Å—are thought to significantly improve the stability of proteins.
(Szilágyi and Závodszky 2000). Thus, the more ion pairs, and more importantly critical
tion pairs, that are found in an enzyme the more stable it becomes. This increase in
stability is thought to allow the enzyme to function in harsh environments.

In this study I examine how the addition of three critical ion pairs that evolved in
a copy of a duplicated enzyme affects function and stability. Threonine deaminase (TD)
is an enzyme responsible for the first step in catalyzing threonine to isoleucine, and in
the Solanaceae family it has duplicated several times leading to 4 copies in Solanum
lycopersicum (tomato) (Fig. 1). One copy evolved an anti-nutritive defensive function
against lepidopteran insect herbivores, a group of significant pests of tomato and other
important Solanaceous plants, by catabolizing threonine in the insect gut and depriving
it of this essential amino acid (Gonzales-Vigil et al. 2011). This defensive copy (TD2) has
been found to be much more stable than the ancestral functioning copy (TD1) of
threonine deaminase, likely due to its primary function of catalyzing threonine in the
harsh environment of the lepidopteran gut. Thus, the TD2 enzyme copy must have
evolved the ability to function in the foreign lepidopteran gut, including being able to
catalyze at high alkalinity (pH 8-10), varying temperatures, and resist digestive
proteases. Here, I examine the role critical ion pairs have in increasing the stability of
TD2 and assess if there are any tradeoffs between stability and activity in evolving the
three critical ion pairs.
Figure 1: Maximum likelihood gene tree of threonine deaminase in the Solanaceae clade.
2. Study system

Threonine deaminase (TD) is an enzyme involved in amino-acid metabolism that is found in plants and microorganisms but seems to be missing in animals. While most organisms that have the enzyme have only one copy of it, many species in the plant family Solanaceae have multiple copies of threonine deaminase. At least two duplications of TD have occurred early in the radiation of the Solanaceae, leading to up to 4 paralogous copies of the enzyme in some species (Rausher and Huang 2016) (Fig. 1). One branch retains the canonical housekeeping function, referred to as TD1, while the TD2 branch evolved a novel function so far characterized only in tomato (Solanum lycopersicum) and wild tobacco (Nicotiana attenuata). Two more copies (TD3 and TD4) exist in some Solanaceous species, but their functionality is yet to be determined. The housekeeping copy, TD1, catalyzes the first reaction in the conversion of threonine to isoleucine by removing an amine from threonine to create ammonia (NH₃) and α-ketobutyrate (α-KB), which is then followed by four additional enzymatic steps to convert α-KB into isoleucine.

Studies in tomato TD2, henceforth SITD2, show that SITD2 performs as an anti-nutritive enzyme against at least three lepidopteran larvae pest of the tomato plant (Chen et al. 2005, 2007; Gonzales-Vigil et al. 2011). While it still catalyzes threonine similarly to SITD1, it does so in the midgut of the caterpillar to deplete threonine, an essential amino acid for lepidopterans, before it is absorbed by the gut (Chen et al. 2007;
Gonzales-Vigil et al. 2011). Four key innovations evolved in tomato *SITD2* that allowed it to transition from a housekeeping enzyme to a defense against herbivores. First, tomato TD2 became much more stable, albeit at an apparent cost of catalytic efficiency, so it can withstand the harsh new environment of the insect gut. The enzyme has much higher heat resistance and is remarkably resistant to gut proteases, specifically chymotrypsins, the primary protease family in lepidopteran midguts (Srinivasan et al. 2006). This is in stark contrast to *SITD1*, which gets completely digested by chymotrypsin proteases (Gonzales-Vigil et al. 2011). Second, it has been optimized to function at a pH very near the alkaline gut of most lepidopteran herbivores of Solanaceous plants (Chen et al. 2007). Third, *SITD2* has evolved the ability to escape allostERIC regulation without compromising function. Canonically functioning TD’s are allostERICally inhibited by isoleucine; however, *SITD2* has its allostERIC region cleaved off by chymotrypsin-like proteases, giving it the ability to constitutively catabolize threonine in the insect gut even in the likely presence of isoleucine. Fourth, *SITD2* has evolved an alternate expression pattern to *SITD1*, exhibiting high expression during flower development and in response to herbivory through the jasmonic acid pathway (Samach et al. 1991, 1995; Chen et al. 2007).
Figure 2: Crystal structure of tomato TD2 (3IAU). Panels show the three pairs of interacting residues (critical ion pairs).

The three critical ion pairs are identified in the cartoon of the cleaved crystal structure of *SlTD2* (Fig. 2). Each critical ion pair is colored differently, with the two interacting amino acids colored the same. These structural features are thought to provide stability to the enzyme and may account for the increase thermostability of *SlTD2* (Gonzales-Vigil et al. 2011). A phylogeny of how the critical ion pairs arose during the evolutionary history of *SlTD2* is shown in Appendix A. Two of the six amino acids involved in creating the three critical ion pairs were already present before the gene duplication event leading to *SlTD1* and *SlTD2*. Therefore, four mutations during the evolution of *SlTD2* lead to creating the three critical ion pairs (Appendix A). The
primary purpose of the investigation reported here was to determine whether these mutations contribute to the enhanced stability of SITD2.
3. Methods

3.1 Protein structure and phylogenetic trees

The crystal structure of the processed form of SITD2 was obtained from the Protein Data Bank (3IAU). PyMol was used to read the pdb file and identify the critical ion pairs based on (Gonzales-Vigil et al. 2011). A cartoon representation was created in PyMol (Fig. 2). Sequences for all paralogs of TD from Solanaceae and orthologs from various outgroups (Fig. 1) are from the master’s thesis work of Jie Huang (2013) and from (Rausher and Huang 2016). The sequences were initially aligned in Mesquite (Maddison and Maddison 2018) and then manually reviewed and corrected. A maximum likelihood tree was created using Garli 2.0 (Zwickl 2006) with 500 bootstrap support using SumTrees version 4.0.0 of the DendroPy 4.0.3 Python package (Sukumaran and Holder 2010). The phylogenetic trees were produced in Figtree (Rambaut 2016) and annotated in Adobe Illustrator. The curated multiple sequence alignment and phylogenetic tree from above were used to infer the ancestral sequences in codeml from PAML 4.5 (Yang 2001, 2007) using the LG matrix model (Le and Gascuel 2008). This was used to infer where on the phylogenetic tree the substitutions arose to create the critical ion pairs.
3.2 Vector construction

A. Schematic of TD2 structure. Numbers refer to amino-acid positions, with position 1 corresponding to the first amino acid after the plastid targeting peptide (PTP). B. Vector used in site-directed mutagenesis. *SlTD2* gene ligated to pET30a(+) vector without the PTP domain of TD2.

A copy of the TD2 gene was obtained from Gregg Howe’s lab and cloned into the pET30(a)+ expression vector (EMD Millipore) with kanamycin resistance. The TD2 gene is a truncated version without the leading PTP sequence (plastid targeting peptide, yellow portion in Figure 3A) which is used by chaperone enzymes to transport *SITD2* intracellularly to the chloroplast before truncation (Samach et al. 1991). Since this region
is removed before the enzyme is biologically active, I chose to express the enzyme without the targeting peptide similar to other studies on threonine deaminase 2 (Chen et al. 2007; Kang et al. 2006) The SlTD2 insert was cut with the restriction enzyme NdeI at the 3’ end, and XhoI at the 5’ end and ligated to the pET30(a)+ vector cut with the same enzymes using NEB T4 Ligase and following the NEB standard protocol for T4 ligase (Fig. 3B).

### 3.3 Site-directed mutagenesis

![Phylogeny diagram](image)

**Figure 4: Phylogeny of TD genes in Solanaceae showing when each of the three critical ion pairs arose by substitution. Legend shows the amino acid abbreviations that will be disrupted by site-directed mutagenesis.**

Critical ion pairs were disrupted by converting one of the pair’s residues to the ancestral state amino acid, i.e. reverting the substitution that created the critical ion pair (Fig. 4). The residue reverted by site directed mutagenesis was chosen to be the residue that arose most recently in SlTD2’s evolutionary history based on ancestral sequence reconstruction in PAML (Appendix A-D). Site directed mutagenesis was done by
Agilent QuikChange II Site-Directed Mutagenesis Kit following the manufacturer’s protocol with NEB® 5-alpha Competent *E. coli* (High Efficiency) for transformation.

**3.4 Expression and purification**

Expression cells (NEB BL21(DE3) Competent *E. coli*) were transformed with the pET30(a)TD2 and with the mutagenized TD2 vectors, respectively, and stored in 25% glycerol at -80°C. Inoculations from these glycerol stocks were grown overnight on agar plates with 50 µg/mL kanamycin at 37°C. A single colony was picked to inoculate a starter culture of 6ml Luria broth with 50 µg/ml kanamycin antibiotic and grown at 30°C overnight with shaking at 225 RPM. A 500 µL sample of the overnight starter culture was used to inoculate 0.5 Liters of Luria broth with 50 µg/mL kanamycin. The culture was grown at 37°C to OD 600 = 0.8, checked through spectrophotometry. The temperature was adjusted to 28°C, Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to 0.5 mM to induce expression and incubated for 15 hours with shaking at 225 RPM. The culture was spun down in a centrifuge at 5000 g to collect the bacterial cells. The cells were resuspended in 250 µL binding buffer (20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, pH 7.4), lysozyme was added with phenylmethylsulfonyl fluoride (PMSF) and Protease Inhibitor Cocktail in DMSO (Sigma Aldrich #P8849-1ML) and incubated for 30 mins on ice. The resuspension was transferred to 15mL Polystyrene Falcon Tubes and placed in a Bioruptor sonicator on the high setting at 30 seconds on/ 30 seconds off intervals for 15 minutes. The lysate was centrifuged at 15,000 g for 10
minutes, then run through a 0.45 µm syringe filter to remove cell and other debris. The lysate was equilibrated with binding buffer (1:2 ratio) and added to a 1 mL HisTrap HP™ purification column. After washing with binding buffer per the manufacturer’s protocol, the purified protein was eluted using elution buffer (20 mM sodium phosphate, 0.5 M NaCl, 500 mM imidazole, 1 mM DTT, 5 mM glycerol, pH 7.4). The first 0.5 mL of eluate was discarded, and the next 2.5 mL of eluate was collected for downstream analysis. SDS-PAGE gels were run to determine purity, and Bradford assay (BioRad #5000201) was run for quantification (Bradford 1976). A buffer exchange into PD-10 buffer, which is elution buffer lacking imidazole is done using Amersham PD-10 columns.

### 3.5 Kinetic assays

Previous studies of closely related threonine deaminase orthologs have shown that they follow Michaelis-Menten laws in the absence of allosteric inhibitors (Sharma and Mazumder 1970; Wessel et al. 2000). The velocity is calculated spectrophotometrically by measuring the amount of the product, α-ketobutyrate, produced. A modified version of the Sharma and Mazumder (1970) and Kang et al. (2006) procedures was used to calculate amount of product produced over time. Standard assays were run in a 200 µL mixture containing 100 mM Tris-HCl buffer pH 9, 100 mM KCl, varying amounts L-threonine, 2 ng of enzyme, and incubated for 10 minutes at 30°C. The amounts of L-threonine were 200, 125, 100, 75, 50, 25, 12.5, and 2.5
mM. The reaction was stopped using 160 µL of cold (4°C) 7.5% trichloroacetic acid and protein precipitates were removed by centrifugation at 10,000g for 2 min. The supernatant was transferred to a new tube and 400 µL of 0.05% of 2-4-dinitrophenylhydrazine in 1 N HCl was added to the reaction and incubated for 10 minutes. 10 mL of 0.4 N NaOH is added, mixed well, and incubated for 20 minutes. The reaction read in a Beckman DU600 spectrophotometer at 505 nm. A non-linear regression was fitted the data (independent: substrate concentration [mM], dependent: velocity [S/min]) using the package DRM in R (Ritz et al. 2015). A one-way ANOVA was performed in python to determine significance with enzyme as a fixed factor. If the ANOVA was significant, a Tukey’s Honest Significance Test was performed to identify which of the enzymes were significant (Appendix F).

3.6 Stability assays

3.6.1 Thermostability assay

This assay incubated the enzyme at a constant temperature and then used an activity assay (See 3.6.1 below) to assess the amount of product produced, where higher product indicates more activity. The enzyme was incubated for 20 minutes at temperatures ranging from 30°C – 90°C.
Table 2: Buffer systems for pH assay

<table>
<thead>
<tr>
<th>pH</th>
<th>buffer system</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>citric acid</td>
</tr>
<tr>
<td>4</td>
<td>citric acid</td>
</tr>
<tr>
<td>5</td>
<td>citric acid</td>
</tr>
<tr>
<td>6</td>
<td>sodium phosphate</td>
</tr>
<tr>
<td>7</td>
<td>sodium phosphate</td>
</tr>
<tr>
<td>8</td>
<td>sodium phosphate</td>
</tr>
<tr>
<td>9</td>
<td>borax-HCl</td>
</tr>
<tr>
<td>10</td>
<td>sodium bicarbonate</td>
</tr>
<tr>
<td>11</td>
<td>sodium bicarbonate</td>
</tr>
<tr>
<td>12</td>
<td>KCl-NaOH</td>
</tr>
<tr>
<td>13</td>
<td>KCl-NaOH</td>
</tr>
</tbody>
</table>

3.6.2 pH assay

The pH assay assesses the product produced by the enzyme at a given pH. Unlike the thermostability assay, the pH assay adds enzyme directly to buffer containing the desired pH. The assays were performed at 30°C with the buffer systems in Table 2.

3.6.3 Protease assays

The protease assay is intended to test the effects that chymotrypsin has on the stability on the enzyme. The enzyme is incubated with α-Chymotrypsin from bovine pancreas (Sigma-Aldrich C4129) at a 1:1 molar ration for 2 hours. It was then followed with an activity assay.
3.6.4 Activity assays

The activity assay was performed under standard conditions: pH 9 and 30°C. 190 µL of reaction buffer (100 mM Tris-HCl, pH 9, 100 mM KCl, 250 mM L-Threonine) was quickly mixed with 10 µL of 20µg/mL enzyme and placed in the water bath for 5 minutes. The reaction was stopped and product measured as in the Kinetic assays. All statistical analysis for the stability assays were performed using both the SciPy and pyvttbl packages in Python (Appendix G). For the temperature, pH, and protease assays I used a balanced design two factor ANOVA with either temperature, pH, or protease as the first fixed factor, respectively, and enzyme (i.e. which mutant) as the second fixed factor to determine their effects and interaction effects on the outcome of the activity assays. The distributions of the residuals were checked for normality.
4. Results

4.1 Kinetic results

Research on threonine deaminase in Solanaceae and other organisms have concluded that in the absence of the allosteric inhibitor (isoleucine), threonine deaminase follows Menten-Michaelis kinetics (Sharma and Mazumder 1970; Wessel et al. 2000). I calculated the Menten-Michaelis kinetic parameters of $K_{cat}$, $K_m$, and the ratio $K_{cat}/K_m$ for $SITD2$ and the mutants. Figure 5 shows the results, with the $K_{cat}$ of $SITD2$, which is in close range of what has been previously reported (Vigil-Gonzales et al. 2011). This suggests that although the assays used to obtain the kinetic data were modified in comparison to the methods in Gonzales-Vigil et al. (2011), the results for $SITD2$ are comparable.
Figure 5: Menten-Michaelis $K_{cat}$ parameter for the four enzymes tested. Bars represent standard errors of the mean.
Table 3: A.) One-way analysis of variance (ANOVA) on the $K_{\text{cat}}$ comparison of the four enzymes. B.) Tukey’s Honest Significance Test to identify which enzymes were significant.

A. 

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Squares</th>
<th>F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>5663.834</td>
<td>3</td>
<td>1887.945</td>
<td>19.80</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Error</td>
<td>762.725</td>
<td>8</td>
<td>95.341</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6426.559</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. Tukey HSD: Table of q-statistics

<table>
<thead>
<tr>
<th></th>
<th>Mut1</th>
<th>Mut2</th>
<th>Mut3</th>
<th>SITD2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mut1</td>
<td>0</td>
<td>8.51 **</td>
<td>0.33 ns</td>
<td>0.77 ns</td>
</tr>
<tr>
<td>Mut2</td>
<td>0</td>
<td>0</td>
<td>8.84 **</td>
<td>9.28 **</td>
</tr>
<tr>
<td>Mut3</td>
<td>0</td>
<td>0</td>
<td>0.45 ns</td>
<td></td>
</tr>
<tr>
<td>SITD2</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

+ p < 0.1 (q-critical[4, 8] = 3.83462999257)
* p < 0.05 (q-critical[4, 8] = 4.52927264085)
** p < 0.01 (q-critical[4, 8] = 6.20441601908)

Figures 5-7 shows the plots for each Menten-Michaelis parameter for SITD2 and each mutant. Significance was tested for each parameter with a one-way ANOVA and followed with a Tukey’s means test to find which mutant differed significantly from the others and from SITD2. The $K_{\text{cat}}$ parameter showed a significant difference (Fig. 5) but the other two parameters did not indicate a significant difference amongst the enzymes tested (Fig. 6, 7). Mutant 2 was significantly different from the rest of the mutants and SITD2, showing a lower $K_{\text{cat}}$ rate than the others (Fig. 5, Table 3). This suggests that removing the critical ion pair between D100 and K245 affected the activity of this SITD2.
mutant, slowing down how much substrate gets catalyzed per second. Mutant 1 and Mutant 3 did not have significant differences from the kinetic parameters of \textit{SlTD2}.

