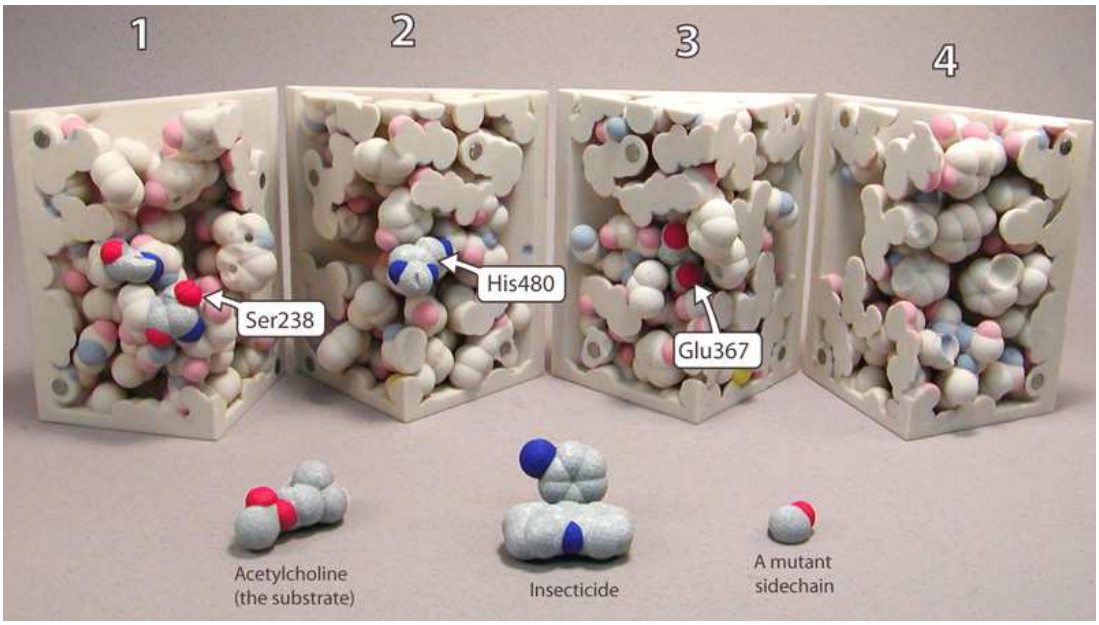
Acetylcholinesterase



Teaching Points

These materials explore the structure-function relationship of a single protein, acetylcholinesterase. The acetylcholinesterase gene and the protein it encodes can be used to demonstrate a number of biological concepts, including enzyme specificity, competitive inhibition, mutation, characteristics of the genetic code, alternate splice sites, natural selection, bioinformatics, and disease transmission. Many of these materials can be presented either at an introductory or an advanced level. For instance, materials are provided to allow students to compare DNA and protein sequences by hand, or these comparisons can be completed using programs available on the Internet. Furthermore, it is not necessary to incorporate ALL the materials in the classroom at one time. We often use acetylcholinesterase as one of a few 'anchor' proteins that serve as examples to help tie together a number of biological concepts. Our hope is that when students leave our classes, they can provide specific examples of a few proteins that illustrate a number of biological concepts.

Nerves send signals to muscles by way of neurotransmitters, such as acetylcholine. Once the signal has been received, the neurotransmitter must be broken down and recycled, or the cell would not be able to distinguish a new signal as it enters the receptors. Acetylcholinesterase, which breaks down acetylcholine into acetic acid and choline, is the target of nerve agents such as sarin as well as insecticides like DDT and Malathion. For more background information on acetylcholinesterase, check out David Goodsell's Molecule of the Month on the Protein Databank website: http://www.rcsb.org/pdb/education_discussion/molecule_of_the_month/download/Acetylcholinesterase.pdf.

Aside from being pesky nuisances, mosquitoes often transmit blood-borne diseases, such as malaria and West Nile virus. In an attempt to control the spread of disease, insecticides have been used widely. Variation in the acetylcholinesterase gene has allowed some mosquitoes to be resistant to insecticides. When mosquitoes are exposed repeatedly to insecticides, those with resistance to the chemicals survive to reproduce, resulting in a population of mosquitoes that are no longer able to be controlled with the insecticide. This material explores the nature of the mutation that confers resistance at the molecular level and is based on a paper published by Weill, et al. (2003).

First, students explore the structure of acetylcholinesterase using computer simulations (located at http://www.scripps.edu/mb/goodsell/jmol/ach) as well as physical models available from the MSOE Model Lending Library. They discover that the catalytic triad consists of serine, histidine and glutamic acid residues that are separated by many amino acids in the single polypeptide, but that the three residues lie physically near each other, buried at the base of a

long, narrow pocket within the protein. Although the substrate fits neatly in the active site and is physically in contact with each of the residues in the catalytic triad, students quickly discover that the only way for the substrate to access the active site is if the protein undergoes a conformational change – appearing to 'breathe' to allow the substrate to reach the catalytic site.

After exploring the active site of acetylcholinesterase, students compare the DNA sequence from two strains of mosquito, one sensitive, and the other resistant, to insecticides. They discover that there are 28 nucleotide differences between the two strains. Next, they align the protein sequences, and discover that there is only ONE difference between the two strains. This activity demonstrates silent mutations – changes at the DNA level that are not expressed as changes in the protein level. This is a great opportunity to review the characteristics of the genetic code – it is a degenerate code – more than one codon can code for a specific amino acid), and that there is third base 'wobble' (often mutations in the third base position still code for the same amino acid).

After examining acetylcholinesterase bioinformatics, students then go on to study the impact of the specific amino acid change on the protein function. They return to the computer simulations and physical models. They discover that the mutation is not in the catalytic site of the protein, but in the long narrow pocket leading to the catalytic site. Insecticides serve as competitive inhibitors, blocking the path to the active site of the molecule. Students discover that the mutation causes a glycine residue to be replaced by a serine residue, and that this minor change still allows the substrate to squeeze through the gorge. They also discover that, although the residue in the resistant strain is only slightly larger, it is sufficiently large to block the inhibitor from entering the gorge.

After students have discovered the nature of the mutation and the impact on protein structure, you might have them go one step further by asking them to propose a design for a better insecticide – one that would inhibit the enzyme regardless of what mutations might arise. Students might suggest an inhibitor that binds directly at the catalytic site, providing another teaching opportunity. (Insects with mutations at the catalytic site would not survive, since the protein would not function normally.)

This manual is organized in sections for those instructors interested in presenting only a portion of the material. Most sections are divided into subsections. Throughout the material, Section A is more appropriate for introductory courses. Students are provided with the aligned sequences and are required to identify the differences between the sequences. In section B, students actually use Internet tools for bioinformatics to align the sequences. There are additional suggestions in section C to allow for more in depth study of bioinformatics.

Models in this Collection

- Two acetylcholinesterase active site boxes
- Acetylcholinesterase backbone structure with embedded box
- CD with documentation

Model Details

- Acetylcholinesterase active site box
    - Fold out model of acetylcholinesterase active site gorge in four sections, including substrate (acetylcholinesterase), competitive inhibitor (insecticide), serine 238 cap, and mutant sidechain
    - Catalytic site residues (ser238, his480 and glu367) in cpk colors; rest of model in muted Goodsell colors (pink is oxygen, pale blue is nitrogen, white is carbon)
    - Magnets join the four sections, as well as the substrate, inhibitor, ser38 cap and mutant sidechain, all of which are in cpk colors
    - Mutant sidechain converts a glycine to a serine residue
    - Based on PDB file 1QON
    - Made of plaster on the Z-Corp printer
- Acetylcholinesterase backbone model
    - Catalytic site residues (ser238, his480 and glu367) ball and stick in cpk colors; aromatic residues lining active site gorge in yellow ball and stick
    - Embedded box indicates where the model was cut to make the active site box
    - Based on PDB file 1QON
    - Made of plaster on the Z-Corp printer

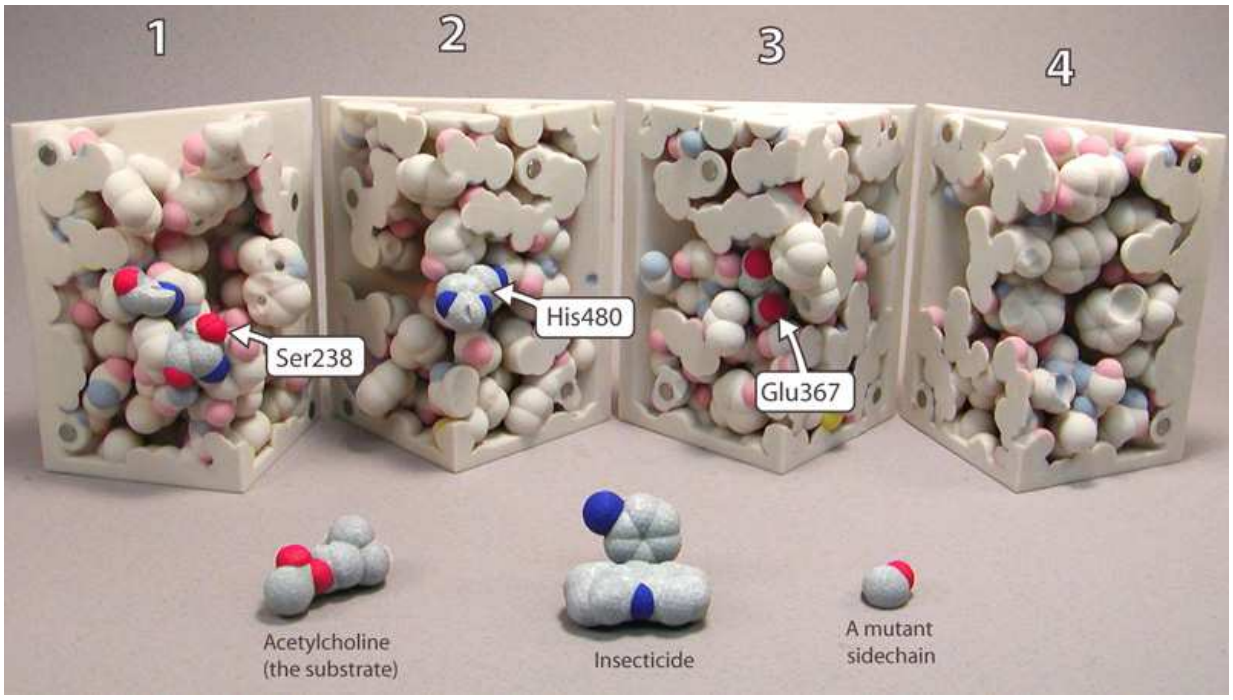
Documentation included

- Teaching points and inventory
- How do the models fit back in the suitcase?
- Evolution in Action: Modeling Insecticide Resistance in Mosquitoes (Instructor's Guide)
- Acetylcholinesterase – information from NCBI
- Worksheets
    - SR and SLA-B DNA and protein sequences
    - Acetylcholinesterase SLAB-SR DNA Alignment
    - Acetylcholinesterase SLAB-SR DNA Alignment Key
    - Alignment of SLA-B and SR protein sequences
    - Alignment of SLA-B and SR protein sequences key
    - Analysis of location of mutations in SR acetylcholinesterase within codons
    - Alignment of Acetylcholinesterase Gene from Strains Sensitive and Resistant to Insecticides
    - Alignment of Acetylcholinesterase Gene from Strains Sensitive and Resistant to Insecticides Key
    - Summary of nucleotide differences in acetylcholinesterase gene among 30 mosquito strains
- Copy of Weill et al. paper: Insecticide resistance in mosquito vectors. Nature 423:136-137.

