Almost the whole of visual science is concerned with illusions — if they are defined as images in which the physical description differs from the image perceived. One of the attractions of illusions, apart from their obvious appeal as geometrical abstract art, is that they direct us to an understanding of visual processing at many levels.

But before the question of why an illusion is perceived can be addressed, we must have a physical description of the image, which is not always straightforward. After all, what is the physical description of a stick that is partially immersed in water? It is straight if the stick itself is measured, but not if a photograph is taken of it.

For this reason (among others) psychologists make a distinction between the distal (or physical) stimulus and the proximal stimulus (its projection onto the retina). An illusion occurs when there is a mismatch between the proximal stimulus and perception.

The distinction between distal and proximal stimuli can be appreciated with the image of black and white concentric circles below. The distal stimulus is circular and stable, but its projection onto the retina is distorted and dynamic. Spokes appear to radiate from the centre and to rotate; they are probably a consequence of slight astigmatic changes in the curvature of the lens, because they are reduced when viewed through a pinhole, they have similar temporal dynamics to measured fluctuations in the curvature of the lens, and they do not occur in those who have no crystalline lens (aphakia).

Aphakic individuals lacking a lens in one eye see such patterns as stable using that eye, but see it with radiating spokes using the eye that has a lens.

In addition, viewing the pattern from a distance will reveal a portrait of Jan Evangelista Purkinje, who first described the radiating spokes phenomenon in 1825. The portrait is carried by low spatial frequency variations in the pattern, which correspond to the coarse grain over the whole pattern; these are normally inhibited by the higher spatial frequencies (the fine grain of the thin lines) when the pattern is close to the eye. When viewed from a distance, the high spatial frequency content (thin lines) is not sharply resolved and the low spatial frequency pattern emerges. Purkinje also demonstrated similar effects with radiating lines (see coloured image above). High-contrast coloured or black and white radiating lines are also visually dynamic — they appear to move.
The motion consists of shimmering points which are perpendicular to the lines of the distal stimulus. Thus, with a radiating pattern they are circular, and can be seen more readily within the orange annuli. If a white surface is observed following fixation on the pattern for around 30 seconds then scintillating dots will seem to be streaming in a direction orthogonal to the previously presented lines.

Regular geometrical patterns produce a range of visual effects; some are based on the optics of the eye (like the spokes and others (such as the scintillations) could be caused by visual processing in the brain. For example, Purkinje’s radiating lines illusion has been examined with the aid of neural imaging. Positron emission tomography (PET) scans indicate that a pattern of radiating lines causes the blood flow in V5 — the area of the brain that is involved in motion processing — to increase. Alternative interpretations, however, relate to eye movement instabilities during fixation. It is likely that several phenomena are induced by these patterns, with a different explanation for each one.

When we know that an image is distorted before it reaches the retina — that the proximal stimulus differs from the distal stimulus — we consider that we have some interpretation of the illusion. Thus, the bent-stick distortion is due to differential refraction through air and water. But in general, explanations for the illusion are sought in higher levels of visual processing. Such illusions might provide an index of neural function, which I have called a ‘neuro-sign’.

Address: Department of Psychology, University of Dundee, Dundee DD1 4HN, U.K.

This is the first in a short series of articles by Nicholas Wade, describing some common visual illusions.

**Correspondence**

**A bacterial antibiotic resistance gene with eukaryotic origins**

James R. Brown*, Jianzhi Zhang† and John E. Hodgson‡§

The rapid emergence of antibiotic-resistant pathogens is mainly due to the horizontal transfer of resistance genes among different bacterial strains or species. No transfers of antibiotic resistance genes have been found to occur between bacteria and eukaryotes, however [1]. Here, we show that a gene known to convey antibiotic resistance in bacteria was apparently transferred from an early evolved eukaryote.

Mupirocin, a topical agent widely used since 1985 against skin surface and nasal colonization by the Gram-positive bacterium *Staphylococcus aureus*, is the natural compound produced by the bacterium *Pseudomonas fluorescens* NCIB 10586 [2]. Pseudomonic acid inhibits protein synthesis by reversibly binding to the active site of isoleucyl-tRNA synthetase (IRS) [3,4]. Sensitivity of *S. aureus* to mupirocin use *in vitro* occurs at a minimal inhibitory concentration (MIC) range of 0.12 to 1.0 mg/L [5,6]. Moderate mupirocin resistance in *S. aureus*, broadly defined as an MIC range of 8 to 256 mg/L, appears to be mediated by different single amino-acid substitutions in IRS, either within or near an ATP-binding motif with the consensus amino-acid sequence KMSKS (in the single amino-acid code) [7]. First observed clinically in the late 1980s, this level of resistance has little impact on the topical use of mupirocin.