![Figure 6: Menten-Michaelis $K_m$ parameter for the four enzymes tested. Bars represent standard errors of the mean.](image-url)
Table 4: One-way analysis of variance (ANOVA) on the $K_m$ comparison of the four enzymes.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Squares</th>
<th>F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>0.02</td>
<td>3</td>
<td>0.007</td>
<td>1.773</td>
<td>0.23</td>
</tr>
<tr>
<td>Error</td>
<td>0.03</td>
<td>8</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.049</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 7: Menten-Michaelis $K_{\text{cat}}/K_m$ ratio for the four enzymes tested. Bars represent standard errors of the mean.
Table 5: One-way analysis of variance (ANOVA) on the \( \frac{K_{cat}}{K_m} \) comparison of the four enzymes.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Squares</th>
<th>F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>94725.893</td>
<td>3</td>
<td>31575.298</td>
<td>3.625</td>
<td>0.064</td>
</tr>
<tr>
<td>Error</td>
<td>69691.911</td>
<td>8</td>
<td>8711.489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>164417.804</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.2 Stability

One of the main objectives of this study was to test the effects of each critical ion pair on stability of the \( SITD2 \) enzyme. Studies have shown that critical ion pairs increase the thermostability of enzymes, particularly in those enzymes from thermophiles and psychrophiles (Szilágyi and Závodszky 2000; Bae and Phillips 2004). I designed a set of tests to understand the effect of critical ion pairs on stability. The general principle was to first stress the enzymes to destabilize them, and then tests how stable the mutants were after the stressed phase in comparison to \( SITD2 \) by testing how active the enzymes were after applying the stressor. If there was a loss of stability, the stress would denature a certain amount of the enzyme which would be noticeable in the activity tests by lowered activity. This assumes that a denatured enzyme is unlikely to refold properly, and indeed, tests on \( SITD2 \) and mutants that were completely denatured by heat did not refold properly when returned to normal temperature (not shown).
4.2.1 Temperature assay

![Graph showing activity vs temperature for each enzyme tested from 30 to 90 °C. Error bars are standard error of the mean.](image)

**Figure 8:** Activity vs Temperature for each enzyme tested from 30 to 90 °C. Error bars are standard error of the mean.

This first stability test performed tested stability across a temperature gradient. I set up an enzyme assay to test the activity left of the enzyme after exposure to a certain temperature for 20 minutes to determine how much of the enzyme stayed active. Figure 8 shows the results of this assay, with temperature ranges tested from 30°C (standard conditions) to 90°C. Adding this much energy to the system increases entropy, and thus, disorder of the enzyme leading to irreversible denaturation. The plot does not show an overall large difference between the four enzymes in activity. Results show that Mutant
Table 6: Two-way ANOVAs of activity by enzyme and temperature.

A. All four enzymes

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.006864</td>
<td>3</td>
<td>16.443852</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Temp</td>
<td>1.531776</td>
<td>8</td>
<td>1376.195747</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*Temp</td>
<td>0.014283</td>
<td>24</td>
<td>4.277325</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Residual</td>
<td>0.010017</td>
<td>72</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. TD2 vs Mutant 1

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.000845</td>
<td>1</td>
<td>6.156742</td>
<td>0.018</td>
</tr>
<tr>
<td>Temp</td>
<td>0.795808</td>
<td>8</td>
<td>724.422152</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*Temp</td>
<td>0.002495</td>
<td>8</td>
<td>2.270912</td>
<td>0.044</td>
</tr>
<tr>
<td>Residual</td>
<td>0.004943</td>
<td>36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. TD2 vs Mutant 2

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.001311</td>
<td>1</td>
<td>7.137253</td>
<td>0.011</td>
</tr>
<tr>
<td>Temp</td>
<td>0.80967</td>
<td>8</td>
<td>550.79041</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*Temp</td>
<td>0.002373</td>
<td>8</td>
<td>1.613944</td>
<td>0.155</td>
</tr>
<tr>
<td>Residual</td>
<td>0.006615</td>
<td>36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D. TD2 vs Mutant 3

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.006678</td>
<td>1</td>
<td>47.143746</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Temp</td>
<td>0.784185</td>
<td>8</td>
<td>692.030888</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*Temp</td>
<td>0.005875</td>
<td>8</td>
<td>5.184284</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Residual</td>
<td>0.005099</td>
<td>36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3 has a noticeable initial percentage drop in activity when temperature increased from 30°C to 40°C. Further, Mutant 3 tends to at a higher activity level than the other enzymes, including SITD2. Additionally, Mutant 2 starts at a lower activity level than all the other enzymes at 30°C and remains below the other enzymes until around 80°C. I performed a two-way analysis of variance (ANOVA) to tease apart the effects that temperature has in conjunction with the enzyme on the measured activity of the temperature assay. Tables 6A-D show the results of the statistical analysis. Table 6A shows that with all enzymes tested there is a significant effect of enzyme, temperature, and a significant interaction between enzyme and temperature.

To further assess which enzymes are driving this and to see if it is different from SITD2, I performed a two-way ANOVA of each mutant against SITD2 (Tables 6B-D). All of these comparisons showed a significant enzyme and temperature effect. Mutant 1 and Mutant 3’s effects were both significantly higher (i.e. increased activity) for the effects of enzyme, temperature, and enzyme*temperature, perhaps suggesting that these two critical ion pairs destabilize SITD2. On the other hand, Mutant 2 was significant but in the opposite direction. Mutant 2 had significant lower activity, suggesting that the second critical ion pair stabilizes the enzyme.
4.2.2 pH assay

In addition to testing the increase in thermostability of each mutation, it is possible that the critical ion pairs could play a role in the high activity of SITD2 in alkaline conditions. Considering that SITD2 is primarily active in the lepidopteran gut with pH’s ranging from 8-10, it is very plausible that the critical ion pairs provide the stability needed to survive the harsh alkaline environment (Berenbaum 1980). Thus, I hypothesized that the critical ion pairs may play a role in increasing stability in higher pH environments. To test pH, an activity assay was run in triplicate at standard
conditions except for varying pH, which varied from acidic to very basic conditions (pH 3 to pH 13). Figure 9 shows the values of the assay with standard error of the mean plotted. For all enzymes no significant activity was captured until pH 6, steadily rises to a peak at pH between 8 and 10, and then decreasing in activity until pH 12 with a sharp decline at pH 13 down to <10% activity. The highest pH activity was at pH 9, suggesting that the enzyme is most efficient near this level of alkalinity, in line with previous accounts of pH optimum in \textit{SlTD2} (Gonzales-Vigil et al. 2011). Similarly to the temperature assay, both Mutant 1 and Mutant 3 show higher levels of activity compared to \textit{SlTD2}, particularly around pH 8-11. Conversely, Mutant 2 indicates it has lower activity than the other mutants and TD2.

A two-way ANOVA was performed to analyze and confirm the differences between the activity levels across the range of pH’s tested. Table 7A shows the results of the two-way ANOVA for all four enzymes, indicating that there is a pH and Enzyme effect, but no significant interaction effect of pH\(\times\)Enzyme. The analysis was then applied to each mutant against \textit{SlTD2} (Table 7B-C). Mutant 1 has a significant Enzyme and pH effect indicating that its higher activity across the pH ranges is driven by the removal of the critical ion pair. This suggests that the first critical ion pair reduces activity across the range of pHs, and perhaps plays a destabilizing role. Similarly, Mutant 3 shows a significant Enzyme, pH, and pH\(\times\)Enzyme interaction effect perhaps suggesting a
destabilizing role. On the other hand, Mutant 2 only has a significant pH effect. This suggests the second critical ion pair may increase stability.
Table 7: Two-way ANOVA on activity by enzyme and pH.

A. All four enzymes

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.01473</td>
<td>3</td>
<td>6.42</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>pH</td>
<td>5.556556</td>
<td>10</td>
<td>726.27</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*pH</td>
<td>0.03481</td>
<td>30</td>
<td>1.52</td>
<td>0.069</td>
</tr>
<tr>
<td>Residual</td>
<td>0.067327</td>
<td>88</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. TD2 vs. Mutant 1

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.005127</td>
<td>1</td>
<td>5.63</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>pH</td>
<td>2.889178</td>
<td>10</td>
<td>317.32</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*pH</td>
<td>0.011949</td>
<td>10</td>
<td>1.31</td>
<td>0.254</td>
</tr>
<tr>
<td>Residual</td>
<td>0.040061</td>
<td>44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. TD2 vs. Mutant 2

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.000244</td>
<td>1</td>
<td>0.25</td>
<td>0.623</td>
</tr>
<tr>
<td>pH</td>
<td>2.66641</td>
<td>10</td>
<td>268.47</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*pH</td>
<td>0.008783</td>
<td>10</td>
<td>0.88</td>
<td>0.555</td>
</tr>
<tr>
<td>Residual</td>
<td>0.043701</td>
<td>44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D. TD2 vs. Mutant 3

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.006775</td>
<td>1</td>
<td>7.05</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>pH</td>
<td>2.774618</td>
<td>10</td>
<td>288.75</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*pH</td>
<td>0.021417</td>
<td>10</td>
<td>2.23</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Residual</td>
<td>0.04228</td>
<td>44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.2.3 Protease resistance assay

Figure 10: Protease resistance assay using chymotrypsin to cleave protein followed by an enzyme activity assay. Control was incubated without chymotrypsin (uncleaved). Bars represent standard error of the mean.

Lastly, protease resistance plays an important role for the SITD2 enzyme to survive the lepidopteran gut. In addition to surviving highly alkaline guts and varying temperatures that the lepidopterans could experience, the enzyme must stay active while being constantly assaulted by digestive proteases in the gut. It is critical that SITD2 be resistant to proteases, but it is a very difficult feat to achieve without sacrificing much activity due to increased stability (Beadle and Shoichet 2002; Tokuriki et al. 2008; Studer et al. 2014). The main digestive protease in the lepidopteran insect gut that feed on Solanaceous plants are serine-based proteases which include...
chymotrypsins. Chymotrypsins are very effective at digesting the tomato housekeeping copy of threonine deaminase (SlTD1). Studies have shown that SlTD1 does not survive through the insect gut (Chen et al. 2007), and my own tests have revealed that SlTD1 is not detectable in SDS-PAGE after exposure to chymotrypsin (Appendix E). However, SlTD2 has evolved a mechanism to evade proteases in the insect gut while staying active enough to deplete free threonine. Impressively, Chen et al. (2007) found that SlTD2 was still active after it had passed through the insect gut. Thus, I set out to test if critical ion pairs played a role in protease resistance in the evolution of SlTD2. I tested this by incubating each of the enzymes with chymotrypsin, followed by an activity assay and compared it to a control that lacked the chymotrypsin in the buffer.

Figure 10 compares the enzyme’s activity in the absence and presence of chymotrypsin protease. Considering SlTD1 is very susceptible to chymotrypsin digestion (Gonzales-Vigil et al. 2011), it is likely to expect a decrease in activity of SlTD2 exposed to chymotrypsin as tradeoff for the ability to remain active post proteolytic cleavage. Surprisingly, SlTD2 and all three mutants had a significant increase in activity after protease exposure (Table 8). Previous research showed that the regulatory domain on the C-terminal end of the protein is cleaved off in the insect gut, but the remaining portion of the enzyme remains active (Gonzales-Vigil et al. 2011). SDS-PAGE analysis (Appendix E) confirmed that the chymotrypsin protease cleaved the enzymes used in this study.
Table 8: Two-way ANOVA on activity by enzyme and proteolytic cleavage.

A. All four enzymes

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.227136</td>
<td>3</td>
<td>20.70</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Protease</td>
<td>0.113561</td>
<td>1</td>
<td>31.05</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*Protease</td>
<td>0.004816</td>
<td>3</td>
<td>0.44</td>
<td>0.728284</td>
</tr>
<tr>
<td>Residual</td>
<td>0.058526</td>
<td>16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B. TD2 vs. Mutant 1

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.029373</td>
<td>1</td>
<td>63.45</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Protease</td>
<td>0.041313</td>
<td>1</td>
<td>89.25</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*Protease</td>
<td>0.000719</td>
<td>1</td>
<td>1.55</td>
<td>0.247857</td>
</tr>
<tr>
<td>Residual</td>
<td>0.003703</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C. TD2 vs. Mutant 2

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.055298</td>
<td>1</td>
<td>7.94</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Protease</td>
<td>0.060095</td>
<td>1</td>
<td>8.63</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Enzyme*Protease</td>
<td>0.000472</td>
<td>1</td>
<td>0.68</td>
<td>0.434271</td>
</tr>
<tr>
<td>Residual</td>
<td>0.055732</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

D. TD2 vs. Mutant 3

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>F</th>
<th>PR(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>0.039652</td>
<td>1</td>
<td>189.77</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Protease</td>
<td>0.041866</td>
<td>1</td>
<td>200.37</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Enzyme*Protease</td>
<td>0.000794</td>
<td>1</td>
<td>3.80</td>
<td>0.087</td>
</tr>
<tr>
<td>Residual</td>
<td>0.001672</td>
<td>8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The two-way ANOVA found that there were no significant interaction effects of the enzyme and proteolytic cleavage (Table 8B-C). However, for all four enzymes and for each individual mutant tested against SlTD2, the data revealed significant effects for enzyme and for proteolytic cleavage. The enzyme effect for Mutant 1 and Mutant 2 compared to SlTD2 were both higher, while the enzyme effect for Mutant 3 was lower. This was a slightly different pattern from the pH and temperature assays, where Mutant 2 was lower and Mutant 3 was higher in activity in both assays. Overall, these results show that the enzyme remains active after cleavage and, more importantly, the enzymes increase their activity for each of the four enzymes tested.
5. Discussion

Overall, this study examined the contribution that each of the three critical ion pairs make to the enzymatic efficiency and stability of the TD2 enzyme in tomato. Removal of the critical ion pairs did not significantly affect the kinetic parameters except in the case of Mutant 2, where it showed a significant drop in $K_{cat}$ compared to $SlTD2$ and the other mutants. However, the catalytic efficiency parameter, $K_{cat}/K_m$, was not significantly different, perhaps suggesting that the evolution of the second critical ion pair likely has little effect on activity and may have been selected for other reasons. However, stability tests revealed a more complicated picture of the effects of the critical ion pairs on $SlTD2$. The temperature assays showed that at 30°C Mutant 2 had lowered activity than $SlTD2$, but Mutant 3 had a higher activity. The pH assay showed that Mutants 1 and 3 were more active than $SlTD2$, and Mutant 2 was lower than all the others. And lastly, the protease assay revealed that the proteolytic cleavage of all the enzymes lead to an increase in activity rate. These results suggest a complex interaction and role for the critical ion pairs in the evolution of the $SlTD2$ enzyme.

5.1 Kinetic parameters

The Menten-Michaelis parameters shown in Figures 5-7 overall suggest that perhaps critical ion pairs do not affect kinetics directly. However, one significant difference in the Mutant 2 $K_{cat}$ parameter was found ($p < 0.001$). The decline in activity of Mutant 2 as observed in Figure 5 could be due to two phenomena occurring at the
molecular level. First, it could be that removing this critical ion pair causes a lowered affinity for the threonine substrate. However, if this was the case, it would also be expected to see a change in the $K_m$ of Mutant 2, which is not the case. Second, the loss of the critical ion pair is expected to make the enzyme more flexible and perhaps the loss of rigidity in one part of the enzyme leads to an enzyme that is too flexible to be efficient at turnover or at performing the deamination efficiently. Similarly, the loss of rigidness of the enzyme could create a more disordered enzyme that may not fold correctly or may not stay folded, leading to fewer active enzymes deaminating the substrate, which would show the pattern I observed of a loss in $K_{cat}$ but not in $K_m$ in this enzyme.

**5.2 Thermostability**

The temperature assay used a range of temperatures, from the standard condition of 30°C all the way to a range only enzymes from thermophilic organisms are known to survive. *SITD2* and the mutants all had quite high activity through about 80°C or so. However, they did not perform the same across the temperature ranges. Both Mutant 1 and Mutant 3 performed significantly higher across the temperature assay than *SITD2*, and Mutant 2 was significantly lower. It is possible that the removal of the critical ion pairs could both decrease stability or increase activity and depending on how dominant one is over the other, the outcome could look like the critical ion pair helps in increasing stability or affects activity but it can be doing both.
Mutant 2 starts with a lower activity than the other enzymes at 30°C, and stays below the other enzymes until they start to sharply decrease in activity at around 60°C. While 60°C is not a realistic temperature that the enzymes would be facing in nature, 30—40°C is a realistic temperature that the enzyme would need to be able to keep stable at in order to catalyze efficiently. Further, a secondary peak is observed at around 75°C, which is likely explained by the increase in energy creating an increase in the activity of the enzyme that allows it to overcome the loss of activity by denaturation of the enzyme (Bisswanger 2014). It is difficult to conclude from these results that the critical ion pairs individually evolved to specifically increase thermostability, but they do affect activity levels at different temperatures and it could be that Mutant 2 did lose stability due to the removal of the second critical ion pair.