Center for BioMolecular Modeling

. . .where teachers come first

Evolution in Action: Modeling Insecticide Resistance in Mosquitoes



Instructor's Manual

Classroom materials were developed by Margaret Franzen, Pellissippi State Technical Community College, Knoxville TN and Center for BioMolecular Modeling, Milwaukee School of Engineering, Milwaukee WI, in collaboration with Lynda Jones, Caitlin Gabel School, Portland OR. JMol files were developed by David Goodsell of The Scripps Research Institute, La Jolla, CA. Physical models were developed in collaboration with Tim Herman, Center for BioMolecular Modeling (CBM), Milwaukee School of Engineering, Milwaukee WI. Funding for this project was provided in part by NSF CCLI grant #0442409 and NIH NCRR SEPA award RR022749.

Feedback on how the materials work in your classroom would be greatly appreciated. Additional materials are being developed; feel free to contact us for the latest updates.
Correspondence should be sent to Margaret Franzen, Center for BioMolecular Modeling, Milwaukee School of Engineering, Milwaukee WI 53202, franzen@msoe.edu.

Concepts Addressed by Acetylcholinesterase Module

- **The active site of a protein is not necessarily formed by adjacent amino acids in the peptide chain**
- **Protein structures are not static, but flexible**
- **Mutation occurs prior to natural selection**
- **Some mutations at the DNA level are not expressed at the protein level (silent mutations)**
- **Not all mutations are 'bad'**
- **Genes can have alternate introns splice sites, leading to more than one protein from a single gene**

Overview of Materials and Teaching Points

These materials explore the structure-function relationship of a single protein, acetylcholinesterase. The acetylcholinesterase gene and the protein it encodes can be used to demonstrate a number of biological concepts, including enzyme specificity, competitive inhibition, mutation, characteristics of the genetic code, alternate splice sites, natural selection, bioinformatics, and disease transmission. Many of these materials can be presented either at an introductory or an advanced level. For instance, matierals are provided to allow students to compare DNA and protein sequences by hand, or these comparisons can be completed using programs available on the Internet. Furthermore, it is not necessary to incorporate ALL the materials in the classroom at one time. We often use acetylcholinesterase as one of a few 'anchor' proteins that serve as examples to help tie together a number of biological concepts. My hope is that when students leave my class, they can provide specific examples of a few proteins that illustrate a number of biological concepts.

Nerves send signals to muscles by way of neurotransmitters, such as acetylcholine. Once the signal has been received, the neurotransmitter must be broken down and recycled, or the cell would not be able to distinguish a new signal as it enters the receptors. Acetylcholinesterase, which breaks down acetylcholine into acetic acid and choline, is the target of nerve agents, such as sarin, as well as insecticides, including DDT and Malathion. For more background information on acetylcholinesterase, check out David Goodsell's Molecule of the Month on the Protein Databank website: http://www.rcsb.org/pdb/education_discussion/molecule_of_the_month/download/Acetylcholinesterase.pdf.

Aside from being pesky nuisances, mosquitoes often transmit blood-borne diseases, such as malaria and West Nile virus. In an attempt to control the spread of disease, insecticides have been used widely. The acetylcholinesterase gene has mutated in some mosquitoes, resulting in resistance to insecticides. When mosquitoes are exposed repeatedly to insecticides, those with resistance to the chemicals survive to reproduce, resulting in a population of mosquitoes that are no longer able to be controlled with the insecticide. This material explores the nature of the mutation leading to resistance at the molecular level and is based on a paper published by Weill, et al. (2003).

First, students explore the structure of acetylcholinesterase, utilizing computer simulations (located at http://www.scripps.edu/mb/goodsell/jmol/ach), as well as physical models, available from the MSOE Model Lending Library. They discover that the catalytic triad consists of serine, histidine and glutamic acid residues that are separated by many amino acids in the single polypeptide, but that the three residues lie physically near each other, buried at the base of a long, narrow pocket within the protein. Although the substrate fits neatly in the active site, physically in contact with each of the residues in the catalytic triad, students quickly discover that the only way for the substrate to access the active site is if the protein undergoes a conformational change – appearing to 'breathe' to allow the substrate to reach the catalytic site.

After exploring the active site of acetylcholinesterase, students compare the DNA sequence from two strains of mosquito, one sensitive, and the other resistant, to insecticides. They discover that there are 28 nucleotide differences between the two strains. Next, they align the protein sequences, and discover that there is only ONE difference between the two strains. This demonstrates silent mutation – changes at the DNA level that are not expressed as changes in the protein level. This is a great opportunity to review the characteristics of the genetic code – it is a degenerate code - more than one codon can code for a specific amino acid), and that there is third base 'wobble' – often mutations in the third base position still code for the same amino acid.

After examining acetylcholinesterase bioinformatics, students then go on to study the impact of the specific amino acid change on the protein function, returning to the computer simulations and physical models. They discover that the mutation is not in the catalytic site of the protein, but in the long narrow pocket leading to the catalytic site. Insecticides serve as competitive inhibitors, blocking the path to the active site of the molecule. Students discover that the mutation is from a glycine residue to a serine residue, and that this minor change still allows the substrate to squeeze through the gorge. They also discover that, although the residue in the resistant strain is only slightly larger, it is sufficiently large to block the inhibitor from entering the gorge.

After students have discovered the nature of the mutation and the impact on protein structure, you could have them go one step further by asking them to propose a design for a better insecticide – one that would inhibit the enzyme regardless of what mutations might arise. Students may suggest an inhibitor that binds directly at the catalytic site, providing another teaching opportunity. (Insects with mutations at the catalytic site would not survive, since the protein would not function normally.)

This manual is organized in sections for those instructors interested in only presenting a portion of the material. Most sections are divided into subsections. Throughout the material, Section A is more appropriate for introductory courses. Students are provided with the aligned sequences and are required to identify the differences between the sequences. In

section B, students actually utilize Internet tools for bioinformatics to align the sequences. In some cases, there are additional suggestions in section C to allow for more in depth study of bioinformatics.

I Background on Acetylcholinesterase

It is strongly recommended that students begin by learning a little about the function of acetylcholinesterase. A great place to start is at the Molecule of the Month website sponsored by the Protein Data Bank. Acetylcholinesterase information is found at: http://www.rcsb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/pdb54_1.html.

A. Utilizing computer visualizations and physical models

1. You can access computer visualizations of acetylcholinesterase structure for both sensitive and resistant strains at: http://www.scripps.edu/mb/goodsell/jmol/ach/. There are three different links on this page, described below. At each of these links, there are JMol images that allow the viewer to rotate the structure, change the appearance (backbone, space fill, etc.), and zoom in on important features. (No additional software is required to view JMol files, which is a distinct advantage over Chime files.) If you click on an atom in the structure, its identity will appear in the footer at the lower left of the browser window.
   a) 'Basics' explores the structure of acetylcholinesterase, showing the long narrow gorge, or pocket, leading to the active site, which is buried deep within the protein. Acetylcholine is in green and is positioned in the active site.
   b) 'Active site' explores in greater detail, the active site of the protein, including the three amino acids that form the catalytic triad (SER200, GLU327 and HIS440)[1]. Additional amino acids that define the boundaries of the gorge are shown in grey and labeled 'other active site residues'. There are two images, one with the substrate, acetylcholine, in green, and the other with the inhibitor, colored yellow.
   c) 'Mutant' focuses on comparing the active site of the wild type (sensitive to insecticide) and the mutant (resistant to insecticide). [This part of the visualization is explored in section VI, after students have determined the nature of the mutation rendering resistance to insecticides.]
2. After exploring the computer models, students can explore the structure in three dimensional space, using the acetylcholinesterase model available from the MSOE Model Lending Library. These models can be reserved for use for 10-14 days (including shipping time); the borrower only pays for return shipping. Additional information about the lending library can be found at: http://www.rpc.msoe.edu/lib/. It is possible to show students the computer visualizations alone, but our experience has been that students really don't understand fully until they manipulate the molecules themselves. These models serve as thinking tools, allowing students to ask much more meaningful questions about the structure-function relationship in this and other proteins.

II Exploring Acetylcholinesterase Gene Structure – Alternate Splice Sites

A. Students can read the handout 'Acetylcholinesterase NCBI' on the CD to discover that there are two different forms of the enzyme, but only one gene. They will learn that which forms are expressed in various tissues, and they can see the map of the two splice variants, one which splices exon 4 to exon 5 (E4-E5) and the other that splices exon 4 to exon 6 (E4-E6).

B. Students can access the information (and much more) themselves by going to the NCBI website (www.ncbi.nlm.nih.gov) and searching in the Gene section for acetylcholinesterase. If they locate the information on the acetylcholinesterase (Yt blood group) from *Homo sapiens,* they will find the information on the handout in the Summary as well as the section 'Genomic Regions, Transcripts and Products'. If there is time, students can dig much deeper by following links in NCBI. Depending on where they explore, they might discover that the two forms can be expressed in one tissue at different times (for instance, under stress), or that the acetylcholinesterase in the blood serves the purpose of binding organophosphates as well as any acetylcholine that was not removed at the synapse.

III Bioinformatics Part 1: DNA Alignment

A. Using two pre-aligned sequences (one sensitive, one resistant to insecticide)

[1] These residue numbers are for acetylcholinesterase from *Torpedo californica* (electric ray) 1AMN. Computer simulations involving docked protein are from *Drosophila melanogaster* (fruit fly) 1QON, with catalytic residues SER238, GLU367 and HIS480, The corresponding catalytic residues in *Culex pipiens* (mosquito) are SER327, GLU453 and HIS567. Although there are some differences in the overall amino acid sequence in the three organisms, they maintain the same overall structure at the active site, allowing for meaningful comparison of structures. When the models were developed, the crystal structure for mosquito acetylcholinesterase had not yet been resolved.

1. See “Worksheets: SLA-B SR DNA Alignment” on the CD. Students use highlighters to denote differences in the two sequences.

B. Using Bioserver to align two sequences (one sensitive, one resistant to insecticide)
   1. Go to the Bioserver website at http://www.bioservers.org/bioserver/.
      a) Under ‘Sequence Servers’ click ‘Register’ (unless you have already registered at the DNA Learning Center and have a username and password). Registration is free, and it allows you to save your sequences for use in further comparisons.
      b) Once you have registered, you may enter the Sequence Server.
   2. Place the acetylcholinesterase sequences in your work area:
      a) Click on ‘Manage Groups”.
      b) Select “Public” under ‘Sequence Sources”
      c) Search for “Resistant Ache” and “Sensitive Ache” posted in 8/05 and click the box to the left of each of these listings.
      d) Click ‘OK’. This will save the sequences to your work area.
   3. After saving both sequences, align them using Clustal W:
      a) Click on the box to the left of the sequences you want to compare (SLA-B and SR).
      b) Select ‘Align:ClustalW’ from the drop down box near the top of the page, then click ‘Compare’.
      c) A new window will open with the alignment. Yellow boxes indicate positions in which the nucleotide sequences are different. You may print the comparison, but some printers don’t pick up the yellow highlighting. After printing, you may need to use a highlighter to mark the differences between the sequences. Alternately, if you have a color printer available, you may cut and paste the alignment into a word processing document (the color highlights transfer), then print the transferred document.