In the early 1990s, much higher levels of mupirocin resistance in *S. aureus* appeared in some clinical settings, where estimated MICs exceed 512 mg/L [5]. Highly resistant strains have a second copy of the IRS gene (IRS-HR), located on an extrachromosomal plasmid [8], the product of which is highly divergent (34% amino-acid identity) from those of sensitive or moderately resistant IRS genes (IRS-S/MR) [9,10]. More recently, another Gram-positive bacterium, *Mycobacterium tuberculosis*, was shown to be highly resistant to mupirocin, and the product of its chromosome-located IRS gene had higher sequence similarity to *S. aureus* IRS-HR proteins than IRS-S/MR proteins [11].

Using the *S. aureus* IRS-HR sequence as a probe against public databases [12], we discovered four more bacterial IRS-HR-type genes. Phylogenetic analyses (Figure 1) show that all six bacterial IRS-HR proteins cluster significantly with eukaryotic IRS proteins that are also highly resistant to mupirocin [3,4]. The present IRS tree contradicts an earlier phylogeny, which showed the Bacteria, Archaea and Eucarya as separate monophyletic groups, with the latter two being sister groups [13]. We suggest that bacterial IRS-HR genes originated from eukaryotes through a horizontal gene transfer that occurred after the Eucarya and Archaea separated. A eukaryotic origin of bacterial IRS-HR genes is also supported by an insertion with the consensus sequence HYPFE, which is shared with eukaryotic IRS proteins but not bacterial IRS-S/MR proteins (Figure 2). Amino-acid compositions are significantly different between the IRS-HR and IRS-S/MR proteins of *S. aureus* (data not shown), further suggesting that the two genes originated from different organisms. A similar gene transfer has been suggested for glutaminyl-tRNA synthetases, for which certain proteobacteria have a eukaryote-like gene [14,15].

The deep phylogenetic position of bacterial IRS-HR proteins relative to eukaryotic IRS proteins suggests that a recent transfer of the
shortly after the divergence of the eukaryotic IRS gene was first transferred to an unknown bacterium. Observations, we propose that a population. To reconcile these genes, however, indicates that the existence of mupirocin-sensitive S. aureus might have been transferred through several bacteria species and only recently was it passed to S. aureus. Since mupirocin (as pseudomonic acid) is a natural product of some widely found strains of P. fluorescens, resistance to this compound might have conferred some competitive advantages to conspecific organisms, thus driving the fixation of IRS-HR genes in different bacterial species, though the possibility of random fixation cannot be excluded. To our knowledge, Treponema pallidum, Borrelia burgdorferi, Clostridium acetobutylicum and Chlamydia trachomatis have not been tested for susceptibility to mupirocin, probably because these organisms are not involved in epidemic infections and several are intracellular, obligate human parasites. The present analysis, however, strongly suggests these organisms to be resistant to mupirocin, as their IRS proteins clearly cluster with IRS-HR sequences with eukaryotes with one minimal length tree (4295 steps) found after 100 random heuristic searches. Numbers represent the percent occurrence of branching points in 1000 bootstrap replications (NJ) or 1000 puzzling quartets (ML). Values less than 50% are not shown, and in the ML tree those nodes are collapsed. The scale bar represents an estimated 0.1 amino-acid substitutions per site. Mitochondrial targeted isoforms of eukaryotes are indicated by mt.

resistance gene from humans or animals as a consequence of intensive clinical use of mupirocin is an unlikely scenario. Plasmids mediating high-level resistance were detected in staphylococci isolated nearly twenty years before the clinical use of mupirocin [16]. The existence of mupirocin-sensitive S. aureus and plasmid-borne IRS-HR genes, however, indicates that IRS-HR is not fixed in the population. To reconcile these observations, we propose that a eukaryotic IRS gene was first transferred to an unknown bacterium shortly after the divergence of the Eucarya and Archaea. The gene might have been transferred through several bacteria species and only recently was it passed to S. aureus.