### 5.3 Stability in high alkalinity

The pH assays tested if a change in pH lead to a large change activity. Enzymes tend to have an optimal pH that maximizes the activity, all else being held constant. *SITD2* appears to be optimized to catalyze around pH 9, similar to the guts of the lepidopteran herbivores it targets (Berenbaum 1980), *SITD1* pH optimum is around pH 8 (Gonzales-Vigil et al. 2011). The pH assay was intended to test the hypothesis of whether any of the critical ion pairs helped in either shifting the optimum pH of *SITD2* or generally provided more stability in order to stay active in higher alkaline environments.
The pH assay did not detect any large shifts in optimal pH (Fig. 9). It may be that the test was too coarse to identify such a shift, as pH’s in increments of whole integers were tested and perhaps the critical ion pairs only changed pH optimums on the order of $10^{-1}$ or smaller. It is very likely very many amino acid substitutions lead to the increase in pH optimum outside of the critical ion pairs. However, the pH assays did detect significant differences between the enzymes and pH levels in terms of activity measured. A similar pattern as seen in the thermostability assay emerged: Mutants 1 and 3 had significantly increased activity compared to $SlTD2$ at almost all pH levels tested where activity was detected (Fig. 9, Table 7). This means that the first and third critical ion pairs are both decreasing activity at these pH levels, which further suggests that perhaps these critical ion pairs are destabilizing. Conversely, Mutant 2 displayed lowered activity than all the other enzymes (Fig 9, Table 7). This supports the original prediction that critical ion pairs increase stability as the second critical ion pair helps increase activity when present.

The results from both the temperature and pH assays lead to an apparent paradox. The results of Mutant 2 suggest that the evolution of the second critical ion pair helped increase stability of $SlTD2$. On the other hand, the results of Mutant 1 and Mutant 3 suggest the opposite, the perhaps destabilize $SlTD2$. The simple explanation is that the role of critical ion pairs in stabilizing enzymes is misunderstood, but it is more likely that their role is much more complex. The effect of the critical ion pair may
depend on the context that they arise in. These epistatic effects are difficult to
understand, particularly intramolecularly, where they can be interacting with any of the
other amino acids in the enzymes, or multiple ones as well. One explanation is that
critical ion pairs 1 and 3 are interacting in an antagonistic epistatic manner with other
amino acids so that it experiences reciprocal sign epistasis. Clearly, the interactions of
these critical ion pairs are much more complex than originally thought.

5.4 Proteolytic cleaving of regulatory region of SITD2

Previous mass-spectrometry analysis has found SITD2 to be largely intact except with the regulatory domain missing, and further tests confirmed that the regulatory domain was cleaved off by proteases but the catalytic region remained unharmed and still active (Chen et al. 2007). I hypothesized that the critical ion pairs played a role in increasing stability of SITD2 catalytic domain in order to escape proteolytic degradation (but allow regulatory degradation to avoid isoleucine regulation through negative feedback, not tested here).

I found a surprising effect during the protease assay. All of the enzymes tested showed an increase in activity after exposure to proteases (Fig. 10). This type of behavior seems to be different than that of other enzymes that have been studied that go through proteolytic cleaving. For instance, many digestive enzymes undergo proteolytic cleaving (for instance, pepsins, trypsins, and chymotrypsins exist in an inactive zymogen form) in order to become active but don’t exhibit any activity before cleavage (Scheid and
Choppin 1974; Park et al. 2008). In this case, the enzyme is very active before proteolysis, but catalytic activity is significantly increased when it undergoes proteolytic cleaving by proteases.

It is likely that the protease resistance of the catalytic domain was a cause of natural selection. While regulatory function is essential for non-defensive TD genes like \textit{SLTD1} to regulate activity in the plant cell, e.g. too much isoleucine is wasteful and may not be needed, in the defensive \textit{SLTD2}, inhibition is no longer necessary and may negatively impact its function. There are two apparent advantages to having the regulatory domain cleaved. The first is that proteolytic cleaving removes inhibition by isoleucine, which is present in the gut of the herbivore, thus allowing the enzyme to perform at maximum efficiency. The second is that it increases the activity of the enzyme, since in the presence of the regulatory domain apparently inhibits activity, even in the absence of isoleucine. The higher activity could be a result of greater flexibility of the enzyme without an extra 195 amino acids added on the C-terminal of the enzyme. One can also imagine that there may have been a tradeoff of trying to keep both the catalytic and regulatory domain protected from digestion, but allowing the regulatory region to be removed could have allowed selection to more easily optimize catalysis and stability of the catalytic. Further studies in assessing the Menten-Michaelis parameters and performing temperature and pH assays for the cleaved enzyme would be informative to better compare the activity of pre and post-cleaved \textit{SLTD2}. In addition, all
three critical ion pairs are found in the catalytic region of the enzyme, and thus while testing singly mutations that removed these ion pairs revealed some small effects of stability, it could be that their effects are larger in the cleaved enzyme.

Of further note is that Mutant 3 resulted in lower activity than all the other enzymes in both treatments of the protease assay (cleaved and uncleaved). This is somewhat surprising based on the previous results of the temperature and pH assays, where Mutant 3 had as high or higher activity than the other enzymes and significantly higher activity than SITD2. This result could be due to the long incubation period for the protease assay (2 hours) in comparison to much smaller incubation periods in the other assays. This could have led to more denaturation if Mutant 3 was less stable (but more active due to increased flexibility as observed in previous tests). This is consistent with the idea that the third critical ion pair provided stability to the enzyme, but at a cost of activity—thus removing the critical ion pair results in an enzyme with lower stability but higher activity.

An alternate argument may propose that the lack of large effect differences between SITD2 and the mutants suggests that the critical ion pairs are effectively neutral and arose and fixed through genetic drift. While it could be the case that the critical ion pairs are not playing a role in affecting activity or stability in the enzyme, it seems somewhat unlikely that all three arose through genetic drift when much evidence in the literature show ionic pairing affects stability, and sometimes activity, due to the nature
of the ionic pair bond and ANOVAs suggested significant effects in the stability of the
different mutants (Bosshard et al. 2004). It is also likely that some of the effect sizes may
have been too small to detect efficiently with the methods of this study. This could be
due to the sensitivity in of the instruments, or the setup or execution of the assays.
Enzymology also uses unrealistically large amounts of substrate and enzyme that would
never be found in biological systems in order to be able to measure the results. This
could either increase or decrease the real differences between enzymes when measured
in the manner as done in this study. Further, it is difficult to recreate the natural system
in an *in vitro* experiment that the enzyme is active in, which could lead to losing the
ability to detect the smaller effect sizes.

Another possibility is that the critical ion pairs could be essentially neutral on
their own but are synergistically affecting either activity or stability of the enzyme.
Epistatic interactions of this nature are widespread in molecular evolution, and the
epistasis could be between two critical pairs or between a critical ion pair and non-
ionically paired intramolecular residue. Finding the latter would be very time intensive
as the vast majority of the residues in *SlTD2* are not ionically paired and at this time no *a
priori* data exists to help limit which residues might be interacting epistatically with the
critical ion pairs. One opportunity to follow up on is to test if any two, or all three, of the
three critical ion pairs are synergistically affecting stability or activity in *SlTD2*. This
scenario may be more likely than any of the others mentioned as synergistic effects have
been found in other ion pairing enzymes, usually with the higher number of ionic pairs an enzyme has the more stable it is (Vogt et al. 1997; Szilágyi and Závodszky 2000; Bosshard et al. 2004).
6. Conclusion

This study hypothesized that the \textit{SI}TD2 enzyme evolved three critical ion pairs to increase stability of the enzyme in its new environment of the lepidopteran herbivore gut, albeit at a potential cost to activity. Kinetic assays showed little effects of the critical ion pairs on activity except for critical ion pair 2, as $K_{cat}$ decreases without this ion pair. The temperature and pH stability assays showed that the Mutations 1 and 3 were more active, but Mutation 2 was less active possibly due to its role in added stability to the \textit{SI}TD2 enzyme. While the protease showed the cleaved version was more active than the uncleaved version for all the enzymes tested. This was surprising and may suggest that this functionality was selected for. Overall, the evidence points to some, if small, role of the critical ion pairs in \textit{SI}TD2 were found to affect both activity and stability, with a possible trend toward an activity vs. stability tradeoff with Mutant 2. Further testing of double and triple mutants and of the cleaved versions of these enzymes against the uncleaved versions may reveal contributions of the critical ion pairs that were not observable in the testing done in this study.
Appendix A: Phylogeny with ion pair residues mapped
Appendix B: PAML codeml.ctl file

codeml.ctl:

seqfile = tdalignment.phy * sequence data filename
treeview = GTRIG.tre * tree structure file name
outfile = lg_mlc * main result file name

noisy = 9 * 0,1,2,3,9: how much rubbish on the screen
verbose = 2 * 0: concise; 1: detailed, 2: too much
runmode = 0 * 0: user tree; 1: semi-automatic; 2: automatic
* 3: StepwiseAddition; (4,5): PerturbationNNI; -2: pairwise

seqtype = 2 * 1: codons; 2: AAs; 3: codons-->AAs
CodonFreq = 2 * 0: 1/61 each, 1: F1X4, 2: F3X4, 3: codon table

* ndata = 10

clock = 0 * 0: no clock, 1: clock; 2: local clock; 3: CombinedAnalysis

aaDist = 1 * 0: equal, + geometric; -: linear, 1-6: G1974, Miyata, c, p, v, a

aaRatefile = lg.dat * only used for aa seqs with model = empirical(_F)
* dayhoff.dat, jones.dat, wag.dat, mtmam.dat, or your own

model = 3
* models for codons:
  * 0: one, 1: b, 2: 2 or more dN/dS ratios for branches
* models for AAs or codon-translated AAs:
  * 0: poisson, 1: proportional, 2: Empirical, 3: Empirical+F
  * 6: FromCodon, 7: AAClasses, 8: REVaa_0, 9: REVaa(nr=189)

NSsites = 0 * 0: one w; 1: neutral; 2: selection; 3: discrete; 4: freqs;
* 5: gamma; 6: 2 gamma; 7: beta; 8: beta & w; 9: beta & gamma;
* 10: beta & gamma + 1; 11: beta & normal > 1; 12: 0 & 2 normal > 1;
* 13: 3 normal > 0

icode = 0 * 0: universal code; 1: mammalian mt; 2-10: see below

Mgene = 0
* codon: 0: rates, 1: separate; 2: diff pi, 3: diff kappa, 4: all diff
* AA: 0: rates, 1: separate

fix_kappa = 0 * 1: kappa fixed, 0: kappa to be estimated
kappa = 2 * initial or fixed kappa
fix_omega = 0 * 1: omega or omega_1 fixed, 0: estimate
omega = .4 * initial or fixed omega, for codons or codon-based AAs

fix_alpha = 0 * 0: estimate gamma shape parameter; 1: fix it at alpha
alpha = 0.5 * initial or fixed alpha, 0: infinity (constant rate)
Malpha = 0 * different alphas for genes
ncatG = 4 * # of categories in dG of NSsites models

getSE = 0 * 0: don’t want them, 1: want S.E.s of estimates
RateAncestor = 1 * (0,1,2): rates (alpha>0) or ancestral states (1 or 2)

Small_Diff = .5e-6
cleandata = 0 * remove sites with ambiguity data (1:yes, 0:no)?
* fix_blength = -1 * 0: ignore, -1: random, 1: initial, 2: fixed
  method = 0 * Optimization method 0: simultaneous; 1: one branch a time

* Genetic codes: 0:universal, 1:mammalian mt., 2:yeast mt., 3:mold mt.,
* 4: invertebrate mt., 5: ciliate nuclear, 6: echinoderm mt.,
* 7: euplotid mt., 8: alternative yeast nu. 9: ascidian mt.,
* 10: blepharisma nu.
* These codes correspond to transl_table 1 to 11 of GENEBANK.
Appendix B: PAML GTRIG.tre file

((((((((((((((((S_MELO_TD1:0.04308098,((((S_ARCANUM_TD1:0.00709077,S_LYCO
_TD1:0.00572077):0.00663446,S_SITI_TD1:0.00496974):0.00324140,(S_BULB_TD1:0.005637
74,S_TUBE_TD1:0.00656244):0.00182541):0.02167222,S_DULC_TD1:0.01649321):0.001967
28):0.00443696,S_AVIC_TD1:0.01309764):0.00774421,(J_PROCUMB_TD1:0.01683931,S_PS
EU_TD1:0.02028851):0.01537710):0.00583645,(CAPSIC_TD1:0.04200949,IOCHRO_TD1:0.
02675163):0.01254769,(L_BARB_TD1:0.04617420,N_PHYS_TD1:0.02296457):0.00659868):0.
00253744):0.02873347,(((N_ATTEN_TD1:0.00403127,(N_SYLV_TD1:0.00072530,N_TABA
_TD1B:0.00825433):0.00300363,N_BENT_TD1:0.01711007):0.00478417,(N_TABA_TD1A:0.00000107,
N_TOME_TD1:0.00000163):0.01162275):0.02615045):0.01506652,PE
TUNIA1:0.05854353):0.04841059,(((CAPSIC_TD3:0.06434373,(S_LYCO_TD3:0.02809184,
S_TUBE_TD3:0.03862511,S_MELO_TD3:0.08399964):0.03150791):0.04566438,(N_ATTEN
_TD3:0.01325955,N_SYLV_TD3:0.00771854):0.01372446,(N_TOME_TD3B:0.00
00107,N_TOME_TD3A:0.00000107):0.01363398):0.04311311):0.02204730,PETUNIA3:0.10
379818):0.06291232):0.02729305,(((AQUILEGIA:0.33141602,(((A_THALIANA:0.03588227,CAPSELLA_RUBELLA:0.07338143):0.24463469,CUCUMUS_SATIVA:0.22384003):0.052
18114,(((CICER_ARIETINUM:0.11589514,MEDICAGO:0.11271417):0.07455121,GLYCINE_MAX:0.10797040,(FRAGARIA_VESCA:0.15115846,MALUS:0.12196614):0.
07911366):0.02553851,(POPULUS_TRICHOCARPA:0.10734544,RICINUS_COMMUNIS:0.
12487268):0.04223310):0.05535949,VITIS_VINIFERA:0.17749354):0.01568796):
0.04665378,(HELIANTHUS_ANNUUS:0.10764313,(LACT_SATIVA:0.00401604,LACT_S
ERIOLLA:0.00708020):0.13242815):0.16775165):0.05342599,MIMULUS_GUTATUS:0.20
60292):0.04568539,((I_BATATAS:0.00370896,I_LAC_1:0.00656587):0.19318607,I_LAC_2:0.
15264047):0.06825943):0.07212409):0.02136572,((((N_ATTEN_TD4:0.01170204,N_TOME
_TD4:0.02027651):0.00006361,N_BENT_TD4:0.03795853):0.04609178,S_TUBE_TD4:0.14828
109):0.02054287,PETUNIA4:0.09444473):0.05785388):0.23972468,(((N_ATTDE_TD2:0.00419
30,N_BENT_TD2A:0.00960504):0.00795493,N_BENT_TD2B:0.03150791,(N_TOME_TD2:0.00
0000204, N_TABA_TD2:0.00000107):0.01803791):0.00855680):0.06517261):0.08042857,(IO
CHRO_TD2:0.04819149,L_BARB_TD2:0.09829218):0.00005208):0.03043091,CAPSIC_TD2:
0.32284480):0.05622179,((S_DULC_TD2:0.08739358,S_MELO_TD2:0.09664139):0.05136292,
S_PSEU_TD2:0.21095288):0.04533327):0.09979614,(S_BULB_TD2:0.05949239,S_TUBE_TD
2:0.02337703):0.01699246):0.06407454,S_LYCO_TD2:0.01387281,S_PERU_TD2:0.01387348)
Appendix D: PAML tdalignment.phy file

69 1605
AQUILEGIA     CAGCCAATA---
ATGGAAAAAGGGTTACAACAGATTCGCTTAATAT---
AGTGAGATCTCTGGAGCCTCGAGAATACCTGATTTTAATGATAGA----------
TACTTGACCAATATAATTAACCTCTAAGGTTATAGTGGTCTATTGAATCACCATTT
CAGGTTGGCAAATGAAGCCTTCTCAGAAGACTTTGGTACTAATATTTGGCTTAAAAGA
GAAGATATGCAACAGCGCAGAATAATGGAGAAATCTCTGCTGACAGCAGTGGATG
ATTGTATGCTCTGTCAAACGCCAGAAATAAGTGGAAATCTCTGCTGAGAGCTATG
GGTGCAACTGTGTTCTCTAAAAAGGGATGCTACTAGATGAAAGCCAAGCACATGCT
AAACAGGAAATGGAGAAGTTCCGAGACTTTTTTACCTCTTTTGATCACACAAGC
GATGCAATTTGTTGGGCGAGAACAGCTTGGGATGGGAGATTTTGCGTCAAGCAGAC
GAATCAGTGCATGCCGTTTATTGTCTTCTGTTGGGTTGGCAGTTACATGAGAGG
TTGCTTCCATATTGGAAGTTGGTGGCTGGGCTAATTGGGATGTGCTCTTCTGCT
AGCTGAGTCAAAATATTGTCTGTTCTCTCTCATGTTGCAAGCTATGCTAGCTG
GAGCAGGTGGGAGGCTTTGCCAGACTGTGTTCTGTGTTAAAAAGTTGTTGGAAG
ACITTTCCGGTTTTGTCGAGAATGTGGTGTAGTGATGTGTTCTCTGCTGACAAAG
CTATATGTGCATCAATAAAGGAATATGGTGGAAACAGAGGAGCATTTGGGAAC
CAGGGGGCGCTCTTGCCTAGCTGGAAGACAGTACATGCAAATTTTATAAGCT
CAAGGCGGAAAATGTCTTGGCAGCTCAACACTGCTGCGCAAACATGCTTTGATA
GACTGAGGCTGGTAAACTGAACCTGGGAGAGAGAGGCTGCC
TTGCAACTTGGCATATCCCAGAGGCTGACCAAGAAGCTGGTAAAAAGTTTGGGAAGT
GGTAAAGATGAATATCATCACAGTGTTCAAAATACAGATATGAGCTAGGGAAGAAA
GAAGCTCTTGGTTCTCAGTGTCGTGAGTGGTGCTTCTGAACCAGTTGTGAGC
TGTTGAGCCGGAATGCTTGGCAGCTCAACACTGCTGCCAAACATGCTTTGATA