IV Bioinformatics Part 2: Protein Alignment

A. Aligning sequences from one sensitive and one resistant strain
   1. Students compare the aligned SLA-B and SR protein sequences and discover that, in spite of 28 differences in the DNA, there is only one difference in the protein. A ‘*’ under the two aligned sequences indicates the amino acids are identical. This handout is called “Alignment of SLA-B and SR protein sequences”

B. Translating the sequences, then aligning the translated sequences
   1. Translating sequences
      a) In BioServer website (http://www.bioservers.org/bioserver/), select SLA-B and click the ‘Open’ button to the right of the sequence. Cut and paste this sequence into a word processing file.
      b) Go to http://bio.lundberg.gu.se/edu/translat.html
      c) Cut and paste the sequence into the box, and click translate. The protein sequence will appear in a new window. Next there will be the DNA sequence, with three rows of protein sequences below the DNA sequence. These protein sequences are staggered and correspond to the three reading frames. The one letter abbreviation for the amino acid is printed directly underneath the first nucleotide in the codon that codes for that amino acid.
      d) Cut and paste the protein sequence (at the top of the output) into the word processing document below the DNA sequence. Be sure to label the source of the DNA.
      e) Repeat steps a-d for the SR sequence.
   2. Aligning translated sequences
      a) Go to www.expasy.ch/tools/sim-prot.html.
      b) Under “Sequence 1” select ‘User-entered sequence” and enter ‘SLA-B’ for ‘Sequence Name’.
      c) Cut and paste SLA-B DNA protein sequence (saved in the word processing file) into the ‘paste your sequence here’ box.
      d) Under “Sequence 2” select ‘User-entered sequence” and enter ‘SR’ for ‘Sequence Name’.
      e) Cut and paste SR DNA protein sequence (saved in the word processing file) into the ‘paste your sequence here’ box.
      f) Click ‘submit’.
      g) The results will show the two aligned sequences. The symbol ‘*’ below the aligned sequences shows identical amino acids in that position.

V Bioinformatics Part 3: Comparison of Multiple Strains (This entire section is more advanced and can be omitted; this section just solidifies the finding that a single amino acid change results in insecticide resistance. If you have time to include this material, it provides very powerful evidence that students are on the right track!)

A. Comparing pre-aligned DNA sequences from 30 strains (19 sensitive, 11 resistant to insecticide)
   1. Students search 30 aligned DNA sequences to locate differences. The SLA-B strain is used as the standard for comparison. Note that most of these sequences only contain exon 3; SLA-B and SR contain the complete gene sequence. Only exon 3 was amplified in the other strains once it was determined that the only amino acid difference between SLA-B and SR was in exon 3.
   2. After all differences are marked with a highlighter, students should locate any differences that are found in ALL of the resistant strains and NONE of the sensitive strains. They will identify one difference at base #739. Students can calculate the amino acid by dividing the nucleotide number by 3 (3 nucleotides per codon), and rounding the value to the next whole number if it is not evenly divisible by 3. They will discover that the amino acid (#247) is the ONLY amino acid difference identified between SLA-B and SR.

B. Aligning the 30 strains using Bioserver
   1. After logging into the Bioserver website (http://www.bioservers.org/bioserver/), students can select additional acetylcholinesterase sequences of strains that are sensitive and resistant for alignment by opening the drop down box(es) that say 'none' and selecting additional sequences. Each time a new sequence is opened in the list, a new drop down box appears.
   2. There are 29 strains sensitive, and 11 strains resistant, to inhibitors. The BioServer only allows eight sequences to be aligned at once, so must complete the alignment in segments, or instructors may assign the alignment of only some sequences. It is best to always include both SLA-B and SR in all alignments.
   3. After sequences are selected, click 'Compare" Align Clustal W. This will align the selected sequences. Note: if you print directly from the output, the color highlights typically don't transfer. You may avoid this problem by cutting and pasting the results into a word processing document, then printing the new document.

VI Bioinformatics Part 4: Understanding the Genetic Code. This section could easily be omitted; it focuses on the location of each of the mutations within the codon and develops an understanding of both the degeneracy of the code and the wobble hypothesis.

A. Using a prepared chart to identify the location of the mutations within the codon
   1. The file 'analysis of location of mutations within codons' provides a table listing all 28 nucleotide differences in the acetylcholinesterase gene between SLA-B and SR strains, listing both the nucleotide number and the codon containing the change. (The codons for SLA-B are listed.) The variant nucleotide is printed in red. The single change, resulting a different amino acid, is highlighted in yellow.
   2. This table can be used to discuss the location of each mutation in the codon (most are in the third base). Students can compare the codons and the amino acids for which they code. This is a perfect opportunity to reinforce the concept of the redundancy of the genetic code. Most of the mutations occur in the third codon; two of the three codons for which the mutation is in the first base, also code for the same amino acid.

B. Preparing their own chart to identify the location of the mutations within the codon
   1. Students can use the data they collected earlier to generate a list of the nucleotide differences and the codons in which each mutation occurs. They will need BOTH the aligned SLA-B and SR DNA sequences, as well as the translated sequences with the amino acids list.
   2. Using the DNA alignment output generated in section II.B.3, highlight the differences in the DNA sequences in resistant and sensitive strands on the translation output sheets generated in section III.B.1.
   3. The top reading frame of the protein sequences is the actual protein sequence. This particular translation program prints the one letter amino acid code directly under the middle base of the codon. Students can thus identify the codon that includes each of the mutations, as well as the amino acid for which it codes. This information can be tabulated; a copy of a suitable data table is found in 'data table VB3'.

VII Protein Structure: Determining the Structure-Function Relationship of the Mutation

A. Utilizing computer visualizations and physical models
   1. Students can now access the computer visualizations again at http://www.scripps.edu/mb/goodsell/jmol/ach/. They may wish to review the 'basics' and 'active

site' links before going on to the 'mutant' section. 'Mutant' focuses on comparing the active site of the wild type (sensitive to insecticide) and the mutant (resistant to insecticide). Both substrate (acetylcholine) and inhibitor are observed in the active site gorge. Students should discover that although both molecules fit in the active site in the wild type protein, the inhibitor is too large to fit in the pocket with the larger serine residue in the mutant protein.

2. After exploring the computer visualizations, students can return to the physical models to study the impact of the mutation on the ability of the substrate and inhibitor to fit in the active site gorge. They will quickly discover that the substrate is small enough to fit in the active site whether the normal or mutant residue is present, but that the inhibitor is too large to fit in the pocket in the presence of the mutation.

B. Proposing a new insecticide structure

1. After students have worked with the computer visualizations and physical models, it is very easy to transition into applying their knowledge. Students can be asked to describe characteristics of an inhibitor that will still function in the presence of the mutation. There are several avenues they may explore – including, but not limited to, a smaller inhibitor, an inhibitor that actually binds at the active site (which CAN'T mutate and still be functional), or perhaps an inhibitor that binds at the top of the pocket, blocking access to the active site.

Most of these teaching points have been developed into a Waksman Challenge for high school students, which was offered in spring 2006. The link to the Waksman Challenge website is: http://wakschallenge.rutgers.edu/index.php. The acetylcholinesterase challenge will eventually appear under the 'archives'. (There is a copy of the challenge on the CD, in the folder 'Waksman Challenge' under 'Insecticide Resistance'.)

Bibliography

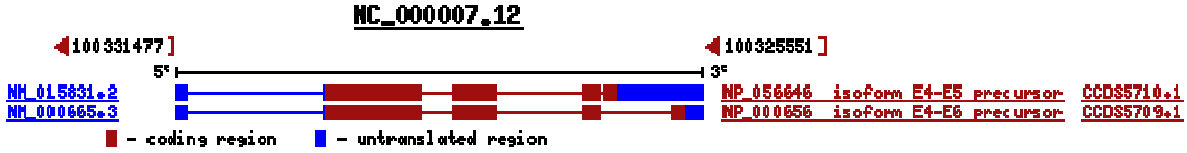
Weill, Mylene, Georges Lutfalla, Knud Morgensen, Fabrice Chandre, Arnaud Berthomieu, Claire Berticat, Nicole Pasteur, Alexandre Philips, Philippe Fort and Michel Raymond. 2003. Insecticide Resistance in Mosquito Vectors. Nature 423: 136-137. Available free online at: http://www-eve.ucdavis.edu/eve102/pdf/Weill_etal_Nature_2003.pdf

Acetylcholinesterase
Information taken from NCBI website; gene information on Acetylcholinesterase – gene ACHE
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene&cmd=Retrieve&dopt=full_report&list_uids=43

Acetylcholinesterase hydrolyzes the neurotransmitter, acetylcholine at neuromuscular junctions and brain cholinergic synapses, and thus terminates signal transmission. It is also found on the red blood cell membranes, where it constitutes the Yt blood group antigen. Acetylcholinesterase exists in multiple molecular forms which possess similar catalytic properties, but differ in their oligmeric assembly and mode of cell attachment to the cell surface. It is encoded by the single ACHE gene, and the structural diversity in the gene products arises from alternative mRNA splicing, and post-translational associations of catalytic and structural subunits. The major form of acetylcholinesterase found in brain, muscle and other tissues is the hydrophilic species, which forms disulfide-linked oligomers with collagenous, or lipid-containing structural subunits. The other, alternatively spliced form, expressed primarily in the erythroid tissues, differs at the C-terminal end, and contains a cleavable hydrophobic peptide with a GPI-anchor site. It associates with the membranes through the phosphoinositide (PI) moieties added post-translationally.

Genomic Regions, Transcripts and Products

(minus strand) Go to reference sequence details



SLA-B DNA sequence

ATGGAGATCCGAGGCCTAATAACCCGATTACTGGGTCCATGTCACCTGCGACATCTGATACT
GTGCAGTTTGGGGCTGTACTCCATCCTCGTGCAGTCGGTCCATTGCCGGCATCATGACATC
GGTAGTTCGGTGGCACACCAGCTAGGATCGAAATACTCACAATCATCCTCGTTATCGTCATC
CTCGCAATCGTCATCGTCGTTAGCTGAAGAGGCCACGCTGAATAAAGATTCAGATGCATTTT
TTACACCATATATAGGTCACGGAGATTCTGTTCGAATTGTAGATGCCGAATTAGGTACATTAG
AGCGCGAGCATATCCATAGCACTACGACCCGGCGGCGTGGCCTGACCCGGAGGGAGTCCA
GCTCCGATGCCACCGACTCGGACCCACTGGTCATAACGACGGACAAGGGCAAAATCCGTG
GAACGACACTGGAAGCGCCTAGTGGAAAGAAGGTGGACGCATGGATGGGCATTCCGTACG
CGCAGCCCCCGCTGGGTCCGCTCCGGTTTCGACATCCGCGACCCGCCGAAAGATGGACCG
GTGTGCTGAACGCGACCAAACCGCCCAACTCCTGCGTCCAGATCGTGGACACCGTGTTCGG
TGACTTCCCGGGGGCCACCATGTGGAACCCGAACACACCGCTCTCGGAGGACTGTCTGTAC
ATCAACGTGGTCGTGCCACGGCCCAGGCCCAAGAATGCCGCCGTCATGCTGTGGATCTTCG
GGGGTGGCTTCTACTCCGGGACTGCCACGCTGGACGTGTACGACCATCGGACGCTGGCCT
CGGAGGAGAACGTGATCGTAGTTTCGCTGCAGTACCGTGTCGCAAGTCTTGGGTTTCTCTTC
CTGGGCACACCGGAGGCACCCGGTAACGCGGGGCTGTTTGATCAGAACCTGGCACTGAGA
TGGGTCCGCGACAACATCCACCGGTTCGGCGGTGACCCCTCGCGGGTCACACTGTTCGGC
GAGAGCGCCGGAGCGGTCTCGGTTTCGCTGCACCTGCTGTCGGCGCTCTCGCGGGACCTG
TTCCAGCGGGCCATCCTCCAGAGTGGCTCCCCGACGGCCCCGTGGGCGCTGGTTTCGCGC
GAAGAAGCTACGCTTAGAGCTCTTCGTCTGGCCGAGGCCGTCAACTGTCCGCACGATGCGA
CCAAGCTGAGCGATGCCGTCGAATGCCTGCGAACCAAGGATCCGAACGAGCTGGTCGACA
ACGAGTGGGGCACGCTGGGGATCTGCGAGTTTCCGTTCGTTCCGGTTGTGGACGGAGCCT
TCCTCGATGAGACACCGCAGCGTTCGTTGGCCAGCGGGCGCTTCAAGAAAACGGACATCCT
GACCGGCAGCAACACCGAGGAGGGTTACTACTTTATCATTTACTATCTAACCGAACTGCTCA
GGAAAGAGGAAGGGGTCACGGTAACACGCGAGGAGTTCCTACAGGCCGTCCGGGAGTTGA
ATCCGTACGTGAACGGTGCCGCCCGGCAGGCCATCGTGTTCGAGTACACGGACTGGATTGA
ACCGGACAACCCGAACAGCAACCGTGACGCGCTGGACAAGATGGTCGGGGATTATCACTTC
ACCTGCAACGTGAACGAATTCGCCCAGCGGTACGCCGAGGAGGGCAACAACGTGTTCATGT
ACCTGTACACGCACAGAAGCAAAGGAAATCCCTGGCCGAGGTGGACCGGCGTGATGCACG
GCGACGAGATCAACTACGTGTTTGGCGAACCGCTGAACTCGGCCCTCGGCTACCAGGACGA
CGAGAAGGACTTTAGCCGGAAAATTATGCGATACTGGTCCAACTTTGCCAAGACTGGCAATC
CCAACCCGAGTACGCCGAGCGTGGACCTGCCCGAATGGCCCAAGCACACCGCCCACGGAC
GACACTATCTGGAGCTGGGACTGAACACGACCTTCGTGGGACGGGGCCCACGATTGCGGC
AGTGCGCTTTCTGGAAGAAATATTTGCCGCAACTAGTAGCAGCTACCTCTAACCTCCAAGTA
ACTCCCGCGCCTAGCGTACCTTGCGAAAGCAGCTCAACATCTTATCGATCCACTCTACTTCT
AATAGTCACACTACTTTTAGTAACGCGGTTCAAGATTTAA