Phylogenies of isoleucyl-tRNA synthetases (IRSs) using (a) neighbor-joining (NJ) and (b) maximum likelihood (ML) methods. Those bacterial IRSs suggested to be highly mupirocin resistant are indicated (HR), and those species in which it has been confirmed by experiments are in red. Trees are based on the non-gapped multiple sequence alignment of amino acids flanked by class 1 aminoacyl-tRNA synthetase signature motifs HIGH and KMSKS (438 amino acids). The NJ tree was generated with the programs PROTDIST (using the Dayhoff PAM120 option) and NEIGHBOR of the PHYLIP package [19]. Maximum log likelihood of the ML tree (using the program PUZZLE 3.0.1 [18] with the jTT option) was -23,862.83 (± 14.99). Maximum parsimony (MP) analysis (using the program PAUP 3.1.1 [20]) also clustered IRS-HR sequences with eukaryotes with one minimal length tree (4295 steps) found after 100 replications (NJ) or 1000 puzzling quartets (ML). Values less than 50% are not shown, and in the ML tree those nodes are collapsed. The scale bar represents an estimated 0.1 amino-acid substitutions per site. Mitochondrial targeted isoforms of eukaryotes are indicated by mt.
research grants to M. Nei, A. Lupas, C. Perry and J. Poupard for their contributions.

We thank K. Coleman, R. Jarvest, E. Lawlor, and others for their assistance.

An alignment of multiple IRS amino-acid sequences is published with this article on the internet.

**Supplementary material**

An alignment of multiple IRS amino-acid sequences is published with this article on the internet.

**Acknowledgements**

We thank K. Coleman, R. Jarvest, E. Lawlor, A. Lupas, C. Perry and J. Poupar for their contributions. This work was supported by NIH and NSF research grants to M. Nei.

**References**


Supplementary material

A bacterial antibiotic resistance gene with eukaryotic origins
[Correspondence]
James R. Brown, Jianzhi Zhang and John E. Hodgson
Current Biology 21 May 1998, 8:365–367

Alignment of multiple isoleucyl-tRNA synthetase (IRS) amino-acid sequences

Columns marked with * were used in the phylogenetic analyses. Abbreviations correspond to IRS sequences from the following species with their GenBank accession numbers or source: ITma (Thermotoga maritima, L37104), IAp (Aquifex pyrophilus, L37096), ICje (Campylobacter jejuni, U15295), IHpy (Helicobacter pylori, AE000642), ISau (Staphylococcus aureus, X74219), ISaR (Staphylococcus aureus Mup R gene, X75439), ITpa (Treponema pallidum, The Institute for Genomic Research, personal communication), IBbu (Borrelia burgdorferi, AE001181), ICtr (Chlamydia trachomatis, Chlamydia Genome Project, personal communication), IMge (Mycoplasma genitalium, U39718), IMpn (Mycoplasma pneumoniae, 2501010), IMtu (Mycobacterium tuberculosis, Z74020), IBsu (Bacillus subtilis, Z99112), ICac (Clostridium acetobutylicum, Genome Therapeutics personal communication), ISpn (Streptococcus pneumoniae, The Institute for Genomic Research, personal communication), ISyn (Synechocystis sp. PCC6803, D90907), IPfl (Pseudomonas fluorescens, X80132), ISmt (Saccharomyces cerevisiae mitochondria (mt.), L38957), ICmt (Caenorhabditis elegans mt., Z81038), IHmt (Homo sapiens mt., assembled from NCBI dbEST), IHi (Haemophilus influenzae, U32777), IEco (Escherichia coli, D10483), INgo (Neisseria gonorrhoeae, University of Oklahoma's Advanced Center for Genome Technology, personal communication), ISce (Saccharomyces cerevisiae, M19992), ITth (Tetrahymena thermophila, M30942), IHa (Homo sapiens, U04953), ICell (Caenorhabditis elegans, Z70310), INlo (Nosema locustae, L37097), IMth (Methanobacterium thermoautotrophicum, M59245), IMja (Methanococcus jannaschii, U67538), IPfu (Pyrococcus furiosus, L37105), IAfu (Archaeoglobus fulgidus, AE001061), ISac (Sulfolobus acidocaldarius, L37106).
<table>
<thead>
<tr>
<th>ID</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITma</td>
<td>DTHGLPIERHSVQELGEKI----KEMSP-AEIRKCEEFALRFVDIQRREEFKRLGVGDM</td>
</tr>
<tr>
<td>IAp</td>
<td>DTHGLPIERAVEKSSKKKIR--KESLPKTEFRKLCREYANRYVNIQKEFIRLGVGLGD</td>
</tr>
<tr>
<td>ICje</td>
<td>DCHGLPVEQQVEVKGLEK-----KKSLSKKEIREFCRQHASEFDVQEREKFNGLGIAWD</td>
</tr>
<tr>
<td>IHy</td>
<td>DCHGLPVEQQVEVLEKKEKEK-----TSLENPLFLRKEHKFLIEIQNEFILQVGLVGLDF</td>
</tr>
<tr>
<td>ISau</td>
<td>DCHGLPVEQALTEQEPEGK----KEMSP-AEIRKCEEFALRFVDIQRREEFKRLGVGDM</td>
</tr>
<tr>
<td>ISaR</td>
<td>DCHGLPVEELKEEGKKIGKQ--DIEKYGIFINEKEKSVALQYKEBERRFSKDLGVGVM</td>
</tr>
<tr>
<td>ITpa</td>
<td>DCHGLPVEVEKLGISGKY-EIENYGIFNKEKCRKVLRYEATWKNKIIRLGVRVMD</td>
</tr>
<tr>
<td>IBlu</td>
<td>DCHGLPVEVEKLGISGKY-IEIENYGIFNKEKCRKVLRYEATWKNKIIRLGVRVMD</td>
</tr>
<tr>
<td>IMge</td>
<td>DCHGLPVEHASKNNNS-----YNSNTSVKKEKHCQCSVQAIQVKEQFQRLGLNNDF</td>
</tr>
<tr>
<td>IMpn</td>
<td>DCHGLPVEHASKKDFQH-----YASLSEKRDLCQFALCSAQAIQVKEQFQRLGLNNDF</td>
</tr>
<tr>
<td>IMtu</td>
<td>DCHGLPVELEKVLGSKSSQT-----YSNLSTVEKRLCHQCEYFQALCSAQAIQVKEQFQRLGLNNDF</td>
</tr>
<tr>
<td>IBbu</td>
<td>DCHGLPVEYEVEKKLGISGKY-EIENYGIFNKEKCRKVLRYEATWKNKIIRLGVRVMD</td>
</tr>
<tr>
<td>ICtr</td>
<td>DCHGLPVEYEVEKKLGISGKY-EIENYGIFNKEKCRKVLRYEATWKNKIIRLGVRVMD</td>
</tr>
<tr>
<td>IICac</td>
<td>DCHGLPVEYEVEKKLGISGKY-EIENYGIFNKEKCRKVLRYEATWKNKIIRLGVRVMD</td>
</tr>
<tr>
<td>ISyn</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IPSl</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IPfl</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>ISmt</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>ICmt</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IHmt</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IHin</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IEco</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>INgo</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>ITth</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IHsa</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>ICel</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>INlo</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IMth</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IMja</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IPfu</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>IAfu</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
<tr>
<td>ISac</td>
<td>DCHGLPVEYIQLKQGVKR------KEMDLVEYKLCREYALSQVDQREDFKRLLGVGLGD</td>
</tr>
</tbody>
</table>