A_THALIANA
CTTCCTTTACCACCTCTCTTAAGGCTCTTCGAGAATACCTGCCGGTTAAGCTA
CCTCGGTTCTCCGAGAAAGCTGACGAGGCTGAGAACGGAAG---
ATGGAGATTGAGCAATATACCTGTCCATACAGGTTATACGACATCGCCATTTGAGT
CACCACCATGTTGCAAGAGAGCTATCTCAAGAGAGATGGCTGTTCTCGATATCT
TAAAAGAGAAGACTTGCAACCTGTATTTCTCAGTTGGAGGCTTACAAT
ATGATGGTGAACCTCCACGAGATGCAATTTGGCGAAGAGGTATCTGCTTCTCA
GCTGGAACCATGCTCAGAGGAGTTGATTATATCTGCTAGTAAACTCGGCGTCGACTG
CTGATAGTTATATGGCTGTTACGACTCTCGAGATAAAGTGGCGAAAGCTATGAGAGA
ATTTGGGCTGAACGGGAGTTGCTATGCTGTACAAGACACAAGCACA
TGCTAAGATACGGCAGCTGAAAGAGAGGCTGAGTTATGTGAATGCTAGAAG
ACCAGCTGAGCGCAAATGCAATGGCTTTGTCGCTAGCATTGGAGAGGAGGAT
ATTGAGCACAGGGGGGATGTGGTTGAGTACGTCAGGCTAAGGAGGAGGCTAGG
AGAGACTTTCGTATAACGAGAAATCTAATGGATGTTGTTGTTCTGACTCTGT
GATGCTATTGTGCTACATAAAGAGATTGTTGGAGAGACAAGCATTTAG
GAAACACGGACGCTCTTCTGACACTCGGAGTTGGAGCACACTGTGAATATATATAT
GGGCTAAAGGAGACGTGATGTGCTAGCCTCAAAACAGGCGCTAAGATGAGAG
GACAGCTAAGAGATTGGACATGCTGAACGTGAAATGCTTCTGACTCTGT
GTTTCTGACTCTCATGCCGAGAATCTCGGACATTTTAAAGCAATTTTGTGAG
CTGTTTGGAGACATGAGGAGTCTTATGACTGAGTTGCTGAGTAC---
GAAAAGGAGGCGTATTCTACTATACAGTCTGCGAGTTTGCCACACAGTGGAGAGCTC
AAAGCACTACAGAAAGAGAATGGGAAATCTTCTCTCAACTTCAACTCAAAACTGCAATCTC
ACCAGTGACTTATGGAAGACTACCTGTGTTACTTGGGAGAGAAGACTCT
GGTGGAGACAGGAGGTTTCTATGCGGATTCCCTTCCGAGAGACCTGGTGTCTAA
TGAACTCTGGGACCTTCTCAGCTCCACAGTTGGAACATGACCCCTTTTTTATACAGGT
GGACAGGTGAGACGCGCGGCAATGTTGCTGTGAGATGCTTACGGGAGATGGAGTGAGG
AGAAAATGGAGGAAATTAAAAACCGAGCTAAAGCTCTTGAGATTACGTACTCTCTT
AGTAAAGTGATGACGAGACTTTTATTTACGCTTCTGATGCAC
CAPSELLA_RUBELLA
CCTCCTCTGCGACGTCAAGGTCTCCTCGGAGCTCCTGTTGCAATACCCCTGCGGTT
ATCTCGGGTCTCGGAGCTACAGGAACCATGGCTGCGGAGAGGACAGGAGAGAG---
ATGGGATATTGGACAGAAATATACGTCTCCACTAAGGTTTACGACATCGCCATTGAGT
CACCACCTCAGTTGCGAAGAGACTGCTCCTAAGAGGTTAGTGTTGCTGATTGCTTCT
CAAAAGAGAACACTTGGCAACCTGTGTTCTGATGG---
CGTGGAGCTTACAATATGATGTTGGAAACTTTTACGCGGATCAGTTGGCAAAAAGGA
GTATCCTCCTCTCAGCTGGAAAACATGCTCAAGAGTTGCTTTATTAGCTGCTAGTA
AACCTGGCTGCAATGCTCTGATTATTATACGCTACAAACCAGAGATAAAGG
ACTTTTTTTTAAAGGCTCTGTTAATCTCGTCTGATGATCCTGTGATGATG
AAGCACAAGACAATGCTAAAGAAACGAGCTGAAGAAGAGAGGGCTGAGTTTTATA
CCTCTTCTGATCACCCTGGATTTATGGCTGCAAAAGGCACTTATGGGTGAG
TCACCTCGTCAGCTGAAGGCTCCGATTCACTGCTATATTTGCTGAACTGGGATG
TGTTTAATAGCTGGCCATCTGCTGCTTATATGTGAAGAGGTTGGCTCCTCGGAGGTAAG
ATCATGGTGTGAGAACAGCTGATGCAAAATGCAATGGCATTGCTGATCAC
CUCUMUS_SATIVA

CTTTCGCCCTCCGCAAAAGGTATCGCGGAGTTCTCTGCACTTCCCGCCTGGCT
ATCTTGGCGCTCCCGGACCCTCCGTCTGCTAATGCGCAGACG---
ATGGGATATTTGACTAGAATATTGGGTTCGAAGGTTTATGATGTGGCTATTGAAT
TGCTGTATTGCTATGCGATTTACTACACCTGAAAATTAGTGGCAATACCTGGAA
AATTTGGGAGCGACGGTTTGTGTTTGGAGTTCTCATAGTGAGGAGCCAAAAGCA
TATGCTAAAAAGGGAGCGTAGAGGAAGGGCGCACATCATCCTCCCTTCGAT
CACCCTGATGTATACAGCGGCGGATGCAAATGGGGAATTTTGCGCTCAA
ATAAAAGGTCCACGTGACTGCAATTTTTGTACCAGTTGGGCGGTGTCTAAATA
GCTGGATATGCTGCTATGTCAAAAGGGGTGTCCCCTCTTGAAGTAAAGATGGGG
TAGAGCCATCTGATGCTATTTAGCATATTGATGAGGAGAAGCCAAAAGAG
GTTAATATGGACAAAGGGCTTGGGCCTTTCGCTAGTGTGGTACGATTAAAGAGTCG
GTGAAGAGACTTTTGGCGCTAAAGATACCTGAGATGGGTAGTCTTTTGTTAG
CCGTGATGCTATATGTGGTCCTCAAAGGATATGTGGGAGGAGAGAGACAT
CTTGGAAACTCAGGAGGCTTTTCTGCTAGTGTGGTGAGGCTTTAAGATCAATTGGG
TATGGCTAAAAAGGTCCATGTAATGTTATAGTTGATTTGCTGCTTGCTGCTGCT
ATTGAAATTGCTTTTACCTTTTCTGGGAGCTGGGCGCAAGCGGCAAGCTGCAA
TTGGCTTGAAGGGAGAAAATGTTTGTGTAATAACGTGGTAAATATGGAAC
TTTGACAAAACGTAGATTTACTAGAATCTGGCAATGTGGGCTGAGAAGAGAG
GCTTGCTGTGTTGCAACTATTGCGCGGAGACACCTGGAAGTTTTAAAAATTGGTG
AATCTGCTGTGCTTCCATTACCAAGATGATATGAGCCACCTCTCTCCATCTCTCTCTCT
GAAAAAGGCTGTTGCATTTTTATATGTGTTGCTGCTGCTTATCCATAGTTGATG
AGGAAATAGAGAGATAAGGATGAGGATCTCTCGCTGCTCTCGGAAATTGATG
AGAATGATCTTTGAAAGGATCACTTTGCGCTATTGATGGGAGGCAATCAATAG
TCCGAAAAAGCAAGTTTTTGTCTTTTATTTTCCAGAAAGGCGGCGGTGCTTGAAG
GAAGTTTCTAGATGCTCGCTGCAATCTGGGAGCTGTAATCTGGTCTATCACAA
GCAAAGGTGAACACTGGTGGCACAATATATCTAGGGGCGCTGAGATTAAAAACCCG
ACAATGGGCGAATCCATAGAAATCGTGGCTGAGGTGACTGCTGAGTATATTCG
TAAACGTGATGATGTTATCTTTCAGCTTCTCTGATCAT
FRAGARIA_VESCA

CAATCGTGTGCTCCCCAAAAGCGGTATCTACAGTATGAGCCAGTT
TTCTTGGGACGCTACATGACTTGGGAAGGAGGACATTGAG---
CTGGGATACCTTGACGACATATTGTGCACAAAAGTGTGATGTGGCAATCGAA
TCACCTTTGGAGCTGGCAACCAAGCAGTATCTGGCGTGGGAAGGACCTTTTG
ATCAAAAGGGGAAGATCTGCGAGCGGGGTGTTTTCGGTAAGATCTGGGAGCTTAT
AATATGATGCTAAAGATCTCAGGAAAGGCAATGTGACAGGGGTTATATGCTG
TCTGCTGAGAATCAGCAAGGGAGATGGGACATGTGCGCAGAAAATTGCGTTGT
GATGCAGTCTTGGGAATATGGTCTGATACACAGGAAATTAAGTGGCAATCTGTT
GAGAGATTTGGGAGCAACAGTTGTACTTTATAGTGATTTCATTACGATGAGGTCTGAA
GCAATATGCTAAGAGCGGGCTAAGAGAGGGAGGGACGGACCTTCATACCACCCTTT
GATCACCAGACATTTATGTGGAATGGGAATGAGGTTATGCGT
AAAATGAAAGACCTTGCTAAATGGTGTCAATTTCGTACTTTTGGGAGGTGAGTCAA
ATAGCTGTATTGCTGTATAGTAAGAGGCTGTCTCCAGAGTAAAGATTATTT
GATG TGAGCTGTAGCTAGATGTCACTTCTCTGTTCAATTAAAGGATATGTTTGAGGGAGAC
GAAATATTTGAGCAGAGGCTCTTTGTTCTTGGGAGGTGAGTCAA
GTATTTAGCTATAAAGAGGCTCTTTGTTCTTGGGAGGTGAGTCAA
GACTTTGATATAACTAGCTAGTTGAGGCTAGCTGAGATGACAGTTTGGGAG
GAATAATGGGCTTATCAGGATGTTGGTGTTCACACAGTTTCTGAACTAGG
GAGCAATGCAAGAGCGATTGGAGTCTTCTGAACTCAGAACATTCAATCTCACGA
AAAGTGACTTGGCCTAAAGATCACCTTCTTCTTCTTGTAGGTAGGCGAGACAGATG
TTGAAATGAGGCTTCTTCTTGTAGGCTTTTCTCTCCAAGAAAACCTGGACTGATTTGAG
AAGTTCTGTGAATTGCTTTGATTTGAGGGAGACAGGGTGAAACTGGGGCAAATGTTGGTACG
GCAAATGGGTACGGTCTTCAATTCGCAACAGAGCACTGCTCTTCGCAACAGAGCACTGC
CTCCATATCCCGCAAAAGAAGTGCTCTCTGCAAAATTCTCTCTGCAATACCCGTCCTCGGT
ACTTTGGTGCCCCGGAACCTAGACATGATAGGCGGATAAAATG---
ATGGAGTATTTGACGAAATATACTGCTGTCAAAAAGGTTGATGACGGTGATTTGAAAT
CCTCTTTGGCAGCCATTGGCGAGATGAGTATGACAGATGACAGGGTTGGCGCT
CAAGGAGAAGAAGATTGCGCCCGTCTTCTTCATTAAAGCCTGGGATAGAAGCTAAA
ATGATGGCAAACCTCCAAAAGGAACAGCCGGTAAAGAGGAATCTACATGCTCTTGGATAC
GCTGGAATCATGCGGCTCCCAAGAGGTGGTGATGCCGCAAAGAGAACCTCTTTTCGGCAC
GCTGTGATTGCAATGGCTTTACTACACCTGAAATTAGGTCATACTGGGAGCTGAG
GGTTGGTGGCAGACATGGGTCTTATAGGGGATTTGAATGCAAGGACCAACAGAAT
ATGTAAATAGGCGCCAAAAGAGAGAGTGATGCCGACATCATCACCTTTTTGCAT
ACCCAGATGTATCATGGACAGAGGAACAGTGTGTTGGGATAGAAATGTACGTTCAA
TGCAAGGATCTGGTCAATCTTTTGTGCTTTGCTGGTGCGGCGCTTAAATAGCG
TGTTAGTTGTGCATATGTAAGAGGCTTTATCCGGAAGAGATCATTGGTGTG
GAACCCCTCTGCAATGCTATGCGATTTTTGTCACATCATATGGCGACTGAGAGT
ATGCTGAGATCAGGTTGGAGGTGGTCTTATGCAAGAGTTGGAGAAGAAAATCTCCGGCTTATGGAAGGAAATTTGTTGAGATGTTGAGTATTCCGTGAGC
GAATTAGATGAGTTTCGGTATCATGCGAACGATCTTGGATACGAGTATGAACTTGAGACCGATAATGAAGCATTCCAGCTTTTAATGCGT
IOCHRO_TD1
CCTTCGGTTCCCTTACTACGAGTATCTCCGAGCTCGTGGCGAGTGGTAACCGGGTTATTTGATACCGAATACTCCGGTTTTAGGTAGTGGTGGTTTGATTGGGTACGAGTATTGTCATGCGCTTTTGACGAATATATTGTCGTCAAAAGTGTATGATGTAGCTTATGAGACGCCTTTGCAGAATGCACCTAAATTGTCTGAAAGGCTTGGAGTTAAAGTGTGGCTCAAAAGAGAGGATCTTCAGCCTGTCTTCTCGTTCAAAATCAGAGGAGCTTACAATATGATGGCAAAACTCCCTAAAGAGCAGTTGGAAAAAGGGGTTATATGCTCATCAGCTGGGAATCATGCACAAGGTGTTGCATTATCCGCTCAGAGACTTGGCTGTAATGCTGTGATTGTTCATGCCTGTTACTACACCAGACATCAAATGGAAATCAGTTAAGAGATTGGGTGCTACTGTTGTTCTGGTGGGGGACTCATATGATGAAGCTCAAGCATACGCCAAGAGCGGGCCGAAGCTGAATGTCGCACATTCATCCCTCCTTTTGACTACCAGATGTCATTACAGGGCAAGGTACAGTAGGGATGGAGATAAATCGCCAACTCAAA
GATAACATTCATGCAATCTTTTGTGTTGCCCTGTGGAAGAGGAGCTTTATAGCTGTGAATTGCTTCTTATTTGAAAAGGGTTGCCCCTGATATAAAAGATTATTGGAGTTGAAACACTTGATGCAAATGCATTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCCTGGACCAAGTTGGGGGGTTTGCAGATGGTGTAGCTGTAAAAGTGGTTGGCGAAGAAACCTATCGTCTCTGCAAGGAATTAATAGATGGTGTAATCTCTAGTGGTCTGAT
GCTATAGTCGATCTCTATAAAAGGACAGTATTGGAAGAGAGGATGAGGCTGCACTCGTTCTGCTTGGAGGAGGGGGTCTTATAGCTGGTATGGCTTCTTATTTGAAAAGGGTTGCCCCTGATATAAAGATTATTGGAGTTGAACCTTGATGCAAATGCATTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCCAGGACTCTTCAGCGCCAGTGGTAAGTGCTCCAGTATCCTTACAGGACAAGCCTGGGAAGTTGATCGAGAATAATCCGACAGGTGGCGATAGCGAAGAAATG--
CTTCAGTACTTGCTTGACATATTGGCTTCGAGAGTGTACGATGTAGCGATCGATTCCCCTTTGCAGCGCGCAACGAAGCTTTCGAAAAAGCTTGGAGTTAACTTTTTTATAAAAGAGAGGATATGCAGTCTGTGTTTTCATTCAAGCTTAGAGGAGCTTACAATTGATGACCCAACTCACAAAGGAGCAATTAGATAAAGGGGTTATAACAGCAT
CAGACTTACAGTTGTAACCTGAACTCGCACATGTTTGGTGAACAGGGCGGAGGCAGCTCTTTTATGGTCAATATGAGCTTTTGTGTCGATTCACTTTTCCTGAGAGGCCCGGTGCTTGTGATGAAGTTTTGGATGCTTTCAGCCCATGTTGGAATATAAGTTTGTTTCATTATCGTGCACAGGAGATACTGGTGCGAATGTGCTTGTCGGGATCCAAGTTCCACAGGATGAGGTGAAATGAGTTTCAGGAACAGGCAAACACTCTTGGTTATGAATATGCAGTGGAGAGTCATCAATGAGGCATTCCAGCTCATAATGCAT
IOCHRO_TD2
GACTCTTCAAGCGCCAGTCTTTTCATCATGCGAACGATCTTGGATACGAGTATGAACTTGAGACCGATAATGAAGCATTCCAGCTTTTAATGCGT
GAGACCAGATAATGAAGCATTCCAGCTTTTAATGCGT
IOCHRO_TD1
CCTTCGGTTCCCTTACTACGAGTATCTCCGAGCTCGTGGCGAGTGGTAACCGGGTTATTTGATACCGAATACTCCGGTTTTAGGTAGTGGTGGTTTGATTGGGTACGAGTATTGTCATGCGCTTTTGACGAATATATTGTCGTCAAAAGTGTATGATGTAGCTTATGAGACGCCTTTGCAGAATGCACCTAAATTGTCTGAAAGGCTTGGAGTTAAAGTGTGGCTCAAAAGAGAGGATCTTCAGCCTGTCTTCTCGTTCAAAATCAGAGGAGCTTACAATATGATGGCAAAACTCCCTAAAGAGCAGTTGGAAAAAGGGGTTATATGCTCATCAGCTGGGAATCATGCACAAGGTGTTGCATTATCCGCTCAGAGACTTGGCTGTAATGCTGTGATTGTTCATGCCTGTTACTACACCAGACATCAAATGGAAATCAGTTAAGAGATTGGGTGCTACTGTTGTTCTGGTGGGGGACTCATATGATGAAGCTCAAGCATACGCCAAGAGCGGGCCGAAGCTGAATGTCGCACATTCATCCCTCCTTTTGACTACCAGATGTCATTACAGGGCAAGGTACAGTAGGGATGGAGATAAATCGCCAACTCAAA
GATAACATTCATGCAATCTTTTGTGTTGCCCTGTGGAAGAGGAGCTTTATAGCTGTGAATTGCTTCTTATTTGAAAAGGGTTGCCCCTGATATAAAGATTATTGGAGTTGAACACTTGATGCAAATGCATTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCCAGGACTCTTCAGCGCCAGTGGTAAGTGCTCCAGTATCCTTACAGGACAAGCCTGGGAAGTTGATCGAGAATAATCCGACAGGTGGCGATAGCGAAGAAATG--
CTTCAGTACTTGCTTGACATATTGGCTTCGAGAGTGTACGATGTAGCGATCGATTCCCCTTTGCAGCGCGCAACGAAGCTTTCGAAAAAGCTTGGAGTTAACTTTTTTATAAAAGAGAGGATATGCAGTCTGTGTTTTCATTCAAGCTTAGAGGAGCTTACAATTGATGACCCAACTCACAAAGGAGCAATTAGATAAAGGGGTTATAACAGCAT
CAGACTTACAGTTGTAACCTGAACTCGCACATGTTTGGTGAACAGGGCGGAGGCAGCTCTTTTATGGTCAATATGAGCTTTTGTGTCGATTCACTTTTCCTGAGAGGCCCGGTGCTTGTGATGAAGTTTTGGATGCTTTCAGCCCATGTTGGAATATAAGTTTGTTTCATTATCGTGCACAGGAGATACTGGTGCGAATGTGCTTGTCGGGATCCAAGTTCCACAGGATGAGGTGAAATGAGTTTCAGGAACAGGCAAACACTCTTGGTTATGAATATGCAGTGGAGAGTCATCAATGAGGCATTCCAGCTCATAATGCAT
IOCHRO_TD2
GACTCTTCAAGCGCCAGTCTTTTCATCATGCGAACGATCTTGGATACGAGTATGAACTTGAGACCGATAATGAAGCAGCTTGGGTTATGAATATGCAGTGGAGAGTCATCAATGAGGCATTCCAGCTCATAATGCAT
CAGCTGGAAATCATGCACAAGGTGTTGCATTAGGTGCTCAGAGACTTAATTGTA
AAGCTAAAAATTGCTAGCCTGTTACTACCCAGAGATCAAGATGGAACCTGTA
AAAACCTTGGATGCTACAGTAGTTCTTGGAGGGTACACATTTGATGAAGCTCAAG
CACATGCACTCACATTAGTCAAACGGAAGGCTCACAATTTGACATCCCAGATCATCG
ATCATTGACATGTCATCAAAAGGCAAGGTACAATTGCAACAGAGGTCAACCGTCACT
AATTTAAGGAC---
ATACATTCAAGTTTTGATCCCCCTGTTATCTACCCAGAGCATATTTGGAGGAAATAT
GTGCAAGTGCTACAGTAGTTCTTGGAGGGTACACATTTGATGAAGCTCAAG
CACATGCACTCACATTAGTCAAACGGAAGGCTCACAATTTGACATCCCAGATCATCG
ATCATTGACATGTCATCAAAAGGCAAGGTACAATTGCAACAGAGGTCAACCGTCACT
AATTTAAGGAC---
CCTTTCTTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAGGCTCTGGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
CTTTCCCTTGCAGGAGCTGAAGCATACTGCAAGTACTACAACCTCAAGGACCAATG
AACATAGTAGTAAATAACCGAGTGCTACTACCAACCCTACAGGAGCAGTAACTG
TTGGCAAGTCTACTGGAATACATTTTGATGGAAGCTTGAATTTTATTTTCTCTCTT
ATGACACTCTCTCAATTTCCATACATTGATCTGACAGATATGACCTTGTCAAAGATCATCTACGCCATTTGATTGGTGGCCGATCTAGTGTACAAGATGAACTTCTTTGCCGGTTTGTTTTCCCAGAGAGGCCTGGTGCCCTAATGAAATTTTTAGACGCTTTCAGCCCACGATGGAATATAAGTTTGTTCCATTACGTTTCACAGGGAGAAGCTGGTGCAAATGTGCTGGTCGGTCTTCAAGTTGCCAGCAACGAGATAGATGAGTTTAAGCAACGAGCCAACAATCTGGGTTATGAGTATGCTGTGGAAACCTGTAACGAGGCTTTCCAACTCTTGATGCAT