SLA-B Protein Sequence

MEIRGLITRLLGPCHLRHLILCSLGLYSILVQSVHCRHHDIGSSVAHQLG 50
SKYSQSSSLSSSSQSSSSLAEEATLNKDSDAFFTPYIGHGDSVRIVDAEL 100
GTLEREHIHSTTTRRRGLTRRESSSDATDSDPLVITTDKGKIRGTTLEAP 150
SGKKVDAWMGIPYAQPPLGPLRFRHPRPAERWTGVLNATKPPNSCVQIVD 200
TVFGDFPGATMWNPNTPLSEDCLYINVVVPRPRPKNAAVMLWIFGGGFYS 250
GTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQ 300
NLALRWVRDNIHRFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAI 350
LQSGSPTAPWALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDP 400
NELVDNEWGTLGICEFPFVPVVDGAFLDETPQRSLASGRFKKTDILTGSN 450
TEEGYYFIIYYLTELLRKEEGVTVTREEFLQAVRELNPYVNGAARQAIVF 500
EYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVFMYLY 550
THRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDDEKDFSRKIMRYW 600
SNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQC 650
AFWKKYLPQLVAATSNLQVTPAPSVPCESSSTSYRSTLLLIVTLLLVTRF 700
KI*

SR DNA Sequence

ATGGAGATCCGAGGCCTAATAACCCGATTACTGGGTCCATGTCACCTGCGACATCTGATACT
GTGCAGTTTGGGGCTGTACTCCATCCTCGTGCAGTCGGTCCATTGCCGGCATCATGACATC
GGTAGTTCGGTGGCACACCAGCTAGGATCGAAATACTCACAATCATCCTCGTTATCGTCATC
CTCGCAATCGTCATCGTCGTTAGCTGAAGAGGCCACGCTGAATAAAGATTCAGATGCATTTT
TTACACCATATATAGGTCACGGAGATTCTGTTCGAATTGTAGATGCCGAATTAGGTACATTAG
AGCGCGAGCATATCCATAGCACTACGACCCGGCGGCGTGGCCTGACCCGGAGGGAGTCCA
GCTCCGATGCCACCGACTCGGACCCACTGGTAATAACGACGGACAAGGGCAAAATCCGTGG
AACGACACTGGAAGCGCCAAGTGGAAAGAAGGTGGACGCATGGATGGGCATTCCGTACGC
GCAGCCCCCGCTGGGTCCGCTCCGGTTTCGACATCCGCGACCCGCCGAAAGATGGACCGG
TGTGCTGAACGCGACCAAACCACCCAACTCCTGCGTCCAGATCGTGGACACCGTGTTCGGT
GACTTCCCGGGCGCGACCATGTGGAACCCGAACACACCCCTCTCGGAGGACTGTCTGTACA
TCAACGTGGTCGTGCCAAGGCCGAGGCCCAAGAATGCCGCTGTCATGCTGTGGATCTTTGG
GGGTAGCTTCTACTCCGGGACTGCCACGTTGGACGTGTACGATCATCGGACGCTGGCCTCG
GAGGAGAACGTGATCGTGGTTTCGCTGCAGTACCGTGTCGCAAGTCTTGGTTTTCTCTTCCT
GGGCACTCCGGAGGCACCTGGTAACGCGGGGCTGTTTGATCAGAACCTGGCACTGAGATG
GGTCCGCGACAACATCCACCGGTTCGGCGGTGACCCCTCGCGGGTCACACTGTTCGGCGA
GAGCGCCGGAGCGGTCTCGGTTTCGCTGCACCTGCTGTCGGCGCTCTCGCGGGACCTGTT
CCAGCGGGCCATCCTCCAGAGTGGCTCCCCGACGGCCCCATGGGCGCTGGTTTCGCGCGA
AGAAGCTACGCTTAGAGCTCTTCGTCTGGCCGAGGCCGTCAACTGTCCGCACGATGCGACC
AAGCTGAGCGATGCCGTCGAATGTCTGCGAACCAAGGATCCGAACGAGCTGGTCGACAATG
AGTGGGGCACGCTGGGGATCTGCGAGTTTCCGTTCGTTCCGGTTGTGGACGGTGCCTTCCT
CGATGAGACACCGCAGCGTTCGTTGGCCAGCGGTCGCTTCAAGAAAACGGACATCCTGACC
GGCAGCAACACCGAGGAGGGTTACTACTTTATCATTTACTATCTAACCGAACTGCTCAGGAA
AGAGGAAGGGGTCACGGTAACACGCGAGGAGTTCCTACAGGCCGTCCGGGAGTTGAATCC
GTACGTGAACGGTGCCGCCCGGCAGGCCATCGTGTTCGAGTACACGGACTGGATCGAACC
GGACAACCCGAACAGCAACCGTGACGCGCTCGACAAGATGGTCGGGGATTATCACTTCACC
TGCAACGTGAACGAGTTCGCCCAGCGGTACGCCGAGGAGGGCAACAATGTGTTCATGTACC
TGTACACGCACAGAAGCAAAGGAAATCCCTGGCCGAGGTGGACTGGCGTGATGCACGGCG
ACGAGATCAACTACGTGTTTGGCGAACCGCTGAACTCGGCCCTCGGCTACCAGGACGACGA
GAAGGACTTTAGCCGGAAAATTATGCGATACTGGTCCAACTTTGCCAAGACTGGCAATCCAA
ACCCGAGTACGCCGAGCGTGGACCTGCCCGAATGGCCCAAGCACACCGCCCACGGACGAC
ACTATCTGGAGCTGGGACTGAACACGACCTTCGTGGGACGGGGCCCACGATTGCGGCAGT
GCGCTTTCTGGAAGAAATATTTGCCGCAACTAGTAGCAGCTACCTCTAACCTCCAAGTAACT
CCCGCGCCTAGCGTACCTTGCGAAAGCAGCTCAACATCTTATCGATCCACTCTACTTCTAAT
AGTCACACTACTTTTAGTAACGCGGTTCAAGATTTAA