**Supplementary material**
**********                   *****************************
#ITma SIYVKFRSKEDPNFF-------------IVIWTTPTWPILSANVIALHPDYEYSVVKVG-
#IApY SIYVKFPLKEIEGKKAY------------AVIWTTPTWPILSANGLMVEADYSLVEVE-
#ICje SIFVAGDKCLKGLSAS-------------AVIWTTPTWPILSANQIAILNFENYVIKEG-
#IHpY SIFVAGFLKESLKDKEVKSAS--------LVIIWTTPWTLSANVAIAUKEAYSTQK-
#ISau SIYVFNVDKQGAVDADAK---------FIIVIWTTPWTSVAINTHPHELKQVYKQV-
#ISaR SIYVFQNLQKDY---------FILSWWTTPWTPILSANLAINKDLNYSKIRVE-
#ITpa AITVRFECTSS??TC----------FILAIWTTPWTPCNAA?ALGPQILYVLIEAN-
#IBbu SLTIKFKIKDKNEE-------LWLVTTPWTPSNGIALVGIEQIEYSKIFDK-
#ICtr SIVVFALQDNQGF-------------LILAATWTPTLSMALAVHPDLYTIVKIK-
#IMge APLYTFKVSKSDFLDNAN-------LIVWTTTPWTLPSNQAIARPFDFYLLFEYN-
#IMpn GLYLFKTIVQSAVLNSGTK-------LIVWTTTPWTLPSNQAIARPFDFYLLFEYN-
#IMtu SRQDFAVGVKQVGGQFDNGL-------DGAYLLVWTTTPWTPLSNALAVUPDITVQVQAG-
#IBsa IYVAFGKDQVSGVLNQLEN-------RIIWTTPWTPANGLISVHPDLYTSVIAVG-
#ICac ATAFVKFVKGEEK-------YLAWTTTPWTLPNQLAVAPIHPELTYAVVQAG-
#ISPn LYYANKVQKQGKVQDLT--------DTYIVIWTTPFTITASRGLTVADYVLQVPA-
#ISyn SIYVSFPITQAGEKAKAEILAPY--DIIAIWTTPWTLPSNALAVALNPETYAVVETE-
#IPfl TIDVAIADDKLAEEGASPASLA--KPAAIWTTPWTPANQALNVHPFEFTYLVAVG-
#ISmt AAYVFQLEKKSSLDCKLGLITNNPIYCLIWTSTPWTLSNARICFQNQDFYSYSSLLRN-
#ICnt AYPFRLKINFSDDVSVWASPEKISIQFFALIWTTPWTPLNASVSAIQYSQFDN-
#IHmt IYVFKPFLKPSQIGISYRWSVPA-----SILWVTQWPWTIPANEACVCMPEKSYAVVKCSK-
#IHmN IYVFRAFQASEDEIEAKFSAQGRGGKLAIIWTTPWTPMSHBRANIAVNADELYNLVQLG-
#IEco SIDVAFQVQDQALAKAFVAVSNNR-GPSIIVWTTTPWTLPSNALAINKQDQVQVQAG-
#INgo AIDVAYFQKTVALAAAFGLACIE-GKAFAWITTPWTPASQAVASAGADVLYQIDTF-
#ISCe AVTFBEIQEKEQKQ-------LWATTTPWTLPSNASLAVNDEFEQYKIYDE-
#ITth SLFAIAKTAEDFKTK------FIWTTTPWTPLSNALVIKDFDYKVSLDA-
#IHsa SVFVTFFLEETVSV--------LWATTTPWTPSNAASVQKMFYVXKIDV-
#ICel AVFVMKLDCCPR-QQ--------LWATTTPWTPLVCNVVPMQYVKKIDV-
#INlo SILVAFLKFRKQLYSLS--------LWATTTPWTPSHMALVQFQSYIAIFR---
#IMth IYVFKPVSGDGY---------ILVWTTTPWTPANMAVAVHPFDYAYVRLD- 
#IMja SIYVFRAFQASEDEIEAKFSAQGRGGKLAIIWTTPWTPMSHBRANIAVNADELYNLVQLG-
#IPfu SIYVFQLEKKSSLDCKLGLITNNPIYCLIWTSTPWTLSNARICFQNQDFYSYSSLLRN-
#IAfu SIYVFQLEKKSSLDCKLGLITNNPIYCLIWTSTPWTLSNARICFQNQDFYSYSSLLRN-
#ISac SIYVFQLEKKSSLDCKLGLITNNPIYCLIWTSTPWTLSNARICFQNQDFYSYSSLLRN-