I_LAC_1

ACGCGGCGCTGACTTGCCGGGTAGTGCTCTCCGCCCTCTTGACCAGTGAGCCTGGATGCGCTTGTTATTCCGAATGAGCCCTATAGCTACAAAGCT----------------

GTAGAGTATTTGACGCGGCTGACTTGCCGGTAGTGTCTCCGGCCTCCTTGACCAGTGAGCCTGGATGCGCTTGTTATTCCGAATGAGCCCTATAGCTACAAAGCT----------------

CTCAAAAAAGGGAGGAGATCTCCAACCTGTATTCTCTTCAAACCTCGAGAGCTTTATAATATGATGCTCAAACCTCCAAGGGAGGTATTTGAGATAAAAGGAGATGCTTTCTCCATCATGCTGGAAACCTGTAACGAGGCTTTCCAACTCTTGATGCAT

I_LAC_2

CCGCCGGCTGAACTGCCGCGAGTGTCTCCGGCCTCCTTGACCAGTGAGCCTGGATGCGCTTGTTATTCCGAATGAGCCCTATAGCTACAAAGCT----------------

CTCAAAAAAGGGAGGAGATCTCCAACCTGTATTCTCTTCAAACCTCGAGAGCTTTATAATATGATGCTCAAACCTCCAAGGGAGGTATTTGAGATAAAAGGAGATGCTTTCTCCATCATGCTGGAAACCTGTAACGAGGCTTTCCAACTCTTGATGCAT
CTACTGGTGCAAATGTGTTAGTCGCGCATCCAAGTCTGAGGTTGATG
AGTTCCGAGGACGGCTGACAGTCTTGGTTATGAATATGCAGTGGAGAGTCTCA
ACGAGGCTTCCGACTCTACAAGCAT
L_BARB_TD1
CCTTCGGCTCCGGTTATATTACGAGTTTTCTTCGAGCGAGTTGCAAGCCGGTT
ACTTGATACCAAAAACTCGCTAGGGGACAGTTGAGAAAAGG
GGTTATATGTCATACGCTGGAACCATGCAAAAGCTTGAGTTCGATTATCCGCTCAG
AGACTCCGGCTGCTGATGCTGTTGTGC leaksCCCTGCTTACTACCCGACACTAAAT
GGAAATCAGTTAAAGAGATTTGGAGTCAAACCTTGAGTTGAGACCTACAGC
ATGAAGCTCAAGCATACGCCAAGAGCGGGCCAAGCTGAGATATAATGCTGAGGCC
AAGTTGAGGAGAGGTCTTTAATAGCTGTAGTGGC-
GCTTATTTGAGAGGTTGC
TTGGAGGAGGAGGTCTTATAGCTGGTATTGC-
GCTTATTTGAAGAGGGTTGCCCCTGATATAAAAGATTATTGGAGTGAACCCCTTG
ATGCAAATGCAATGGCATTATCATTGCAGCATGGTCAGAGAGTAATGCTGGACC
AAGTTGGGGGTTTTCAGAGTAGCTGGTTAAAGTGGTT------------------------

J_PROCUMB_TD1
CCGTCAAGCTCCGGTTGCTACGAGTTTTCTCCAGAGCTGTTGCAAGCTGCAACCGGGAC
ATTTGATAACGAATGTCGCCGTTTTAGGATGTGTGGTAAGAGATCGGCTAGGAT
ATTAAACAAACAATCTTGTGTCAAAAAGTGTATGATGCTTATAGCTGAGACCGCTT
ACAGAAACGCGCCTAAGACTGTAAGGAGCTGGGAGTTAAAGTGGCTACTAGGCTCAG
AAGAGGATCTTACGCTGTTTTTCTCCTGCCTTTG------------------------
GAAAAGGAGCATACTAGAGCAGCTGCAGGTGCACTTGCTCTTGCTGGAGCCGAGGC
ATACTGCAAGTATTAGCGCTCAAGGTTGAAAAAGTGATGAGCAGCAAATACATG
AGCCAACATGAACTTTGACACGCTTCTTGACTTTACATGTTGGACTTCAACAG
ATTTTGAAGTGAAGGATTGGAGGAGTGGAAACTCAAGCTGTAATTGACAGC
APAAAAGTCCGCTGAAATGGTAGGACCAAATGCAATCATCAGCTGAAAGATTG
GTTGAACAAATGTTATGAAACATGATGAGATTGAACCTACCCAGAGAGAT
CTGGCTCTTTGATGAGATTTTTAGATGTCCGGCAGCCATGATAGTTG
TTTCAATTATCTGCAGCAGGAGATCAGTCTGTGCAAAATGTGCTCTGAGG
CTCCAGAGGAGGAGTTGTAAGTTATGAGCAGCTTCTGGCTTATGGAAG
TTTGATACCGAATAGTCCGGTTCTAGGTAGTGTTTGTGACGGATGA
TTGACAACCAACATCTGTCGCAAGGATTGATGAGATTGAACCTACCCAGAG
AGAAAAGCAGCTCTAGTACTTACAGTGTTTGTGCTTCACACAACATTAG
GTTGAACAAATGTTATGAAACATGATGAGATTGAACCTACCCAGAGAGAT
CTGGCTCTTTGATGAGATTTTTAGATGTCCGGCAGCCATGATAGTTG
TTTCAATTATCTGCAGCAGGAGATCAGTCTGTGCAAAATGTGCTCTGAGG
CTCCAGAGGAGGAGTTGTAAGTTATGAGCAGCTTCTGGCTTATGGAAG
TTTGATACCGAATAGTCCGGTTCTAGGTAGTGTTTGTGACGGATGA
TTGACAACCAACATCTGTCGCAAGGATTGATGAGATTGAACCTACCCAGAG
AGAAAAGCAGCTCTAGTACTTACAGTGTTTGTGCTTCACACAACATTAG
GTTGAACAAATGTTATGAAACATGATGAGATTGAACCTACCCAGAGAGAT
CTGGCTCTTTGATGAGATTTTTAGATGTCCGGCAGCCATGATAGTTG
TTTCAATTATCTGCAGCAGGAGATCAGTCTGTGCAAAATGTGCTCTGAGG
CTCCAGAGGAGGAGTTGTAAGTTATGAGCAGCTTCTGGCTTATGGAAG
TTTGATACCGAATAGTCCGGTTCTAGGTAGTGTTTGTGACGGATGA
TTGACAACCAACATCTGTCGCAAGGATTGATGAGATTGAACCTACCCAGAG
AGAAAAGCAGCTCTAGTACTTACAGTGTTTGTGCTTCACACAACATTAG
GTTGAACAAATGTTATGAAACATGATGAGATTGAACCTACCCAGAGAGAT
CTGGCTCTTTGATGAGATTTTTAGATGTCCGGCAGCCATGATAGTTG
TTTCAATTATCTGCAGCAGGAGATCAGTCTGTGCAAAATGTGCTCTGAGG
CTCCAGAGGAGGAGTTGTAAGTTATGAGCAGCTTCTGGCTTATGGAAG
TTTGATACCGAATAGTCCGGTTCTAGGTAGTGTTTGTGACGGATGA
TTGACAACCAACATCTGTCGCAAGGATTGATGAGATTGAACCTACCCAGAG
AGAAAAGCAGCTCTAGTACTTACAGTGTTTGTGCTTCACACAACATTAG
GTTTTTAGATGCCTTCAGCCCACGTTGGAATATAAGTTTGTTTCATTATCGTGCACAGGGAGATACTGGTGCAAATGTGCTAGTCGGGATCCAAGTTCCACCGGATGAGGTAGTTGAGTTGAGGGACGGGCAGACAGTCTTGGTTATGAATATGCAATGGAGAGTCTCAATGAGGCATTCCAGCTCATCAATATGCAT

S_LYCO_TD1
GCACCGGTTCGGTGTGACGATTCGCTCCAGTCGTGAGCTGAGCGTGCAC

S_MELO_TD1
GGTAAGGTTCCGTCGGTGCGAGTTTCTCCCAGCTCGTTGCAGTGTGAGCCGGGAT

72
GAGGATCTTCAGCCGGTCTTCTCATTCAAAATCAGAGGAGCTTACAATATGATGGCAAAACTCTCTAAAGAGCAGTTGGAAAAAGGGGTAATATGCTCATCAGCTGGAAACCATGCACAAGGTGTTGCATTATCCGCTCAGAGACTTGGCTGCGATGCTGTGATTGCCATGCCTGTTACTACGCCAGACATCAAATGGAAATCAGTTAAGAGATTGGGTGCTACTGTTGTTCTTGTTGGGGACTCATATGATGAAGCTCAAGCATATGCCAAAAGCGGGCTGTAGCTGAAGGCCGCACATTCATCCCTCCTTTGATCACCCAATGTCATCGCCGGGCAAGGTACAGTAGGGATGGAGATAAATCGCCAACTCAAAGATAACATTTCATGCAATCTTTGTGCCTGTTGGAGGAGGGGGCTTATAGCTGGTATTGCTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAACTTATCGTATCTGTGAGGAATTAATAGATGGCGTAGTCCTAGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAGAAGCTGTTGCTTATTTGAAAAGAGTTGCCCCCGATATAAAGGTTATTGGAGTTGAACCACTTGATGCAAATGCGTTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGTGGTTTTGCAGATGGTGTAGCTGTTA
AAGAGCTCTAGTACTTTACAGTGTTGGTCTTCACATGAAACTAGAACTTGAAGAATGGTGGACAGGATGAAATCAGCAGATCTGAAAACCATTAATCTTACAGACAATGACTTGGTCAAAGATCATCTAAGGCATTTGATGGGTGGTAGAACAAATGTTCATAATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAGCCCGGTGCTTTGATGAAAGTTTTAGATGCTTTCAGCCCATGTTGGAATATAAGTTTGTTTCATTATCGTGCACAGGGAGACACTGGTGCAAATGTGCTAGTTGGGATCCAAGTTCCAAAGGATGAGGTAGTTGAGTTCCAGAGACGGGCAGACAGTCTTGGTTATGAATATGCAGTGGAGAGTCTCAATGAGGCATACCAGCTCATAATGCAT

CCATCGGCTCCGTTCTGCTACGAGTTTTTCTCCGAGCTGCGTTGCAGTGCGACCCGGGGTACTTGATACCGAATAGTCCGGTTCTAGGTACTGGTGTTGGGTACGAGTTTTGACAAACATCTTGTCGTCAAAAGTGTATGATGTAGCTTATGAGACGCCTTTGCTGCGCAATGAAGGATCTTCAGCCCGTCTTCTCGTTCAAAATCAGAGGAGCTTACAATATGATG

GCAAAACTACCAAAAGTGAGTTGGGACTCTATATGAGCTCAAGCATACGCCAACAGAGGTTCCAGGGACGAGCAGACAGTCTTGGTTATGAATATGCAGTGGAGAGTCTCAATGAGGCATTCCAGCTCATAATGCAT