SR Protein Sequence

MEIRGLITRLLGPCHLRHLILCSLGLYSILVQSVHCRHHDIGSSVAHQLG 50
SKYSQSSSLSSSSQSSSSLAEEATLNKDSDAFFTPYIGHGDSVRIVDAEL 100
GTLEREHIHSTTTRRRGLTRRESSSDATDSDPLVITTDKGKIRGTTLEAP 150
SGKKVDAWMGIPYAQPPLGPLRFRHPRPAERWTGVLNATKPPNSCVQIVD 200
TVFGDFPGATMWNPNTPLSEDCLYINVVVPRPRPKNAAVMLWIFGGSFYS 250
GTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQ 300
NLALRWVRDNIHRFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAI 350
LQSGSPTAPWALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDP 400
NELVDNEWGTLGICEFPFVPVVDGAFLDETPQRSLASGRFKKTDILTGSN 450
TEEGYYFIIYYLTELLRKEEGVTVTREEFLQAVRELNPYVNGAARQAIVF 500
EYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVFMYLY 550
THRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDDEKDFSRKIMRYW 600
SNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQC 650
AFWKKYLPQLVAATSNLQVTPAPSVPCESSSTSYRSTLLLIVTLLLVTRF 700
KI*

```
SLA-B  ATGGAGATCCGAGGCCTAATAACCC GATTACTGGGTCCATGTCACCTGCG
SR     ATGGAGATCCGAGGCCTAATAACCC GATTACTGGGTCCATGTCACCTGCG  50

SLA-B  ACATCTGATACTGTGCAGTTTGGGG CTGTACTCCATCCTCGTGCAGTCGG
SR     ACATCTGATACTGTGCAGTTTGGGG CTGTACTCCATCCTCGTGCAGTCGG 100

SLA-B  TCCATTGCCGGCATCATGACATCGG TAGTTCGGTGGCACACCCAGCTAGGA
SR     TCCATTGCCGGCATCATGACATCGG TAGTTCGGTGGCACACCCAGCTAGGA 150

SLA-B  TCGAAATACTCACAATCATCCTCGT TATCGTCATCCTCGCAATCGTCATC
SR     TCGAAATACTCACAATCATCCTCGT TATCGTCATCCTCGCAATCGTCATC 200

SLA-B  GTCGTTAGCTGAAGAGGCCACGCTG AATAAAGATTCAGATGCATTTTTTA
SR     GTCGTTAGCTGAAGAGGCCACGCTG AATAAAGATTCAGATGCATTTTTTA 250

SLA-B  CACCATATATAGGTCACGGAGATTC TGTTCGAATTGTAGATGCCGAATTA
SR     CACCATATATAGGTCACGGAGATTC TGTTCGAATTGTAGATGCCGAATTA 300

SLA-B  GGTACATTAGAGCGCGAGCATATCC ATAGCACTACGACCCGGCGGCGTGG
SR     GGTACATTAGAGCGCGAGCATATCC ATAGCACTACGACCCGGCGGCGTGG 350

SLA-B  CCTGACCCGGAGGGAGTCCAGCTCC GATGCCACCGACTCGGACCCACTGG
SR     CCTGACCCGGAGGGAGTCCAGCTCC GATGCCACCGACTCGGACCCACTGG 400

SLA-B  TCATAACGACGGACAAGGGCAAAAT CCGTGGAACGACACTGGAAGCGCCT
SR     TAATAACGACGGACAAGGGCAAAAT CCGTGGAACGACACTGGAAGCGCCA 450

SLA-B  AGTGGAAAGAAGGTGGACGCATGGA TGGGCATTCCGTACGCGCAGCCCCC
SR     AGTGGAAAGAAGGTGGACGCATGGA TGGGCATTCCGTACGCGCAGCCCCC 500

SLA-B  GCTGGGTCCGCTCCGGTTTCGACAT CCGCGACCCGCCCGAAAGATGGACCG
SR     GCTGGGTCCGCTCCGGTTTCGACAT CCGCGACCCGCCCGAAAGATGGACCG 550

SLA-B  GTGTGCTGAACGCGACCAAACCGCC CAACTCCTGCGTCCAGATCGTGGAC
SR     GTGTGCTGAACGCGACCCAAACCACC CAACTCCTGCGTCCAGATCGTGGAC 600

SLA-B  ACCGTGTTCGGTGACTTCCCGGGGG CCACCATGTGGAACCCGAACACACC
SR     ACCGTGTTCGGTGACTTCCCGGGGC CGACCATGTGGAACCCGAACACACC 650

SLA-B  GCTCTCGGAGGACTGTCTGTACATC AACGTGGTCGTGCCACGGCCCAGGC
SR     CCTCTCGGAGGACTGTCTGTACATC AACGTGGTCGTGCCAAGGCCGAGGC 700

SLA-B  CCAAGAATGCCGCCGTCATGCTGTG GATCTTCGGGGGTGGCTTCTACTCC
SR     CCAAGAATGCCGCTGTCATGCTGTG GATCTTTGGGGGTAGCTTCTACTCC 750
```

```
SLA-B  GGGACTGCCACGCTGGACGTGTACG ACCATCGGACGCTGGCCTCGGAGGA
SR     GGGACTGCCACGTTGGACGTGTACG ATCATCGGACGCTGGCCTCGGAGGA  800

SLA-B  GAACGTGATCGTAGTTTCGCTGCAG TACCGTGTCGCAAGTCTTGGGTTTC
SR     GAACGTGATCGTGGTTTCGCTGCAG TACCGTGTCGCAAGTCTTGGTTTTC  850

SLA-B  TCTTCCTGGGCACACCGGAGGCACC CGGTAACGCGGGGCTGTTTGATCAG
SR     TCTTCCTGGGCACTCCGGAGGCACC TGGTAACGCGGGGCTGTTTGATCAG  900

SLA-B  AACCTGGCACTGAGATGGGTCCGCG AACAACATCCACCGGTTCGGCGGTGA
SR     AACCTGGCACTGAGATGGGTCCGCG AACAACATCCACCGGTTCGGCGGTGA  950

SLA-B  CCCCTCGCGGGTCACACTGTTCGGC GAGAGCGCCGGAGCGGTCTCGGTTT
SR     CCCCTCGCGGGTCACACTGTTCGGC GAGAGCGCCGGAGCGGTCTCGGTTT 1000

SLA-B  CGCTGCACCTGCTGTCGGCGCTCTC GCGGGACCTGTTCCAGCGGGCCATC
SR     CGCTGCACCTGCTGTCGGCGCTCTC GCGGGACCTGTTCCAGCGGGCCATC 1050

SLA-B  CTCCAGAGTGGCTCCCCGACGGCCC CGTGGGCGCTGGTTTCGCGCGAAGA
SR     CTCCAGAGTGGCTCCCCGACGGCCC CATGGGCGCTGGTTTCGCGCGAAGA 1100

SLA-B  AGCTACGCTTAGAGCTCTTCGTCTG GCCGAGGCCGTCAACTGTCCGCACG
SR     AGCTACGCTTAGAGCTCTTCGTCTG GCCGAGGCCGTCAACTGTCCGCACG 1150

SLA-B  ATGCGACCAAGCTGAGCGATGCCGT CGAATGCCTGCGAACCAAGGATCCG
SR     ATGCGACCAAGCTGAGCGATGCCGT CGAATGTCTGCGAACCAAGGATCCG 1200

SLA-B  AACGAGCTGGTCGACAACGAGTGGG GCACGCTGGGGATCTGCGAGTTTCC
SR     AACGAGCTGGTCGACAATGAGTGGG GCACGCTGGGGATCTGCGAGTTTCC 1250

SLA-B  GTTCGTTCCGGTTGTGGACGGAGCC TTCCTCGATGAGACACCGCAGCGTT
SR     GTTCGTTCCGGTTGTGGACGGTGCC TTCCTCGATGAGACACCGCAGCGTT 1300

SLA-B  CGTTGGCCAGCGGGCGCTTCAAGAA AACGGACATCCTGACCGGCAGCAAC
SR     CGTTGGCCAGCGGGTCGCTTCAAGAA AACGGACATCCTGACCGGCAGCAAC 1350

SLA-B  ACCGAGGAGGGTTACTACTTTATCA TTTACTATCTAACCGAACTGCTCAG
SR     ACCGAGGAGGGTTACTACTTTATCA TTTACTATCTAACCGAACTGCTCAG 1400

SLA-B  GAAAGAGGAAGGGGTCACGGTAACA CGCGAGGAGTTCCTACAGGCCGTCC
SR     GAAAGAGGAAGGGGTCACGGTAACA CGCGAGGAGTTCCTACAGGCCGTCC 1450

SLA-B  GGGAGTTGAATCCGTACGTGAACGG TGCCGCCCGGCAGGCCATCGTGTTC
SR     GGGAGTTGAATCCGTACGTGAACGG TGCCGCCCGGCAGGCCATCGTGTTC 1500
```

```
SLA-B  G A G T A C A C G G A C T G G A T T G A A C C G G     A C A A C C C G A A C A G C A A C C G T G A C G C
SR     G A G T A C A C G G A C T G G A T C G A A C C G G     A C A A C C C G A A C A G C A A C C G T G A C G C  1550

SLA-B  G C T G G A C A A G A T G G T C G G G G A T T A T     C A C T T C A C C T G C A A C G T G A A C G A A T
SR     G C T C G A C A A G A T G G T C G G G G A T T A T     C A C T T C A C C T G C A A C G T G A A C G A G T  1600

SLA-B  T C G C C C A G C G G T A C G C C G A G G A G G G     C A A C A A C G T G T T C A T G T A C C T G T A C
SR     T C G C C C A G C G G T A C G C C G A G G A G G G     C A A C A A T G T G T T C A T G T A C C T G T A C  1650

SLA-B  A C G C A C A G A A G C A A A G G A A A T C C C T     G G C C G A G G T G G A C C G G C G T G A T G C A
SR     A C G C A C A G A A G C A A A G G A A A T C C C T     G G C C G A G G T G G A C T G G C G T G A T G C A  1700

SLA-B  C G G C G A C G A G A T C A A C T A C G T G T T T     G G C G A A C C G C T G A A C T C G G C C C T C G
SR     C G G C G A C G A G A T C A A C T A C G T G T T T     G G C G A A C C G C T G A A C T C G G C C C T C G  1750

SLA-B  G C T A C C A G G A C G A C G A G A A G G A C T T     T A G C C G G A A A A T T A T G C G A T A C T G G
SR     G C T A C C A G G A C G A C G A G A A G G A C T T     T A G C C G G A A A A T T A T G C G A T A C T G G  1800

SLA-B  T C C A A C T T T G C C A A G A C T G G C A A T C     C C A A C C C G A G T A C G C C G A G C G T G G A
SR     T C C A A C T T T G C C A A G A C T G G C A A T C     C A A A C C C G A G T A C G C C G A G C G T G G A  1850

SLA-B  C C T G C C C G A A T G G C C C A A G C A C A C C     G C C C A C G G A C G A C A C T A T C T G G A G C
SR     C C T G C C C G A A T G G C C C A A G C A C A C C     G C C C A C G G A C G A C A C T A T C T G G A G C  1900

SLA-B  T G G G A C T G A A C A C G A C C T T C G T G G G     A C G G G G C C C A C G A T T G C G G C A G T G C
SR     T G G G A C T G A A C A C G A C C T T C G T G G G     A C G G G G C C C A C G A T T G C G G C A G T G C  1950

SLA-B  G C T T T C T G G A A G A A A T A T T T G C C G C     A A C T A G T A G C A G C T A C C T C T A A C C T
SR     G C T T T C T G G A A G A A A T A T T T G C C G C     A A C T A G T A G C A G C T A C C T C T A A C C T  2000

SLA-B  C C A A G T A A C T C C C G C G C C T A G C G T A     C C T T G C G A A A G C A G C T C A A C A T C T T
SR     C C A A G T A A C T C C C G C G C C T A G C G T A     C C T T G C G A A A G C A G C T C A A C A T C T T  2050

SLA-B  A T C G A T C C A C T C T A C T T C T A A T A G T     C A C A C T A C T T T T A G T A A C G C G G T T C
SR     A T C G A T C C A C T C T A C T T C T A A T A G T     C A C A C T A C T T T T A G T A A C G C G G T T C  2100

SLA-B  A A G A T T T A A
SR     A A G A T T T A A                                 2125
```

```
SLA-B  ATGGAGATCCGAGGCCTAATAACCC GATTACTGGGTCCATGTCACCTGCG
SR     ATGGAGATCCGAGGCCTAATAACCC GATTACTGGGTCCATGTCACCTGCG  50

SLA-B  ACATCTGATACTGTGCAGTTTGGGG CTGTACTCCATCCTCGTGCAGTCGG
SR     ACATCTGATACTGTGCAGTTTGGGG CTGTACTCCATCCTCGTGCAGTCGG 100

SLA-B  TCCATTGCCGGCATCATGACATCGG TAGTTCGGTGGCACACCAGCTAGGA
SR     TCCATTGCCGGCATCATGACATCGG TAGTTCGGTGGCACACCAGCTAGGA 150

SLA-B  TCGAAATACTCACAATCATCCTCGT TATCGTCATCCTCGCAATCGTCATC
SR     TCGAAATACTCACAATCATCCTCGT TATCGTCATCCTCGCAATCGTCATC 200

SLA-B  GTCGTTAGCTGAAGAGGCCACGCTG AATAAAGATTCAGATGCATTTTTA
SR     GTCGTTAGCTGAAGAGGCCACGCTG AATAAAGATTCAGATGCATTTTTA 250

SLA-B  CACCATATATAGGTCACGGAGATTC TGTTCGAATTGTAGATGCCGAATTA
SR     CACCATATATAGGTCACGGAGATTC TGTTCGAATTGTAGATGCCGAATTA 300

SLA-B  GGTACATTAGAGCGCGAGCATATCC ATAGCACTACGACCCGGCGGCGTGG
SR     GGTACATTAGAGCGCGAGCATATCC ATAGCACTACGACCCGGCGGCGTGG 350

SLA-B  CCTGACCCGGAGGGGAGTCCAGCTCC GATGCCACCGACTCGGACCCACTGG
SR     CCTGACCCGGAGGGGAGTCCAGCTCC GATGCCACCGACTCGGACCCACTGG 400

SLA-B  TCATAACGACGGACAAGGGGCAAAAT CCGTGGAACGACACTGGAAGCGCCT
SR     TAATAACGACGGACAAGGGGCAAAAT CCGTGGAACGACACTGGAAGCGCCA 450

SLA-B  AGTGGAAAGAAGGTGGACGCATGGA TGGGCATTCCGTACGCGCAGCCCCC
SR     AGTGGAAAGAAGGTGGACGCATGGA TGGGCATTCCGTACGCGCAGCCCCC 500

SLA-B  GCTGGGTCCGCTCCGGTTTCGACAT CCGCGACCCGCCCGAAAGATGGACCG
SR     GCTGGGTCCGCTCCGGTTTCGACAT CCGCGACCCGCCCGAAAGATGGACCG 550

SLA-B  GTGTGCTGAACGCGACCCAAACCGCC CAACTCCTGCGTCCAGATCGTGGAC
SR     GTGTGCTGAACGCGACCCAAACCACC CAACTCCTGCGTCCAGATCGTGGAC 600

SLA-B  ACCGTGTTCGGTGACTTCCCCGGGGG CCACCCATGTGGAACCCGAACACACC
SR     ACCGTGTTCGGTGACTTCCCCGGGGC CGACCATGTGGAACCCGAACACACC 650

SLA-B  GCTCTCGGAGGACTGTCTGTACATC AACGTGGTCGTGCCACGGCCCAGGC
SR     CCTCTCGGAGGACTGTCTGTACATC AACGTGGTCGTGCCAAGGCCGAGGC 700

SLA-B  CCAAGAATGCCGCCGTCATGCTGTG GATCTTCGGGGGTGGCTTCTACTCC
SR     CCAAGAATGCCGCTGTCATGCTGTG GATCTTTGGGGGTAGCTTCTACTCC 750
```

```
SLA-B  GGGACTGCCACGCTGGACGTGTACG  ACCATCGGACGCTGGCCTCGGAGGA
SR     GGGACTGCCACGTTGGACGTGTACG  ATCATCGGACGCTGGCCTCGGAGGA  800

SLA-B  GAACGTGATCGTAGTTTCGCTGCAG  TACCGTGTCGCAAGTCTTGGGTTTC
SR     GAACGTGATCGTGGTTTCGCTGCAG  TACCGTGTCGCAAGTCTTGGTTTTC  850

SLA-B  TCTTCCTGGGCACACCGGAGGCACC  CGGTAACGCGGGGCTGTTTGATCAG
SR     TCTTCCTGGGCACTCCGGAGGCACC  TGGTAACGCGGGGCTGTTTGATCAG  900

SLA-B  AACCTGGCACTGAGATGGGTCCGCG  ACAACATCCACCGGTTCGGCGGTGA
SR     AACCTGGCACTGAGATGGGTCCGCG  ACAACATCCACCGGTTCGGCGGTGA  950

SLA-B  CCCCTCGCGGGTCACACTGTTCGGC  GAGAGCGCCGGAGCGGTCTCGGTTT
SR     CCCCTCGCGGGTCACACTGTTCGGC  GAGAGCGCCGGAGCGGTCTCGGTTT 1000

SLA-B  CGCTGCACCTGCTGTCGGCGCTCTC  GCGGGACCTGTTCCAGCGGGCCATC
SR     CGCTGCACCTGCTGTCGGCGCTCTC  GCGGGACCTGTTCCAGCGGGCCATC 1050

SLA-B  CTCCAGAGTGGCTCCCCGACGGCCC  CGTGGGCGCTGGTTTCGCGCGAAGA
SR     CTCCAGAGTGGCTCCCCGACGGCCC  CATGGGCGCTGGTTTCGCGCGAAGA 1100

SLA-B  AGCTACGCTTAGAGCTCTTCGTCTG  GCCGAGGCCGTCAACTGTCCGCACG
SR     AGCTACGCTTAGAGCTCTTCGTCTG  GCCGAGGCCGTCAACTGTCCGCACG 1150

SLA-B  ATGCGACCAAGCTGAGCGATGCCGT  CGAATGCCTGCGAACCAAGGATCCG
SR     ATGCGACCAAGCTGAGCGATGCCGT  CGAATGTCTGCGAACCAAGGATCCG 1200

SLA-B  AACGAGCTGGTCGACAACGAGTGGG  GCACGCTGGGGATCTGCGAGTTTCC
SR     AACGAGCTGGTCGACAATGAGTGGG  GCACGCTGGGGATCTGCGAGTTTCC 1250

SLA-B  GTTCGTTCCGGTTGTGGACGGACCC  TTCCTCGATGAGACACCGCAGCGTT
SR     GTTCGTTCCGGTTGTGGACGGTGCC  TTCCTCGATGAGACACCGCAGCGTT 1300

SLA-B  CGTTGGCCCAGCGGGCGCTTCAAGAA  AACGGACATCCTGACCGGCAGCAAC
SR     CGTTGGCCCAGCGGTCGCTTCAAGAA  AACGGACATCCTGACCGGCAGCAAC 1350

SLA-B  ACCGAGGAGGGTTACTACTTTATCA  TTTACTATCTAACCGAACTGCTCAG
SR     ACCGAGGAGGGTTACTACTTTATCA  TTTACTATCTAACCGAACTGCTCAG 1400

SLA-B  GAAAGAGGAAGGGGTCACGGTAACA  CGCGAGGAGTTCCTACAGGCCGTCC
SR     GAAAGAGGAAGGGGTCACGGTAACA  CGCGAGGAGTTCCTACAGGCCGTCC 1450

SLA-B  GGGAGTTGAATCCGTACGTGAACGG  TGCCGCCCGGCAGGCCATCGTGTTC
SR     GGGAGTTGAATCCGTACGTGAACGG  TGCCGCCCGGCAGGCCATCGTGTTC 1500
```

```
SLA-B  GAGTACACGGACTGGATTGAACCGG  ACAACCCGAACAGCAACCGTGACGC
SR     GAGTACACGGACTGGATCGAACCGG  ACAACCCGAACAGCAACCGTGACGC 1550

SLA-B  GCTGGACAAGATGGTCGGGGATTAT  CACTTCACCTGCAACGTGAACGAAT
SR     GCTCGACAAGATGGTCGGGGATTAT  CACTTCACCTGCAACGTGAACGAGT 1600

SLA-B  TCGCCCAGCGGTACGCCGAGGAGGG  CAACAACGTGTTCATGTACCTGTAC
SR     TCGCCCAGCGGTACGCCGAGGAGGG  CAACAATGTGTTCATGTACCTGTAC 1650

SLA-B  ACGCACAGAAGCAAAAGGAAATCCCT  GGCCGAGGTGGACCGGCGTGATGCA
SR     ACGCACAGAAGCAAAAGGAAATCCCT  GGCCGAGGTGGACTGGCGTGATGCA 1700

SLA-B  CGGCGACGAGATCAACTACGTGTTT  GGCGAACCGCCTGAACTCGGCCCTCG
SR     CGGCGACGAGATCAACTACGTGTTT  GGCGAACCGCCTGAACTCGGCCCTCG 1750

SLA-B  GCTACCAGGACGACGAGAAGGACTT  TAGCCGGAAAATTATGCGATACTGG
SR     GCTACCAGGACGACGAGAAGGACTT  TAGCCGGAAAATTATGCGATACTGG 1800

SLA-B  TCCAACTTTGCCAAGACTGGCAATC  CCAACCCGAGTACGCCGAGCGTGGA
SR     TCCAACTTTGCCAAGACTGGCAATC  CAACCCGAGTACGCCGAGCGTGGA 1850

SLA-B  CCTGCCCGAATGGCCCAAGCACACC  GCCCACGGACGACACTATCTGGAGC
SR     CCTGCCCGAATGGCCCAAGCACACC  GCCCACGGACGACACTATCTGGAGC 1900

SLA-B  TGGGACTGAACACGACCTTCGTGGG  ACGGGGCCCACGATTGCGGCAGTGC
SR     TGGGACTGAACACGACCTTCGTGGG  ACGGGGCCCACGATTGCGGCAGTGC 1950

SLA-B  GCTTTCTGGAAGAAATATTTGCCGC  AACTAGTAGCAGCTACCTCTAACCT
SR     GCTTTCTGGAAGAAATATTTGCCGC  AACTAGTAGCAGCTACCTCTAACCT 2000

SLA-B  CCAAGTAACTCCCGCGCCTAGCGTA  CCTTGCGAAAGCAGCTCAACATCTT
SR     CCAAGTAACTCCCGCGCCTAGCGTA  CCTTGCGAAAGCAGCTCAACATCTT 2050

SLA-B  ATCGATCCACTCTACTTCTAATAGT  CACACTACTTTTAGTAACGCGGTTC
SR     ATCGATCCACTCTACTTCTAATAGT  CACACTACTTTTAGTAACGCGGTTC 2100

SLA-B  AAGATTTAA
SR     AAGATTTAA                  2125
```

Alignment of SLA-B and SR protein sequences

99.9% identity in 702 residues overlap; Score: 3713.0; Gap frequency: 0.0%

```
SLA-B,          1 MEIRGLITRLLGPCHLRHLILCSLGLYSILVQSVHCRHHDIGSSVAHQLGSKYSQSSSLS
SR,             1 MEIRGLITRLLGPCHLRHLILCSLGLYSILVQSVHCRHHDIGSSVAHQLGSKYSQSSSLS

SLA-B,         61 SSSQSSSSLAEEATLNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHIHSTTTRRRGLTR
SR,            61 SSSQSSSSLAEEATLNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHIHSTTTRRRGLTR

SLA-B,        121 RESSSDATDSDPLVITTDKGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAE
SR,           121 RESSSDATDSDPLVITTDKGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAE

SLA-B,        181 RWTGVLNATKPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVVVPRPRPKNAAVM
SR,           181 RWTGVLNATKPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVVVPRPRPKNAAVM

SLA-B,        241 LWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQ
SR,           241 LWIFGGSFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQ

SLA-B,        301 NLALRWVRDNIHRFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPW
SR,           301 NLALRWVRDNIHRFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPW

SLA-B,        361 ALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEWGTLGICEFPFVP
SR,           361 ALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEWGTLGICEFPFVP

SLA-B,        421 VVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVTREEFL
SR,           421 VVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVTREEFL
```

```
SLA-B,      481 QAVRELNPYVNGAARQAIVFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAE
SR,         481 QAVRELNPYVNGAARQAIVFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAE

SLA-B,      541 EGNNVFMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDDEKDFSRKIMRYW
SR,         541 EGNNVFMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDDEKDFSRKIMRYW

SLA-B,      601 SNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQCAFWKKYLPQL
SR,         601 SNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQCAFWKKYLPQL

SLA-B,      661 VAATSNLQVTPAPSVPCESSSTSYRSTLLLIVTLLLVTRFKI
SR,         661 VAATSNLQVTPAPSVPCESSSTSYRSTLLLIVTLLLVTRFKI
```

Alignment of SLA-B and SR protein sequences - KEY

99.9% identity in 702 residues overlap; Score: 3713.0; Gap frequency: 0.0%

```
SLA-B,          1 MEIRGLITRLLGPCHLRHLILCSLGLYSILVQSVHCRHHDIGSSVAHQLGSKYSQSSSLS
SR,             1 MEIRGLITRLLGPCHLRHLILCSLGLYSILVQSVHCRHHDIGSSVAHQLGSKYSQSSSLS
                  ************************************************************

SLA-B,         61 SSSQSSSSLAEEATLNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHIHSTTTRRRGLTR
SR,            61 SSSQSSSSLAEEATLNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHIHSTTTRRRGLTR
                  ************************************************************

SLA-B,        121 RESSSDATDSDPLVITTDKGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAE
SR,           121 RESSSDATDSDPLVITTDKGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAE
                  ************************************************************

SLA-B,        181 RWTGVLNATKPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVVVPRPRPKNAAVM
SR,           181 RWTGVLNATKPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVVVPRPRPKNAAVM
                  ************************************************************

SLA-B,        241 LWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQ
SR,           241 LWIFGGSFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQ
                  ****** ********************************************************

SLA-B,        301 NLALRWVRDNIHRFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPW
SR,           301 NLALRWVRDNIHRFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPW
                  ************************************************************

SLA-B,        361 ALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEWGTLGICEFPFVP
SR,           361 ALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEWGTLGICEFPFVP
                  ************************************************************

SLA-B,        421 VVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVTREEFL
SR,           421 VVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVTREEFL
                  ************************************************************
```

```
SLA-B,      481 QAVRELNPYVNGAARQAIVFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAE
SR,         481 QAVRELNPYVNGAARQAIVFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAE
                ************************************************************

SLA-B,      541 EGNNVFMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDDEKDFSRKIMRYW
SR,         541 EGNNVFMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDDEKDFSRKIMRYW
                ************************************************************

SLA-B,      601 SNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQCAFWKKYLPQL
SR,         601 SNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQCAFWKKYLPQL
                ************************************************************

SLA-B,      661 VAATSNLQVTPAPSVPCESSSTSYRSTLLLIVTLLLVTRFKI
SR,         661 VAATSNLQVTPAPSVPCESSSTSYRSTLLLIVTLLLVTRFKI
                *********************************************
```

Analysis of location of mutations in SR acetylcholinesterase within codons

Acetylcholinesterase mRNA compared in SLA-B (sensitive to insecticides) and SR (resistant to insecticides). Nucleotides that differ between the two strands are listed by number. SLA-B condon sequences are listed on top, with SR sequences below, with the nucleotide which differs highlighted in red. Only one nucleotide difference, at base 739, results in a change in amino acid sequence, from glycine to serine. This residue is highlighted in yellow.