GCTATATGTGCATCTATATAAGGACATGTGTTGGAAGAAGAGGAAGCAACTCTTACGGCTGCTTATTTGAAAAGGGTTGCCCTGATATAAAGATTATTGGAGTTGAACTTGATGCAAATGCATTGGCATTATCATTACAGCATGGTCAGAGAGTAATGCTGGACCAAGTTGGGGTCTTTGCAGATGGTGTAGCTGTTAAAGTGGTTGGTGAAGAATTTATCGTCTCTGCAAGGAATTGATAGATGGTGTAGTCCTCGTTGGTCGTGATTGCTATTATGTGCATCTATAAAGGACATGTTTGAAGAGAAGAGAAGCATACTTGAGCCTGCAGGTGCACTTGCTCTTGCTGGAGCTGAGGCATACTGCAAGTACTATGGCCCTCAAGGATGAAAACGTCGTAGCATTACTAGTGGAGCCAACATGAACTTTGACAAGACTTAGATTGGTAACTGAACTCGCAGATGTTGGTAGACAGCGGGAAGCTGTTC

TATGACTTTATGCCAGAAAAAACCGGGAGCTTAAAAATTTGGTAAATGGTAAGGACCAATGAATATCACTGAATTCAAGTACAGATAATTCTGATAAAAGAAAAGACTCTAGTACTTTACAGTGTTGGTCTTCACACAAAAATTGAACCTGAAGAAATGGTGGAAGAGTAAATCAGCAGATCTGAAAACCATTAATCTTACAGACAATGACTTGGTCAAAGATCATCTAAGGCATTTGATGGGTGGTAGAACAAATGTTCATAATGAGCTTTTGTGTCGATTCACTTTTCCCGAGAAGCCCGGTGCTTTGATGAAAGTTTTAGATGCTTTCAGCCCATGTTGGAATATAAGTTTGTTTCATTATCGTGCACAGGGAGACACTGGTGCAAATGTGCTAGTTGGGATCCAAGTTCCAAAGGATGAGGTAGTTGAGTTCCAGAGACGGGCAGACAGTCTTGGTTATGAATATGCAGTGGAGAGTCTCAATGAGGCATTCCAGCTCATAATGCAT
N_ATTEN_TD1

N_BENT_TD1
78

GCCTGCAGGTGCACTTGCTTTGGAGCTGAGGCAATACTGCAAGTACTATGGCCTCAAGGGTGAAAACATAGTAGCAATAACTAGTGAGCCAACATGAACTTTGACAGACTTAGATTGGTAACCGAACTCGCAGATGTTGGTAGACAACGGGAAGCTGTTCTTGCTACTTTTATGGCAGAAAAGCCAGGTAGCTTCAAGAAGTTCTCTGAAATGGTAGGGCCAATGAATATCACTGAATTCAAGTACAGATACAATTCTGGGAAAGAAAAAGCTCTTGTGCTTTACAGTGTTGGTCTTCACCAAAATTAGAACTTGAAGAAATGGTGGAGAGGATGAAATATGCAAATCTGCAAACCATTAATCTTACAGATAATGACTTGGTCAAAGATCATCTAAGGCATTTGATGGGTGGCAGAACAAATGTCCTCAATGAGCTTCTGTGTCGATTCACTTTTCCTGAGAAGCCCGGTGCGTTGATGAAGTTTTAGATGCTTTCAGCCCACGTTGGAATATAAGTTTGTTTCATTACCGTGCACAGGGAGATACTGGTGCAAATGTGCTAGTCGGCATCCAAGTTCCACAGGCTGAGGTCGATGAGTTCCAGGGACGGGCTGACAGTCTTGGTTATGAATATGCAGTGGAGAGTCTCAATGAGGCTTTCCAGCTCATAATGCAG

N_TABA_TD1A

CCTTCGGTCCGTTGCTACGAGGTTTTCGCCGAGCTCGTTGCAGTGCGAACCGGGATACCTAATACCGAATAGTCGGTTTAGGTACCGGTTGGAAGATGTTGTTACGAGTATTGAGAAATATCGCTTTATCTGCTACGAGCCACGCTCAAGACCTGTGTTGATGTTTCGTTACTACACACAGGATATTAAGTGAAGGTTGAGGATCTTCAGCCTGTGTTGTTTATAGGATGGAAGAGATGCTGGCAAGATGCTGGACCAAGTTGGGGTTCGCAGATGGTGACTTGTTAAAGTGGTCGGTGAAGAAACTTATCGTCTCTGCAAGGAATTAATAGATGGCGTGGTCCTTGTTGGTCGTGATGCTATATGTGCATCTATAAAGGACATGTTTGAAGAAAAGAAGCATACTAGAGCCTGCAGGTGCACTTGCTCTTTGGAGCTGAGGCAATACTGCAAGTACTATGGCCTCAAGGGTGAAAACATAGTAGCAATAACTAGTGAGCCAACATGAACTTTGACAGACTTAGATTGGTAACCGAACTCGCAGATGTTGGTAGACAACGGGAAGCTGTTCTTGCTACTTTTATGGCAGAAAAGCCAGGTAGCTTCAAGAAGTTCTCTGAAATGGTAGGGCCAATGAATATCACTGAATTCAAGTACAGATACAATTCTGGGAAAGAAAAAGCTCTTGTGCTTTACAGTGTTGGTCTTCACCAAAATTAGAACTTGAAGAAATGGTGGAGAGGATGAAATATGCAAATCTGCAAACCATTAATCTTACAGATAATGACTTGGTCAAAGATCATCTAAGGCATTTGATGGGTGGCAGAACAAATGTCCTCAATGAGCTTCTGTGTCGATTCACTTTTCCTGAGAAGCCCGGTGCGTTGATGAAGTTTTAGATGCTTTCAGCCCACGTTGGAATATAAGTTTGTTTCATTACCGTGCACAGGGAGATACTGGTGCAAATGTGCTAGTCGGCATCCAAGTTCCACAGGCTGAGGTCGATGAGTTCCAGGGACGGGCTGACAGTCTTGGTTATGAATATGCAGTGGAGAGTCTCAATGAGGCTTTCCAGCTCATAATGCAG
GCCCGGTGCTTTTAGAAGTTTTTAAAATGTCTTTCAGCCCACGTTGGAATATAAGTTTGTTTCATTACCGTGCACAGGGAGACACTGGTGCAAATGTGCTAGTTGGCATCC
AGAAATATGCAATGGAGAGTCTCAATGAGGCTTTCCAGCTCAAATGAGGGCTCAACAGAGAGGATCTTCAGCC-
N_TABA_TD1B
CCTTCGTTCCGTTGCTACGAGTTTTCCGCGAGCTCCTTGACGTGCAAGACGGGAGT
ATTTAAACTCCAATAGTCCCGGGTTTATTAGGCCGCTGTAAGTTGTTACAGAGTA
TTTGACGAACATCTTGTGCTTAAAGTGTATGAGCTGCTTTAAGGAGACTCCTTGG
CAGAAACGCGCCTAAAAGTGCTGACAGAGAGATTGGGAGTTAATGTGTGGCTCAAAG
AGAGGATCTTCAGCC-
GTCTTTCTCGTTCAAGCTCGAGAGCTTTCAACAAATATGATGCGA AAAAACTTTCTAAGG
AGCAGTTGGAAGAGGGTTATATGCTCATCAGCTGGAAGATCATCGGACAAGGGTGA
TTGCAGTATCTCATACACCTTTTGATACCCCTGTAGTCATCATAGAGACA
AGTACAGTGGGAATGGAGATATGCAAGATGGTTGCTCAGGCTTTAAGGGTTTCTG
AGTCGAGATGCTAGTGCTTCAAGCTCAGAGGAGCTTACAATATGATGGCAAACTTTCTAAGG
AGAGGATCTTCAGCC-
N_TOME_TD1
CCTTCGTTCCGTTGCTACGAGTTTTCCGCGAGCTCCTTGACGTGCAAGACGGGAGT
ATTTAAACTCCAATAGTCCCGGGTTTATTAGGCCGCTGTAAGTTGTTACAGAGTA
TTTGACGAACATCTTGTGCTTAAAGTGTATGAGCTGCTTTAAGGAGACTCCTTGG
CAGAAACGCGCCTAAAAGTGCTGACAGAGAGATTGGGAGTTAATGTGTGGCTCAAAG
AGAGGATCTTCAGCC-
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
N_TOME_TD1
CCTTCGTTCCGTTGCTACGAGTTTTCCGCGAGCTCCTTGACGTGCAAGACGGGAGT
ACTTTAAACTCCAATAGTCCCGGGTTTATTAGGCCGCTGTAAGTTGTTACAGAGTA
TTTGACGAACATCTTGTGCTTAAAGTGTATGAGCTGCTTTAAGGAGACTCCTTGG
CAGAAACGCGCCTAAAAGTGCTGACAGAGAGATTGGGAGTTAATGTGTGGCTCAAAG
AGAGGATCTTCAGCC-
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------
N_TOME_TD1
CCTTCGTTCCGTTGCTACGAGTTTTCCGCGAGCTCCTTGACGTGCAAGACGGGAGT
ACTTTAAACTCCAATAGTCCCGGGTTTATTAGGCCGCTGTAAGTTGTTACAGAGTA
TTTGACGAACATCTTGTGCTTAAAGTGTATGAGCTGCTTTAAGGAGACTCCTTGG
CAGAAACGCGCCTAAAAGTGCTGACAGAGAGATTGGGAGTTAATGTGTGGCTCAAAG
AGAGGATCTTCAGCC--
ATGGAAATCAGTTGAGAGATTGGGTGCTACTGTTATTCGCAAGGGGGACACATA
TGATGAGGCTCAAGCATAATGCGGAAAGCTCACGGCCAAAGACAGAGGCCGCACT
TCGTCCTCCTTTTATATCCCCCCGTTATATGAGGCTCAAGCATATGCGAAAGCTCAGGCCAAAGCAGAAGGCCGCACAT
TCGTCCCTCCTTTTATATCCCCCCGTTATATGAGGCTCAAGCATATGCGAAAGCTCAGGCCAAAGCAGAAGGCCGCACAT
ACCATGTTGCAAAGAATTTGATGTTGAACATGATGTTGCAATATTTGTACCTGCTCTTGAGCTAGGGCAG
CTATTAAGTATGACAGGAAAGAACCTCCTGGTCTTCTGCAAGGAAATGATTAGATG
GGGTTGGTTAGTTGGTCTGATGCTATATGTGACATCGATAAAAGGAGATCGTCTG
AGAGAAAAGCATGGAAGCTTGCAAGGGTAAACATGTAAGCAATAACTA
GTGAGCCCAATATGAATTTGTGACAGACCTTACAGGTGAATGCTAATGTGCAACTTAGCAGGATG
TTGGTAGGGCGCTGGAAGTCTGGTTCTGCTACCTATATGCCCCGAAACGAGCCAGG
GCTTCAGAAAGCTTCTATATGATGGTGGGAAAGGATGAATTTACTGAAATTA
ACAGATATAATTCTGGAAGGAATATTGGGAAGACAGATCATCACAACCTG
CATACCATATAATTTTACAGATACAGGATAATGCTAAAGATCTCTGCTAGCT
ATGGGTTGTCAATCCCTTCTGCAAGCTTGTGACCTCGTTCTGTCATCATTCCCGA
GAGGGCTGGTCTTGTAGAAAGTTTTTATGATGCTTTTAGCCCTGTCGTAATATT
GGTTTCTCCTACCCGGGCAACGGGGAAACTGGTTGCAATATCTTACGGCAAG
TCCAGGATCTTACCATATGATGGCAAACCTCCTCAGAGGAGCAGTTAGAAAAG
GGGTTATATGTGTTACATCGTTGAAGGAAACTCGTGGAAGGAGCTTGTGG
AAGAGACCTTGGCTGGAATGTCATGTGATGCTTATGAGGGAAGTCTGCTGATC
AAGGGAATATCGTGGAGATTTGGGTCTACTGTGTTGTTGGGGGCAAGCAT
ATGAGAGCAGTCATAGCTGCAATGCGTGCAGTCAATTGGGAAAATGGTGCTAAG
TCTCCCTCCTTTGATACCCAGATATCACAACGGGCAAGGAGCTTGTGG
TGAGATGGAATCTGCGAATAGGTTGTTGCAAGGAAATTTTGCTCTGCAAGGAA
GCTATATGATTGGGAGAGGAAAGTTTCCGGCTCTGTCAGGGAATATTAGATG
GGGTTGGGCTCCTAGTTAACATGCTATATGCTGCTAGTACAACTG

84
CTGATAGGGAGATCAGATGTCCACAAATGAGCATCTCTGTGATTTCATTTTCCCATACAAACCCTGGGTGTTTTGATGAATTTGAAGTTTTTAGACGCTTTTCAGCCCACGTTGGAATAAAGTTTGTTACATTACATCATGCACAGGGAGAAACTGGCGGGAATGTGCTAATCGGACTCCAAGTTCCACAGAATGAGGTTGATGAATTCCGGGTCTCGAGCTGAGAGTCTTGGTTATGTGTGTGCTGTTGAGACTCTCTCAATGAAAGCTTTCCACAAACTGCATACTGCTGATG
S_MELO_TD3
GTTGATTCATTTGTTCCGATAGTTCATCCGAGGTTGCGATTGCAATGCAGGAAGGTTATTTGTTACATTGATACCAGTAGTACATCCGAGTTCTACAATGCGAGGAAGGTATTTGTTACCGAATTATCCGTTGCTACTGTGTGTGTTGTCGAAGGACAGCTTAGAAGATT
TATGAAATCATGCAAATGATGCGGCAAGAATATTGAGTCAGATGCGGAAATTTGTTGCGAATGTGCTGGTCGGACTCCAAGTTCCACAGAATGAGGTTGATGAATTCCGGGTCTCGAGCTGAGAGTCTTGGTTATGTGTGTGCTGTTGAGACTCTCTCAATGAAAGCTTTCCACAAACTGCATACTGCTGATG
S_TUBE_TD3
GATGATACATTGATACCAGATGTTGAGAGATGGAATTCTTCGCAAATTCGATACTGCTGACATCTCTGTGATTTCATTTTCCCATACAAACCCTGGGTGTTTTGATGAATTTGAAGTTTTTAGACGCTTTTCAGCCCACGTTGGAATAAAGTTTGTTACATTACATCATGCACAGGGAGAAACTGGCGGGAATGTGCTAATCGGACTCCAAGTTCCACAGAATGAGGTTGATGAATTCCGGGTCTCGAGCTGAGAGTCTTGGTTATGTGTGTGCTGTTGAGACTCTCTCAATGAAAGCTTTCCACAAACTGCATACTGCTGATG
N_BENT_TD2A

TCTCCATCAGCGTATTCAATGCCCGTGTCCTTTGCAGGGCCAGGGAGCGAGGG
AATGGTAGAGAATAATCCATCAGGAGGCTACAGAGGGAATTG---
TTTCAATCTTCTTGAATTGCTTCTCAGATGTAGTGAATGCAATGATTTG
CCCCTTGCACAATGCAAGCTTCCAAGAAGCTTGGAGTTAAACTTTGGAATC
AAAAGAGAGGATATGCGAGTTCTGTTTATTTAGGTGCTAGAGACCTAAAATG
GCTACGTAGTTGTAGTTCCACACAGGAGATCAAAGAGTGAGGTAAAGAAAG
AACCTGAGGGTAATAGTTTTCTACAGGTAGTGCTCATAGTAAATTCAAC
GCTGCAGAAGGATGAATGCTCAGCAGCTTTCTTGGATC
ACCAGATGTGACATGAGGAATTCAGGTCAGTTAAAAATGGAGATTATCGTCAAC
TTAAAGAT---
ATATATGCGATTTTGTGCTGTAGAAGGGAGAGGGTTAATAGCTGTTTGGCTG
CATATTCAAAAAGGTGGTGCTCTCTACAAACAGATTATTGGAGGACCATTTGG
TGCAAGTTCAATGACACAAATCTTGTACCATAGGAGAAGAGAATAGTTGAACA
AGTTGACAAATTTCGAGATGGCTAGCTTGCTACTAGTTGGAAGAAACTTTGC
GCTTTGTGCAAAGATTTAATAGCGAGAATGTCTTACTAGATGAGCTATTA
GTGCAGCGTAAGAGGATGGTACGACGAAGGAAAGCATATTAGAGACATCA
GGTGCAGCTTGCATGGTAGCTGGAAGTGAATTAGCTTGATTTCCGCT
ACTTTATGCGCAAGAAGACAGGAAAGCTTTCAAAATTTGTGAAAGTTGGGA
CCTATGAAATATTACTGAAATTTAAGTACAGATATAATTCTGGAAGAAAAAAGACT
CTGTGACTCTACAGTTTTGTTAATAAAAATTCGAGATCCAAATTTGG
AGAGGATGAATACTCAACCAGCTAATACTGCTATCCTCACAATTTAATTCTG
TCAAAGAACTCTAAGGCGATTTGGATGCGAGATCAGAAACAAGTAAAGGAG
ATTCTTGCTCATTCTTCCAGAGAAGCTTGAGCTTGAAGAAGGTTCTTACTG
ATGTTTTACGCCCCTCGCTGAAATATAAAGTGGTTTTGTTCTATTATCGAGAACGGAGA
ACTTGAGCAAAGTGTGTTTCTTGAGATTCTATCTCAGTAAATTCTGGAAGATT
GTTCCCGAGTTCCAGAATCAATTCTTTGTTATATCATGAAATACGAGAGATCTCAAC
GAGGCCTTCCAAAATTTAATG---
N_ATTE_TD2

TCTCCATCAGCGTATTCAATGCCCGTGTCCTTTGCAGGGCCAGGGAGCGAGGG
AATGGTAGAGAATAATCCATCAGGAGGCTACAGAGGGAATTG---
TTTCAATCTTCTTGAATTGCTTCTCAGATGTAGTGAATGCAATGATTTG
CCCCTTGCACAATGCAAGCTTCCAAGAAGCTTGGAGTTAAACTTTGGAATC
AAAAGAGAGGATATGCGAGTTCTGTTTATTTAGGTGCTAGAGACCTAAAATG
GCTACGTAGTTGTAGTTCCACACAGGAGATCAAAGAGTGAGGTAAAGAAAG
AACCTGAGGGTAATAGTTTTCTACAGGTAGTGCTCATAGTAAATTCAAC
GCTGCAGAAGGATGAATGCTCAGCAGCTTTCTTGGATC
ACCAGATGTGACATGAGGAATTCAGGTCAGTTAAAAATGGAGATTATCGTCAAC
TTAAAGAT---
ATATATGCGATTTTGTGCTGTAGAAGGGAGAGGGTTAATAGCTGTTTGGCTG
CATATTCAAAAAGGTGGTGCTCTCTACAAACAGATTATTGGAGGACCATTTGG
TGCAAGTTCAATGACACAAATCTTGTACCATAGGAGAAGAGAATAGTTGAACA
AGTTGACAAATTTCGAGATGGCTAGCTTGCTACTAGTTGGAAGAAACTTTGC
GCTTTGTGCAAAGATTTAATAGCGAGAATGTCTTACTAGATGAGCTATTA
GTGCAGCGTAAGAGGATGGTACGACGAAGGAAAGCATATTAGAGACATCA
GGTGCAGCTTGCATGGTAGCTGGAAGTGAATTAGCTTGATTTCCGCT
ACTTTATGCGCAAGAAGACAGGAAAGCTTTCAAAATTTGTGAAAGTTGGGA
CCTATGAAATATTACTGAAATTTAAGTACAGATATAATTCTGGAAGAAAAAAGACT
CTGTGACTCTACAGTTTTGTTAATAAAAATTCGAGATCCAAATTTGG
AGAGGATGAATACTCAACCAGCTAATACTGCTATCCTCACAATTTAATTCTG
TCAAAGAACTCTAAGGCGATTTGGATGCGAGATCAGAAACAAGTAAAGGAG
ATTCTTGCTCATTCTTCCAGAGAAGCTTGAGCTTGAAGAAGGTTCTTACTG
ATGTTTTACGCCCCTCGCTGAAATATAAAGTGGTTTTGTTCTATTATCGAGAACGGAGA
ACTTGAGCAAAGTGTGTTTCTTGAGATTCTATCTCAGTAAATTCTGGAAGATT
GTTCCCGAGTTCCAGAATCAATTCTTTGTTATATCATGAAATACGAGAGATCTCAAC
GAGGCCTTCCAAAATTTAATG---
GCTACGATTGTGATCATGCTTCGATTAACAGAAGGGAACGCAATAGGTTGAGCAAGCATTAAG
AACCTGCTGAGAAAGTACAGCTTGCTAAGTACATGACACATTTGAAGACCTCAGGAA
CATGCTTTAAAGTTGGCTGAAGAGAGTGCATTCACTTCATCCGCCTTTCCATGC
ACCCAGATGTGATCATAGGGAACAGTTGAATTGGAACGGAGATTAATCGTCAAC
TTAAAGAT---
ATACATGCAATTTTGTGCCGTAGGGAGGGAGGTTTTAAATGCTGGTGTTGCTG
CATATTCCAAAGGTTGGTCCTCCATACAAAGATTATGGAGTTGAGCCATTTGG
TGCAAGTTCAATAGCAACAACTCTTTGATCCATAGGGAAGAGAATGAAATTTGGAACA
AGTTGATAATTTTGACATGGCGATGCTGGTGCAGTATGTGACACATTTGATAAAGCTCAGGAA
CATGCTTTAAAGTTGGCTGAAGAGGAAGTGCATTCACTTCATCCGCCTTTCCATGC
ACCCAGATGTGATCATAGGGAACAGTTGAATTGGAACGGAGATTAATCGTCAAC
TTAAAGAT---
ATACATGCAATTTTGTGCCGTAGGGAGGGAGGTTTTAAATGCTGGTGTTGCTG
CATATTCCAAAGGTTGGTCCTCCATACAAAGATTATGGAGTTGAGCCATTTGG
TGCAAGTTCAATAGCAACAACTCTTTGATCCATAGGGAAGAGAATGAAATTTGGAACA
AGTTGATAATTTTGACATGGCGATGCTGGTGCAGTATGTGACACATTTGATAAAGCTCAGGAA
CATGCTTTAAAGTTGGCTGAAGAGGAAGTGCATTCACTTCATCCGCCTTTCCATGC
ACCCAGATGTGATCATAGGGAACAGTTGAATTGGAACGGAGATTAATCGTCAAC
TTAAAGAT---
ATACATGCAATTTTGTGCCGTAGGGAGGGAGGTTTTAAATGCTGGTGTTGCTG
CATATTCCAAAGGTTGGTCCTCCATACAAAGATTATGGAGTTGAGCCATTTGG
TGCAAGTTCAATAGCAACAACTCTTTGATCCATAGGGAAGAGAATGAAATTTGGAACA
AGTTGATAATTTTGACATGGCGATGCTGGTGCAGTATGTGACACATTTGATAAAGCTCAGGAA
CATGCTTTAAAGTTGGCTGAAGAGGAAGTGCATTCACTTCATCCGCCTTTCCATGC
ACCCAGATGTGATCATAGGGAACAGTTGAATTGGAACGGAGATTAATCGTCAAC
TTAAAGAT---
L_BARB_TD2
GACTCTTCAGGACCCGTTGGGTAGTGCTCCTCCATCCCTTGACAGGAGGACTGGA
AAATTGATAGAGAAATACCAACCGGGTGCGTAGAGGAGAATTTG---
TTCCAGTACTTGGCTGCAAGAGGTAGTGTTGCTATTCAGTC
CCCGTGTGCAATGGCAATTGTTTGAATAAGTTGCTACTGTTGCTATGAC
TTTCCAAAGGCTGGCTGGTGTGTTTCCATTGGAATAGTTGCTACTGTTGCTATGAC
TTTCCAAAGGCTGGCTGGTGTGTTTCCATTGGAATAGTTGCTACTGTTGCTATGAC
TTTCCAAAGGCTGGCTGGTGTGTTTCCATTGGAATAGTTGCTACTGTTGCTATGAC
CAPSIC_TD2
AATCCCTTTGACACGGTATGATGCTAGATCCTTAGAGAAGCCGAC
GAGTTGATCGTGAAATAATCCACCTGCGGGCGCACAAAGACGAC------
ATTCAATATTTTCTGACATGTTGTCGTCGCGGGGTGGAAGGCTGCTAAGAAT
CTCCGTTGCAACAATCGATGAAATATTTCCGAAAGGGGTGTAACACGCTAC
AGCTGAAAATCATGCAATGCTTCATGGAGATCTCGATGACTCACG
TCACACCTTCTATATGAAACAAGGGCAAGGACAGTTGA
AATACTGGGGCCTACGTAATCTTTAGGGCCACACATTTGGATAGCAGCTCA
ATATGCAATGACGATGTCTGAAGATGAAGATCTCACATTCATCCCACCATAG
TCATCCTGATATCATAAAAGGCAAAGGTAATCGACAGATTAGCCGTC
TTTGAAGCATAATCAGTACATGCAATATTTGTACCTATTGGTGGGGGCTAGCG
GCCGGTATAGCTACCTATATGAAACAAGGGCAAGGACAGTTGA
GTTGAACCATATGTTGCTAGTCTCAATAAATGGAATGGTACTTGTTGA
TCGTGATCTAGTATGCTACATAAAAAGGTATGTAATGGAAGAAAGAAATAT
ATTAGAGACATCAGGTACCTTGGCCTAGCTGGAGCTAGCATTGAAATAAT
TTATAATATAAAAATGAAAAATTTGGCTAAGGCAAGGAAATATGGA
TTTGAGCAAAATCTAAATATGGAATGGGCAAGGCTATAAAAGGATGTAATGGA
AGCTTATAGCTACTTTTTATGCGCAAGGAAACCTGGGAAAGCTTCAAGAATC
CATGTTGAAGTTTTTACTCATTAAATATCATGTTTGTTCAGATTCGTCAC
ATTTTGCTGGACTTTATGGAATATTTTGTTGCTGACATTAGCTATAGCTAC
AACATACAGAATGCAACATCAGTACATGCAATATTTGTACCTATTGGTGGG
CTACGGGAGATCGATCCTCGTCATGAAATTTTATGTCAATTTATTTCCCTG
AGACCTTTATTAGCTACTTTTACGCCAAGGAAACCTGGGAAAGCTTCAAGA
ATTTCGTGAAATTTTACTCATTAAATATCATGTTTGTTCAGATTCTGTCAC
ATTTTTTTCGACATTTGCTTTAGG--------
ATACAAATACGAATGCAAAATGTCGCAATGATGCTTGAAGACTTTTCA
ACATGTTGCTGCAATAATAATCTAGAAAATTTTGTCAACTTTTATGG
TGGCACATCGATCCTCGTCATGCAAAATTTATGCTAAAAATTTTTCCTGAG
AGCTGGAGCTTTAAGATGTCGTTAGCTTTTTAGCCTCTGAGGAATATAACTT
TATTCCATTATCGTACGAGGAGAGGAGTACGAGTAAGCGAGTTTGCTGCTG
AGTCCAAAATGGGAAAATGCGAAATTTTCAATATCAGTAAAAATTTTGTTG
TACATGAAATTGAGATCCACAATGAGGCTTTACAAAAACTTATAATGGAG
S_DULC_TD2
GATTCTTCAGCGCTTTACCCGAGTGAACAGTAATTTGCGAGCTCCACGCGG
AGTTGATAGAAAAATAAACGAATTTTCTACGATCCAGATGAATTA---
TTGCAATATTGGCTGCAATGTTGGCATCACCAGTATAAATGTCGCGAAAGTAA
CCCCCTGGAGCCCTGCAAAGAGATTTCAGATAAGCTTGGAGATTAAATTTTATAT
AAAAAGAGGAGTTAACAAGTGTGTTTTCATTTACGAGATGAGGTAGCTTACAA
CATGATGAGTAGCTCTCTACAGAGAGCAATTAGATAAAAGGGGTGTAACGACATC
AGCTGGAAAATCATGCAAAGGTGTTGCGACTATGTGCTGCAACACATTAAATGCA
GGCTACAAATTGCATGCCTGAGACCCACCCAAAGTACAGGTTAAGCGGCTTAG
ATCCTAGGTGGAATATAAACCTGTTGGACGACATTGTGAAAGCTCAGTCAAGC
AGAGGCTGAGGTTTGGTGAAGATTGCAATTTAGTAATATCTATCAATTGCTTGG
TGATCCAGGAGTCATCAAAGGCAAGGGTATCATGGGAAATGAGAAAAACATC
AACATGGGAC---
ATCTATGCGGTTTTGCTGCTGTTGCTGTTGCTGCTTCTTAATAGCTTGGTTCG
TTATTTCAAAAAGTTGCCCCAAAATCAAAAAAATTTATGGAGTTAACCATTGGT
GCAGCCTATAAGCATAATTCTATATATAAAAGGTCAAAGATAAAATTTGAAAAT
GTTGATACCTTTTTGGGATGGTGGGTACCACAAAGTTGAGTTATACAAATTTCA
AAAATGGTACAGTTATGTTATGATACTTGGACGTATATTGAG
GTCGAATAAAAGGATGTTGATGATAACGGAAGAATATAGAGACATACAGGG
TGCAGTTTCCATAGCTGGAGTCAAACATATTGCAAATAATACATCCAAAAT
GAAATATTGCAAAATGTTGAGCACATATGGATTTCCAGCAAAACTAA
ATTAGTTCAGACTAGCAGGAGTGGTCCAGGAGAAGCTTTTGGTACAT
TGTATGCCAGAGAAACAGGAACCTTTCTTAGAAATTTTCAGCTCTCT
TTCAATATTACGAAATCTACCATGACATAGATGCTGCAAGAAGAAGCTGTA
GCTCTATACAGTGTGTTGATTTGATGAAATAGCAGTATATTGCAATTTGAAG
AAGGGTAAATCAGAGGATTTCCATACGTGTGATCTTCACATCATGATGTTGCA
ATGAACACCTAAGGCAATTTGGTGGTGGGCTGCAATCAATCTAGTATGAGCTT
TCGATCAATTTTCTCTGAAATTTGCTGAGCTCTAAGAAAAATTTCTATATGCT
TTCTCCCTCTTGTTGGAATATAAATCTTTTGGCTCCGGCTACGGAACAGGAAGTAT
ATAGCAAGTGTGTTGAGTTTCCAGCTAAAGATTGAGTTGAAATTTCCAAAAT
GAAATCAGAGCTATATAAGGTGGTATGCAATATTGCTTGGAAATAGCTGAG
TTATAAGCTTATAAGGAAT_S_MEO_TD2
GATTTCTCTGCCCCCTTGAGGGTGAATGAACCTCTGACCTGGAAAAGCCCTGGA
GAGTTGATAGAAAAATCATGCAAATTCTTGATCCAGATGAATTA---
AGGCAATATTGGTCTGATATTGAGTATCAGGCTGAGGTATGCTGAGTCCAGGAGT
CCCCGTGGAAACGTCAAAACAAGATTGAGATGGTCTGCTGAAATTTTCTAT
CAAAAAAGAGGATAGACAGCTGTTTTCATCTAAGTATAGGAGGCTTACAA
CATGATGAGCAGCCTATCTCAGAGGAGGATGGAATAAGGGGTGTAACAGCAGTCA
AGCTGAAAATCATGCAAAAGGTGTTGCAAGATGCTGCTAAAGACTATAATGTC
AGCTCAATTTTGCAGCTCATAAACCCACAAATCAAGGTTGAGCGGACATAG
AACCTAGGTTGGAATGAAAATCTTGAGGTTAGGTAACACATTTGATGAAAGCTCAGC
TTATGCCTCTGCAATTGGTTCAAGAAAAAGGTACAAATACATCCCACCATTTGA
TATCCTGGAGTCATCAAAGGCTACAAATTGGAGTGGAAATAAACCGC
AGCTTAAGGAC---
ATCCATGCAATATTTATACCTGTTTGGTGGTTGCTAATATCTGTGTTGCTGC
TTTTTTCAACAGTAGCCCAAAATACAAAAAATTATTGGGTTGAAACCATATGGT
GCAGCTTCATGACTGTGTCATTACTCGAAGGTTAAAAAGATAAAAATTGGGCAAT
GTGATACATTCTGCTAGTGGTGTGCTGACTGACTGAAATAGTGCAATTTTCA
AAAAATAGCAGAAAATAGGAAGCGGSACACATTTGACTGACAGATGATGTTAGT
GCTGGACTAAAGGATGTGTATGATGAAGGGAGGAACATATTAGAGACATCAGG
TGCACTTGCCTAGCTGGATCTGCAAGCTATTTGCAATTGCAATTGCTCAGCAAACATAA
ATTAGTGCCAGAAATTAGCAGATTGGTGCAGAAAGAGAGTGGTATTTATGCTAAGCTC
TTTTATGCAGAACAAAGCTAGGAAAGCTTCCACAAATTCATTAAATTGTTTGGCTCT
TTCATATTCTGACATCTCGACATGACTGTGTCATTACTCGAAGGTTAAAAGAATAAAATTGGACAAT
GTTGATACTTTTGCTGATGGTGTTGCTGTTGCACTTGTTGGTGAATACAATTTTCA
AAAATGCCATGAATTAATTGATGGAATGGTACTCGTCCATCGTGATGGTATTAGT
AGATGAATTCCAGAAGGATTTCACAACGTGTGATCTCTCACATCATGATTITGGCCAA
GGAAACATCTAAGCATTGTTGTTGCTGAGGCCCTTAATCTCTAGTGAGATTAT
CTTGCAATTTATTTTCCTGAAATTTGCTGAGCCTAAATAGATTTCTTATAGCTCT
TCTCCCTCCTGTTGGAAATATAAATTTGGCTCAGACAGGGAGAAATAGA
TGCAAGTTCATTTGTTGAGTTTCAATTGACCTGACCTGCAGGAGGAGGAGGAGTCAAA
AAAACAAAGCCAAATAAACTTTGGTTATTCATATGCATTGGATATGGCTGAATGAG\nGCTCAATTGCTATTGGAATCTGCAAGCTATTTGCTGATTTATGCTAAGCT
S_PSEU_TD2           ---ATT---GCG------------GTGCGCT---TCA---
CATCTGGAGCCCGGAGAGTTTGTGACCACACATCGCCAACCCGGGGCCGATAGAGA
AGATTTG---
TTACACATATTTATTTGACATGTTGGCATCGCCAGTCTACGATGTCGCGATGGAAT
CCCCGTTGGAAAAAGCGAGAGATTCTCAGCAACACGCTGGAATATTTTTT
TCAAAAGAGAGGAGAAGGTTATTTGCTTACATCTTACGATTGAGACTTACA
ACATGATGAGAAACACCTCAGAGGAGCAATATTAAAGAAAAAGGAGATTATATACAGCAA
TCAGCTGAAATCTGACAAAGGTGGTTGCAATTAGTTTATATATATAA---
AAATGTCCAGCTACAAATTGCTGTCATCTACACTACACCCACAATACTGAGATTGAT
GCAGTTGAAGCTTGAATCTCGAAGACTCATTAAAGGGAAAAATTTTGAGTGA
GCTCAATTGCTATTTGGAATCTGCAAGCTATTTGCTGATTTATGCTAAGCT
S_PSEU_TD2           ---ATT---GCG------------GTGCGCT---TCA---
CATCTGGAGCCCGGAGAGTTTGTGACCACACATCGCCAACCCGGGGCCGATAGAGA
AGATTTG---
TAGACCATCTAAACATTTTGGTGCTGGTGAGCCTAATATACGTGATGAAATTTT
CGGTGAATTTATGTACCTGAGAAGGCTGAAACTTTGAAAAACATTTTACCTAGCGCT
TTCTCCCTCGTTGGAATATATAACTTTATGCGGTTACCCTAAATCAGAGATATAA
ATGCAAGTGTGTAATGGGATTCCTCCTAAATGAGATGGATGAGTTCA
AAAATCAAGCTGATAAGCTTTGGTTATCCATATGAAGCTTGAATATTTAATAGGAG
CTTTAAACCTGTGAGTACT
S_LYCO_TD2
GATATTACAGCTCCTCGGGTGGAAAATGTTCCGCTATTTGCAAAAAAGTTCCCGGA
GAGTTGATAGTAATAAAAACAAACGGGGGGCGATACGAGTGAATTG---
TTTCAATATTTAGTTGATATTTCTGACGCTCGCCAGTGTATGAGCCTGAATC
GCCATTAGAGCTAGCAGAGAAGGCTGCTTCTCGGATAGCTGAGTAAACTTTATATC
AAAAGAAGGATAAACACAGGGGTTCAAGGAGCTAGGCTTAC
S_BULB_TD2
GATTCTTCAGCTCCAGTGGTAAAGTGTTCCTGCAAACTTGCCAAAAACCCCTGGGG