Nucleotide number	Codon	Nucleotide number	Codon	Nucleotide number	Codon	Nucleotide number	Codon
402	GUC GUA	696	CCC CCG	846	GGG GGU	1314	GGG GGU
450	CCU CCA	714	GCC GCU	864	ACA ACU	1518	AUU AUC
573	CCG CCA	732	UUC UUU	876	CCC CCU	1554	CUG CUC
624	GGG GGC	739	GGC AGC	1077	CCG CCA	1599	GAA GAG
627	GCC GCG	763	CUG UUG	1182	UGC UGU	1632	AAC AAU
651	CCG CCC	777	GAC GAU	1218	AAC AAU	1689	ACC ACU
691	CGG AGG	813	GUA GUG	1272	GGA GGU	1827	CCC CCA

```
Base     451
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S03      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S04      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S06 A    AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S06 B    AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S07      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S08      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S09      AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S10      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S11      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S12      AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S13      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S14      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S15      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S16      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S17      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S18      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S19      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
SLA-B    AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
Resist.
SR       AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R01      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R02      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R03      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R04      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R05      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R06      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R07      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R08      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R09      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R10      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
```

```
Base     511
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S03      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S04      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S06      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S06      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S07      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S08      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCAACCAAA
S09      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S10      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S11      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S12      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S13      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S14      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S15      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S16      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S17      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S18      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S19      CTCCGGTTTC GACATCCACG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
SLA-B    CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
Resist.
SR       CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R01      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R02      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R03      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R04      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R05      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R06      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R07      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R08      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R09      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R10      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
```

```
Base     571
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S03      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S04      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S06      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S06      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S07      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S08      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S09      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S10      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S11      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S12      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S13      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S14      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S15      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S16      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S17      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACAGTGTTCG GTGACTTCCC GGGGGCCACC
S18      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S19      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACAGTGTTCG GTGACTTCCC GGGGGCCACC
SLA-B    CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
Resist.
SR       CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGCGCGACC
R01      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R02      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R03      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R04      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R05      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R06      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R07      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R08      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R09      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R10      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
```

```
Base     631
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S03      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S04      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S06      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S06      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S07      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S08      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S09      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S10      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S11      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S12      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S13      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S14      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S15      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S16      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S17      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S18      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S19      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
SLA-B    ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
Resist.
SR       ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R01      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R02      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R03      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R04      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R05      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R06      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R07      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R08      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R09      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R10      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
```

```
Base     691
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S03      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S04      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S06      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S06      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S07      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S08      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S09      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S10      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S11      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S12      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S13      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTATTCC
S14      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S15      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
S16      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
S17      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S18      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
S19      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
SLA-B    CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
Resist.
SR       AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R01      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R02      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R03      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R04      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R05      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R06      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R07      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R08      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R09      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R10      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
```

```
Base     751
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S03      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGA CCTCGGAGGA GAACGTGATC
S04      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGA CCTCGGAGGA GAACGTGATC
S06      GGGACTGCCA CGCTGGACGT GTATGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S06      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S07      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S08      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S09      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S10      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S11      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S12      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S13      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S14      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S15      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S16      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S17      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S18      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S19      GGGACTGCCA CGTTGGACGT GTACGACCAT CGGACGCTGG CCTCGGAAGA GAACGTGATC
SLA-B    GGGACTGCCA CGCTGGACGT GTACGACCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
Resist.
SR       GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R01      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R02      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R03      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R04      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R05      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R06      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R07      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R08      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R09      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R10      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
```

```
Base     811
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S03      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S04      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S06      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S06      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S07      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S08      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S09      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S10      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S11      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S12      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S13      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S14      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S15      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S16      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S17      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S18      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S19      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
SLA-B    GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGGTTTC TCTTCCTGGG CACACCGGAG
Resist.
SR       GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGTTTTC TCTTCCTGGG CACTCCGGAG
R01      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R02      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R03      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R04      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R05      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R06      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R07      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R08      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R09      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R10      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
```

```
Base     451
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S03      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S04      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S06      AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S06      AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S07      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S08      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S09      AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S10      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S11      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S12      AGTGGAAAGA AGGTGGACGC CTGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S13      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S14      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
S15      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S16      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S17      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S18      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
S19      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
SLA-B    AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
Resist.
SR       AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R01      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R02      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R03      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R04      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R05      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R06      AGCGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCTCC GCTGGGTCCG
R07      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R08      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R09      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
R10      AGTGGAAAGA AGGTGGACGC ATGGATGGGC ATTCCGTACG CGCAGCCCCC GCTGGGTCCG
```

```
Base     511
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S03      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S04      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S06      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S06      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S07      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S08      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCAACCAAA
S09      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S10      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S11      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S12      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S13      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S14      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S15      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S16      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S17      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S18      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
S19      CTCCGGTTTC GACATCCACG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
SLA-B    CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
Resist.
SR       CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R01      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R02      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R03      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R04      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R05      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R06      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R07      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R08      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R09      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
R10      CTCCGGTTTC GACATCCGCG ACCCGCCGAA AGATGGACCG GTGTGCTGAA CGCGACCAAA
```

```
Base     571
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S03      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S04      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S06      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S06      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S07      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S08      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S09      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S10      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S11      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S12      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S13      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S14      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S15      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S16      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S17      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACAGTGTTCG GTGACTTCCC GGGGGCCACC
S18      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
S19      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACAGTGTTCG GTGACTTCCC GGGGGCCACC
SLA-B    CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
Resist.
SR       CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGCGCGACC
R01      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R02      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R03      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R04      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R05      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R06      CCGCCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R07      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R08      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R09      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
R10      CCACCCAACT CCTGCGTCCA GATCGTGGAC ACCGTGTTCG GTGACTTCCC GGGGGCCACC
```

```
Base     631
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S03      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S04      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S06      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S06      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S07      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S08      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S09      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S10      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S11      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S12      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S13      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S14      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S15      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S16      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S17      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S18      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
S19      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
SLA-B    ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
Resist.
SR       ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R01      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R02      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R03      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R04      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R05      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R06      ATGTGGAACC CGAACACACC GCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R07      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R08      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R09      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
R10      ATGTGGAACC CGAACACACC CCTCTCGGAG GACTGTCTGT ACATCAACGT GGTCGTGCCA
```

```
Base     691
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S03      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S04      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S06      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S06      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S07      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S08      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S09      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S10      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S11      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S12      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S13      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTA TTCC
S14      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S15      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
S16      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
S17      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
S18      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
S19      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTGG CTTCTACTCC
SLA-B    CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTGG CTTCTACTCC
Resist.
SR       AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R01      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R02      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R03      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R04      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R05      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R06      CGGCCCAGGC CCAAGAATGC CGCCGTCATG CTGTGGATCT TCGGGGGTAG CTTCTACTCC
R07      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R08      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R09      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
R10      AGGCCGAGGC CCAAGAATGC CGCTGTCATG CTGTGGATCT TTGGGGGTAG CTTCTACTCC
```

```
Base     751
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S03      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGA CCTCGGAGGA GAACGTGATC
S04      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGA CCTCGGAGGA GAACGTGATC
S06      GGGACTGCCA CGCTGGACGT GTATGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S06      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S07      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S08      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S09      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S10      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S11      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S12      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S13      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S14      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
S15      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S16      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S17      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S18      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
S19      GGGACTGCCA CGTTGGACGT GTACGACCAT CGGACGCTGG CCTCGGAAGA GAACGTGATC
SLA-B    GGGACTGCCA CGCTGGACGT GTACGACCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
Resist.
SR       GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R01      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R02      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R03      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R04      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R05      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R06      GGGACTGCCA CGCTGGACGT GTACGACCAC CGGACGCTGG CCTCGGAGGA GAACGTGATC
R07      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R08      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R09      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
R10      GGGACTGCCA CGTTGGACGT GTACGATCAT CGGACGCTGG CCTCGGAGGA GAACGTGATC
```

```
Base     811
         1234567890 1234567890 1234567890 1234567890 1234567890 1234567890
Sens.
S01      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S03      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S04      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S06      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S06      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S07      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S08      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S09      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S10      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S11      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S12      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S13      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S14      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGG
S15      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S16      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S17      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S18      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
S19      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
SLA-B    GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGGTTTC TCTTCCTGGG CACACCGGAG
Resist.
SR       GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGTTTTC TCTTCCTGGG CACTCCGGAG
R01      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R02      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R03      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R04      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R05      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R06      GTAGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R07      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R08      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R09      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
R10      GTGGTTTCGC TGCAGTACCG TGTCGCAAGT CTTGGT
```

Summary of all nucleotide differences among 30 strains of mosquito

Base	471	498	528	553	573	603	651	691	696	714	732	739	763	774	777	813	846
Sens.																	
S01	A	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	G
S03	A	T	G	G	A	C	G	C	C	C	C	G	C	C	C	A	T
S04	A	T	G	G	A	C	G	C	C	C	C	G	C	C	C	A	T
S06 A	C	C	G	G	G	C	G	C	C	C	C	G	C	T	C	A	T
S06 B	C	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	G
S07	A	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	T
S08	A	C	G	A	G	C	G	C	C	C	C	G	C	C	C	A	G
S09	C	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	G
S10	A	T	G	G	A	C	G	C	C	C	C	G	C	C	C	A	T
S11	A	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	T
S12	C	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	G
S13	A	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	G
S14	A	T	G	G	G	C	G	C	C	C	C	G	C	C	C	A	G
S15	A	C	G	G	A	C	C	A	G	T	T	G	T	C	T	G	T
S16	A	C	G	G	A	C	C	A	G	T	T	G	T	C	T	G	T
S17	A	C	G	G	A	A	C	A	G	T	C	G	T	C	T	G	T
S18	A	C	G	G	A	C	C	A	G	T	T	G	T	C	T	G	T
S19	A	C	A	G	A	A	C	A	G	T	T	G	T	C	C	G	T
SLA-B	A	C	G	G	G	C	G	C	C	C	C	G	C	C	C	A	G
Resist.																	
SR	A	C	G	G	A	C	C	A	G	T	T	A	T	C	T	G	T
R01	A	T	G	G	G	C	G	C	C	C	C	A	C	C	C	A	T
R02	A	T	G	G	G	C	G	C	C	C	C	A	C	C	C	A	T
R03	A	C	G	G	A	C	C	A	G	T	T	A	T	C	T	G	T
R04	A	T	G	G	G	C	G	C	C	C	C	A	C	C	C	A	T
R05	A	T	G	G	G	C	G	C	C	C	C	A	C	C	C	A	T
R06	A	T	G	G	G	C	G	C	C	C	C	A	C	C	C	A	T
R07	A	C	G	G	A	C	C	A	G	T	T	A	T	C	T	G	T
R08	A	C	G	G	A	C	C	A	G	T	T	A	T	C	T	G	T
R09	A	C	G	G	A	C	C	A	G	T	T	A	T	C	T	G	T
R10	A	C	G	G	A	C	C	A	G	T	T	A	T	C	T	G	T

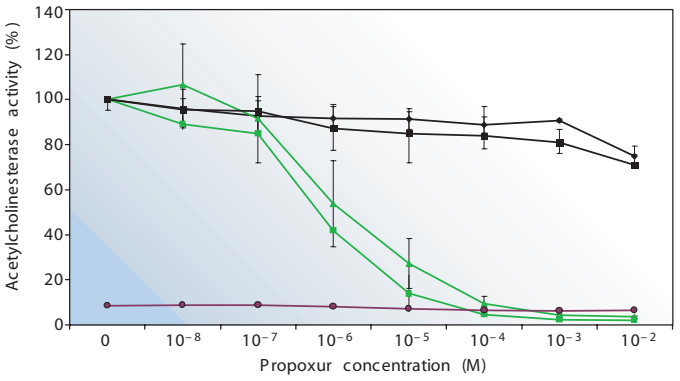


Figure 1 Residual acetylcholinesterase activity of susceptible (green squares) and resistant (black diamonds) mosquitoes assayed in homogenates and lysates from transfected S2 cells in the presence of increasing concentrations of the carbamate insecticide Propoxur. S2 cells were transfected with the recombinant pAc5.1/V5-His vector (Invitrogen) either alone (purple circles) or with expression of either sensitive acetylcholinesterase-1 (green triangles) or insensitive G119S-mutant enzyme (black squares). Residual enzyme activity was assayed after incubation with Propoxur for 15 min (ref. 5). Three independent experiments were carried out using different volumes of lysate. c

Comparative genomics

Insecticide resistance in mosquito vectors

Resistance to insecticides among mosquitoes that act as vectors for malaria (*Anopheles gambiae*) and West Nile virus (*Culex pipiens*) emerged more than 25 years ago in Africa, America and Europe; this resistance is frequently due to a loss of sensitivity of the insect's acetylcholinesterase enzyme to organophosphates and carbamates[1]. Here we show that this insensitivity results from a single amino-acid substitution in the enzyme, which we found in ten highly resistant strains of *C. pipiens* from tropical (Africa and Caribbean) and temperate (Europe) areas, as well as in one resistant African strain of *A. gambiae*. Our identification of this mutation may pave the way for designing new insecticides.