96
ATATGCAGTGGAACTGAGTGAANAAAGACAAGTCACATACATCCCCACCATTTGGCT
CAGCTTCCAGAGTTCATCAAAGGGCCAAAGGTGAATAATGGGAAGGATATAAACATGT
CAAATTTAAGGC---
ATTCATGCTGTTTGTACCTGGTGTGGTGGGTTCTAATATCTGGTGTTGCTGCT
CTTTTTCAAAAATGCCCAGATACAAAAATTATGGGAGTTGAGCCATATATGCT
GCAGCTTCATGACATTTGTCATGGTATGAAAGGAGCATAGAGTAAATTGGAAAAT
GTTGATACTTTTGTGATTTGTAGGCTGTAAGTGGGCACTTGTGTTGGAATACACTTTTGC
AAAAATGCCAAAGTAAATGATGGAATGGTACTTGTTAGAAATGATGGTATTAG
TGCTGCAATATAAGGATTGTAAGTGAAGGAAGGAACATATTAGAAGACATACAG
GTGCAAGTGGTACATACAGTGGGACTATTTCAACATTTAAATTGATGTGATGGGATTCCAAGTCCCTAAAGTGAGATGGATGAGTG
CAAACTGCAATTTATATGAAGACTTTTAATTTCTCACAATAATGATGGTGT
GCAGAAACATATAAATGAGTGTTGCTGCTAGCAAACTATGATGAGGTTTACAT
GTTGGAATATTTTATATTTTCTAGAAAGGTGAGCTGAAGTTGGAATTGATAGCT
CTTTCCCTCCCTGGGAATAATAATTATGAGCCTGGATCGAGGAGATAT
AATGGGAATGTGTTATGTGAGCTACAGTCAGGTAAGAGAAGAGCATATTAGA
GTTTTGTCAGAGTTGCTTGGGATATTTGGGTACAAAGTATATGATGTAGGCCCATGAATCTCCGTT
ACAACTTGCGCCAAAGCTTTCACAGAAGTTGGGGTTAATGTTTGCTCAAACG
AGAGGATCTTCTTACATTAGTCTAAGGAGGTGTAGCTCAACATATGATG
GTCAATCTGACAAGGACAGTGAGTTAAAGAGGGGTTATATGCTCATCGGCTGGG
AATCATGCAACAGGTTGGCTTACATTAGCTTGGCAGAAATTTTGTGATGCTGCT
TTTGTGATGCTCTTATACACCAGAATTTAAGTGGAAAATCAGTTAAAGAGATTG
GGCGCTACTTTTGTGGGAGATCGTTATGAGCAGCTCAAGCATATGCCA
AAAAACAGGTCAGAGGAAGGCGGGRGCCAGTCATTCCCCTTTTGATCATCGAGG
AAGTCTCATGGACAGAAGGTACATTGGTATGGAAATTTATCGCAACACAAGG
ATAAATTCTCATGCAATATTTGTGCTTGGTGTAGGAGTAGTAAATGCT
CGATCAAGTGGGAATGTTGCCAGATTTGTAGGCTGTTAAATGTTGTTGGCAAGA
ACCTTTCCGTCTCTGTAAGGAATTTGGATAGACGGGGGTGTCTCTTATAAGTCGAGAT
GCTATATGGCATCATAAAAGGACATGTTGGAGAAGAAAAAGGACATATTAGA
GCCTGCAGGTGCACTTGCTAGCTGGAGCTGAAGCATACTGCAAGTACTATGGCCTCAAGGATGAAAACGTAGTAGCAATAACTAGCGGAGCCAATATGAACTTTGACAGACTAAGATTAATATCCGAACTTGCAGATGTTGGTAGAAAGCAGGAAGCTGTGCTAGCTACTTTTATGCCAGAAGAGCAGGGAAGCTTGAAACAATTTTGTGAACGGGACAAACTGTGGCACAAATGT TT TATGACTCTTTTCACAGCCCGTTGGAATATAAGTTTGATGACATATTCCGACA GCCTCCGACTCCATTGCTAGTATTTTCTCCGAATTCTGTTACGGTGAGCCTGGGA TCTTATGATACCGAATTATCCGGTGGGCGGTAATGGAGGTGAGAACGGGTTTCAGTATCTAGTGGATATATTGGGTACGAAAGTGTACGATGTAGCAAATGAATCGCAT TGCAGCTTGGCAGTCGTTTACAGTGTTGGGAGGCTTCAAGTTCCACAGGATGAGTTGATGAATTCCAGTATCGCGCTGCCAGTCTTGGTTATGAGTATGTGGTGGAGAGTAAATGATGCTTTCAAACTTCTAATGCAT...
TTAGATACTTTTCAGCCCGCTTGGGAATATAAGTTTGATCCATTATCGCGGACAGG
GACAAGATTGCTGAAATGTTTATAGTTGGGATACAAGTTCCAGAGGCTGAGTTG
ATGAGTTCCAGGGTCGAGCTGCCAATCTTGGTTATGAATATGTGGTGGAGAGTCT
AAATGATGCTTTTCAAGCTTATAATG

N_ATTEN_TD4

CCTCCGACTTCTTTGCTAGTATGTTTTCTCCGAATTGTTACAGTGAGCAGCCTGGCT
ACTTGATACGCCAATTATCCCGTAAATGAGTTGGGACAGATTAGTACGATGTAAGGAAATAGTCCATTGAGCTTTGCTG
CAGAGGACAAATTGGTGCAAATGTTTTAGTTGGGATACAAGTTCCAGAGGCTGAGTTG
ATGAGTTCCAGGGTCGAGCTGCCAATCTTGGTTATGAATATGTGGTGGAGAGTCT
AAATGATGCTTTTCAAGCTTATAATGC

N_BENT_TD4

CGCTATATGTCATTGAATGCTGGAAATATAGGATGCTGGACTGCTGCTTACTACACCAAAGAGTTAATGGAACATAGAGTTAAATCAGCTG
CGCTATATGTCATTGAATGCTGGAAATATAGGATGCTGGACTGCTGCTTACTACACCAAAGAGTTAATGGAACATAGAGTTAAATCAGCTG
CGCTATATGTCATTGAATGCTGGAAATATAGGATGCTGGACTGCTGCTTACTACACCAAAGAGTTAATGGAACATAGAGTTAAATCAGCTG
CGCTATATGTCATTGAATGCTGGAAATATAGGATGCTGGACTGCTGCTTACTACACCAAAGAGTTAATGGAACATAGAGTTAAATCAGCTG
CGCTATATGTCATTGAATGCTGGAAATATAGGATGCTGGACTGCTGCTTACTACACCAAAGAGTTAATGGAACATAGAGTTAAATCAGCTG
Appendix E: SDS-PAGE of chymotrypsin exposed SlTD2 and Mutants

SDS-PAGE (10%, Tris-Glycine buffer) was run to determine that all of the enzyme incubated with chymotrypsin was cleaved. Two incubations of each enzyme were run.

Lane 1: Standard; 2: Control; 3: SlTD2; 4: SlTD2; 5: Mut1; 6: Mut1; 7: Mut2; 8: Mut2; 9: Mut3; 10: Mut3. Control (a) shows normal size of TD2 and Mutants, ~55kD (lane 2), (b) shows the cleaved TD2 and the Mutants ~40kD. (c) is the chymotrypsin band ~25kD. All chymotrypsin exposed enzymes were fully cleaved before the activity assays. Note: Gel was damaged before photo taken but it did not affect results.
Appendix F: One-way ANOVA code in Python

```python
from pyvttbl import Anova1way
print("############################")
print("ANOVA for Kcat")
print("############################")
d = [[262.939894114,267.7981937091,255.9825599502],
    [214.8584119933,210.3275957801,217.5930038867],
    [262.6191172172,255.746857644,273.867290266],
    [284.0514285714,253.3485714286,262.3885714286]]
conditions_list = 'CIP1 CIP2 CIP3 TD2'.split()
D=Anova1way()
D.run(d, conditions_list=conditions_list)
print(D)

print("\n\n\n\n############################")
print("ANOVA for Km")
print("############################")
m = [[0.36751,0.41894,0.484027],
     [0.363046,0.45325,0.505294],
     [0.298005,0.394395,0.403787],
     [0.493134,0.418784,0.52156]]
conditions_list = 'CIP1 CIP2 CIP3 TD2'.split()
D=Anova1way()
D.run(m, conditions_list=conditions_list)
print(D)

print("\n\n\n\n############################")
print("ANOVA for Kcat/Km")
print("############################")
c = [[715.4632366847,639.2280367335,528.8600841486],
     [591.8214551141,464.04323393,340.6265340311],
     [881.2574192285,648.4536001825,678.2469292222],
     [576.0126630316,604.9623945246,503.084154131]]
conditions_list = 'CIP1 CIP2 CIP3 TD2'.split()
D=Anova1way()
D.run(c, conditions_list=conditions_list)
print(D)
```
import pandas as pd
from statsmodels.formula.api import ols
from statsmodels.stats.anova import anova_lm
from statsmodels.graphics.factorplots import interaction_plot
import matplotlib.pyplot as plt
from scipy import stats

# temperature assay anovas

datafile="temp2WAnova.csv"
data = pd.read_csv(datafile)

def eta_squared(aov):
aov['eta_sq'] = 'NaN'
aov['eta_sq'] = aov[:-1]['sum_sq']/sum(aov['sum_sq'])
return aov

def omega_squared(aov):
    mse = aov['sum_sq'][-1]/aov['df'][-1]
aov['omega_sq'] = 'NaN'
aov['omega_sq'] = (aov[:-1]['sum_sq']-(aov[:-1]['df']*mse))/(sum(aov['sum_sq'])+mse)
return aov

fig = interaction_plot(data.Temp, data.Enzyme, data.Product, colors=['red','blue','green','grey'], markers=['D','^','o','v'])
fig

formula = 'Product ~ C(Enzyme) + C(Temp) + C(Enzyme):C(Temp)'
model = ols(formula, data).fit()
aov_table = anova_lm(model, typ=2)
eta_squared(aov_table)
omega_squared(aov_table)
print(aov_table)
d1 = data[data.Enzyme != 'CIP2']
d1 = d1[d1.Enzyme != 'CIP3']
d2 = data[data.Enzyme != 'CIP1']
d2 = d2[d2.Enzyme != 'CIP3']
d3 = data[data.Enzyme != 'CIP1']
d3 = d3[d3.Enzyme != 'CIP2']

formula = 'Product ~ C(Enzyme) + C(Temp) + C(Enzyme):C(Temp)'
model = ols(formula, d1).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
print("
\n\nTD2 vs CIP1")
print(aov_table)

formula = 'Product ~ C(Enzyme) + C(Temp) + C(Enzyme):C(Temp)'
model = ols(formula, d2).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
print("
\n\nTD2 vs CIP2")
print(aov_table)

formula = 'Product ~ C(Enzyme) + C(Temp) + C(Enzyme):C(Temp)'
model = ols(formula, d3).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
print("
\n\nTD2 vs CIP3")
print(aov_table)

###pH Assay Anovas###

datafile="pH2WAnova.csv"
```
data = pd.read_csv(datafile)

#fig = interaction_plot(data.pH, data.Sample, 
data.Product, colors=['red', 'blue', 'green', 'grey'], markers=['D', '^', 'o', 'v'])

formula = 'Product ~ C(Sample) + C(pH) + C(Sample):C(pH)'
model = ols(formula, data).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
print(aov_table)

d1 = data[data.Sample != 2]
d1 = d1[d1.Sample != 3]
d2 = data[data.Sample != 1]
d2 = d2[d2.Sample != 3]
d3 = data[data.Sample != 1]
d3 = d3[d3.Sample != 2]

formula = 'Product ~ C(Sample) + C(pH) + C(Sample):C(pH)'
model = ols(formula, d1).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
print("\n\n\nTD2 vs CIP1")
print(aov_table)

#formula = 'Product ~ C(Enzyme) + C(Temp) + C(Enzyme):C(Temp)'
model = ols(formula, d2).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
```
print("\n\n\nTD2 vs CIP2")
print(aov_table)

#formula = 'Product ~ C(Enzyme) + C(Temp) + C(Enzyme):C(Temp)'
model = ols(formula, d3).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
print("\n\n\nTD2 vs CIP3")
print(aov_table)

### Chymotrypsin anova

datafile="chymo2WAnova.csv"
data = pd.read_csv(datafile)
data

fig = interaction_plot(data.Protease, data.Enzyme, data.Product, colors=['red','blue','green','grey'], markers=['D','^','o','v'])

formula = 'Product ~ C(Enzyme) + C(Protease) + C(Enzyme):C(Protease)'
model = ols(formula, data).fit()
aov_table = anova_lm(model, typ=2)

eta_squared(aov_table)
omega_squared(aov_table)
print(aov_table)

d1 = data[data.Enzyme != 'CIP2']
d1 = d1[d1.Enzyme != 'CIP3']
d2 = data[data.Enzyme != 'CIP1']
d2 = d2[d2.Enzyme != 'CIP3']
d3 = data[data.Enzyme != 'CIP1']
d3 = d3[d3.Enzyme != 'CIP2']

formula = 'Product ~ C(Enzyme) + C(Protease) + C(Enzyme):C(Protease)'
model = ols(formula, d1).fit()
aov_table = anova_lm(model, typ=2)
eta_squared(aov_table)
omega_squared(aov_table)
print("\n\n\nTD2 vs CIP1")
print(aov_table)

model = ols(formula, d2).fit()
aov_table = anova_lm(model, typ=2)
eta_squared(aov_table)
omega_squared(aov_table)
print("\n\n\nTD2 vs CIP2")
print(aov_table)

model = ols(formula, d3).fit()
aov_table = anova_lm(model, typ=2)
eta_squared(aov_table)
omega_squared(aov_table)
print("\n\n\nTD2 vs CIP3")
print(aov_table)
References


Rambaut, A. 2016. FigTree v1.4.3.


threonine deaminase gene of tomato: isolation, structure, and upregulation in floral 

paramyxovirus glycoproteins. Activation of cell fusion, hemolysis, and infectivity 
by proteolytic cleavage of an inactive precursor protein of Sendai virus. Virology 
57:475–490.


Srinivasan, A., A. Giri, and V. Gupta. 2006. Structural and functional diversities in 

1218.

A. 111:2223–8.

Sukumaran, J., and M. T. Holder. 2010. DendroPy: A Python library for phylogenetic 

moderately thermophilic and extremely thermophilic protein subunits results of a 


Vogt, G., S. Woell, and P. Argos. 1997. Protein thermal stability, hydrogen bonds, and 

effector-binding sites in threonine deaminase by site-directed mutagenesis, kinetic, 


Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. University of Texas at Austin.