Acetylcholinesterase terminates synaptic transmission by hydrolysing the neurotransmitter acetylcholine; its inactivation by insecticides leads to paralysis and death. Mosquitoes, however, show widespread and strong resistance to this type of insecticide. They have two genes that encode different isoforms of acetylcholinesterase: *ace-1*, which has no homologue in the fruitfly *Drosophila melanogaster* and is closely linked to resistance in *C. pipiens*; and *ace-2*, a homologue of the unique *Drosophila ace* gene[2]. The generally mild insensitivity of acetylcholinesterase-2 in *D. melanogaster* is due to the combined weak effect of several mutations[3].

To identify mutations involved in resistance in mosquitoes, we determined the complete *ace-1* messenger RNA coding sequence of two *Culex pipiens* strains: one susceptible and one resistant (results not shown). *C. pipiens ace-1* encodes a putative 702-amino-acid protein, which is 81% identical to its *A. gambiae* homologue and 39% identical to *D. melanogaster* acetylcholinesterase-2. Complementary DNAs from the susceptible and resistant strains differ at 27 n cleotide positions, only one of which generates an amino-acid substitution: the GGC (glycine) codon at position 119, according to the nomenclature for *Torpedo* acetylcholinesterase[1], is replaced by an AGC (serine) codon in resistant mosquitoes (mutation G119S).

From three-dimensional modelling, we find that this mutated residue lies within the active 'gorge' of the enzyme, close to the catalytic site and abutting the oxyanion hole (results not shown). To evaluate the biochemical effect of the mutation *in vitro*, we assayed the catalytic properties and insecticidal sensitivity of wild-type and mutant recombinant acetylcholinesterase-1 that was expressed in S2 *Drosophila* cells. The G119S mutant showed the same insensitivity to Propoxur insecticide as resistant-strain acetylcholinesterase-1 (Fig. 1). A single mutation in *ace-1* must therefore be responsible for the insensitivity of the enzyme.

To determine whether the G119S mutation is present in other *C. pipiens* strains with insensitive acetylcholinesterase, we sequenced exon 3 of *ace-1* in several resistant and susceptible strains derived either from the temperate or the tropical/subtropical form of the *C. pipiens* species complex (*C. p. pipiens* and *C. p. quinquefasciatus*, respectively). All of the resistant strains carried the G119S substitution, regardless of their origin. Moreover, although 23 nucleotides were polymorphic, a unique haplotype was found to be associated with the resistance within each subspecies (see supplementary information). This indicates that the same G119S mutation has occurred independently at least twice in *C. pipiens*, once in each subspecies.

We also investigated the recent emergence of insensitive acetylcholinesterase in the main African malaria vector *Anopheles gambiae*[4], with the use of the *ace-1* genomic sequences of a resistant (YAO) and a susceptible (KISUMU) strain. The coding sequences differed at 18 nucleotide positions, two of them being non-synonymous. In the YAO strain, one mutation that resulted in the replacement of a valine residue by alanine in the amino-terminal region has no equivalent in *Torpedo* acetylcholinesterase and did not seem to affect the enzyme's catalytic properties (results not shown). The other was the same G119S substitution as in *C. pipiens* (results not shown), indicating that this single point mutation has occurred independently at least three times in the *ace-1* gene: twice in the *C. pipiens* complex and once in *A. gambiae*.

The discovery of the *ace-1* mutation that is responsible for insecticide resistance in mosquitoes opens the way to new strategies for pest management. The development of new insecticides that can specifically inhibit the G119S mutant form of acetylcholinesterase-1 will be crucial in overcoming the spread of resistance.

Mylène Weill*, Georges Lutfalla†, Knud Mogensen†, Fabrice Chandre‡, Arnaud Berthomieu*, Claire Berticat*, Nicole Pasteur*, Alexandre Philips§, Philippe Fort§, Michel Raymond*

**Institut des Sciences de l'Évolution (UMR 5554), CC 065, Université Montpellier II, 34095 Montpellier, France*
e-mail: weill@isem.univ-montp2.fr
†Défenses Antivirales et Tumorales (UMR 5124), and §Centre de Recherche en Biochimie des Macromolécules (UPR1086), CNRS, 34293 Montpellier, France
‡IRD/LIN, BP 64501, 34394 Montpellier, France

1. Toutant, J. P. *Progr. Neurobiol.* **32,** 423–446 (1989).
2. Weill, M. *et al. Proc. R. Soc. Lond. B* **269,** 2007–2016 (2002).
3. Mutéro, A., Pralavorio, M., Bride, J. M. & Fournier, D. *Proc. Natl Acad. Sci. USA* **91,** 5922–5926 (1994).
4. N'Guessan, R. *et al. Med. Vet. Entomol.* **17,** 1–7 (2002).
5. Bourguet, D., Pasteur, N., Bisset, J. & Raymond, M. *Pest. Biochem. Physiol.* **55,** 122–128 (1996).

Supplementary information accompanies this communication on *Nature*'s website.
Competing financial interests: declared none.

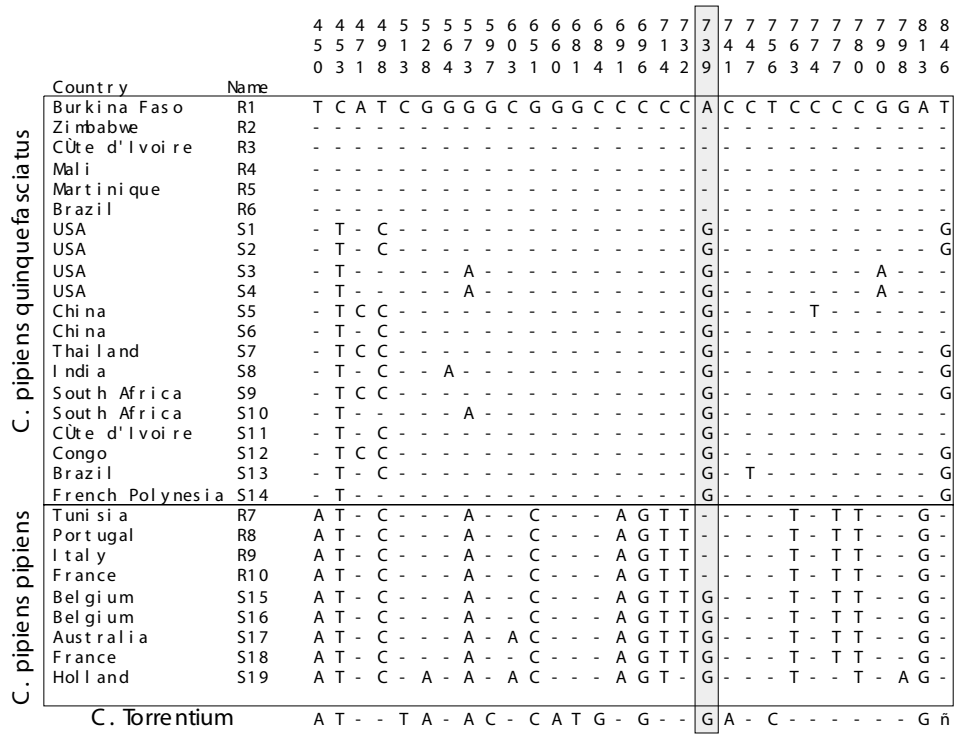


Figure S1. Nucleotide polymorphism in exon 3 of *ace-1* in *Culex pipiens* mosquitoes. Samples are listed according to subspecies (*C. p. pipiens* or *C. p. quinquefasciatus*), presence of insensitive AChE (Ri) or not (Si), and country of origin. Only polymorphic sites are indicated, and a dash indicates similarity with the top sequence. The position of the G119S mutation is boxed. *Culex torrentium*, the closest known species to the *C. pipiens* complex, is presented as an outgroup. See Supplementary Information for references for all samples.

Table S1. Name, country of origin and reference of the strains used in Figure S1.

Taxa	Name in Fig. S1	Real name	Country	Reference
C. p. quinquefasciatus	R1	BO	Burkina-Faso	M. Raymond, unpublished
	R2	HARARE	Zimbabwe	M. Raymond, unpublished
	R3	SUPERCAR	Côte d'Ivoire	1
	R4	DJI	Mali	M. Raymond, unpublished
	R5	MARTINIQUE	Martinique	2
	R6	RECIFE	Brazil	Collected in 1995 by A.-B. Failloux (Pasteur Institute, Paris, France)
	S1	PRO-R	USA	3
	S2	S-LAB	USA	3
	S3	TEM-R	USA	4
	S4	TRANS-P	USA	5
	S6	LING	China	6
	S7	MAO	China	7
	S6	THAI	Thailand	8
	S8	MADURAI	India	9
	S9	BSQ	South Arica	Collected in 1991 by A. J. Cornel (Sth Afr. Inst. Med. Res., South Africa)
	S10	BED	South Arica	Collected in 1991 by A. J. Cornel (Sth Afr. Inst. Med. Res., South Africa)
	S11	BOUAKE	Côte d'Ivoire	10
	S12	BRAZZA	Congo	11
	S13	BRESIL	Brazil	M. Raymond, unpublished
	S14	MOOREA	French Polynesia	12
C. p. pipiens	R7	ESPRO	Tunisia	13
	R8	PRAIAS	Portugal	14
	R9	PADOVA	Italy	15
	R10	BARRIOL	France	16
	S15	BRUGES-A	Belgium	17
	S16	BRUGES-B	Belgium	17
	S17	KILLCARE	Autralia	18
	S18	BLEUET	France	19
	S19	HETEREN	Holland	M. Raymond, unpublished
C. torrentium	-	UPPSALA	Sweden	20

1. Chandre, F. Thèse de Doctorat, Université Paris XII-Val de Marne (1998).
2. Bourguet, D., Raymond, M., Bisset, J., Pasteur, N. & Arpagaus, M. *Biochem. Genet.* **34**, 351-362 (1996).
3. Georghiou, G.P., Metcalf, R.L. & Gidden, F.E. *Bull. Wld Hlth Org.* **35**, 691-708 (1966).
4. Georghiou, G. & Pasteur, N. *J. Econ. Entomol.* **71**, 201-205 (1978).
5. Priester, T.M. & Georghiou, G.P. *J. Econ. Entomol.* **71**, 197-200 (1978).
6. Weill, M., *et al. Journal of the American Mosquito Control Association* **17**, 238-244 (2001).
7. Qiao, C.-L., Marquine, M., Pasteur, N. & Raymond, M. *Biochem. Genet.* **36**, 417-426 (1998).
8. Guillemaud, T., Rooker, S., Pasteur, N. & Raymond, M. *Heredity* **77**, 535-543 (1996).
9. Nielsen-Leroux, C., *et al. J. Med. Entomol.* **39**, 729-735 (2002).
10. Magnin, M., Marboutin, E. & Pasteur, N. *J. Med. Entomol.* **25**, 99-104 (1988).
11. Beyssat-Arnaouty, V. Thèse de Doctorat, Université de Montpellier II (1989).