Supplementary material
Supplementary material

***** **** **************************** **** *******

#ITma ------RVILAD-YVSLETGTGCVHIAPGHGEEDYIYGHVQYGL------PIVSPVDEEFGRF----------
#IApv ETLKNNRWSFVSEFVSDLGTVHMAIAPGHQIDYTVGK-RYNL------EPYAPLDDSGRF----------
#ICje -------LIMGE-HLMDGSSGLVHTAPGHGEDDYYACL-KYGI-------EVLMPVDDSCCDEI3KL-----
#IHy ------VALGE-HVGLDEGTGAVHTAPGHGEEDYLGK-RYNL-------EVLMSVDKEGIDIEGIHNLW--
#ISau -------VINGD-HVTTDAGTGCVHTAPGHGEDDYIVGQ-KYEL-----PVISPIDKGFV----------
#ISaR ---NAYYVVDGE-FVSTSEGTGIVHTAPAHGEDDYLGQ-RYKL-----PVINVTGREGTY----------
#ITpa SEEGLCFTVDAFVSTEDGTGVVHVAPAFGEDDYEVLQDIDAGI-----SIQCPILDAEGRF----------
#IBbu ------KVHTAD-YVTTDDGTGIVHTAPAFGEDDYRILKK-HTNV-----DIIDPLDAEGRF----------
#ICtr KELEAFRILPID-FIEESEGTGIVHMAPAFGEDFFACP-EHNV------PLVCPVDEQGRCY--------
#IMge -------VLMGI-HVSDKNGETGIVHSSPAFGIDDFYLCQ-KNKL-------KEVLSIDEKGFV----------
#IMpn -------VLLGN-HVLNSBGTGLVHTAPLYGLDFYLCQ-NNKL------NEALVSLDEKGFV----------
#IMtu -------QVLAGD-FVTLDDGTGIVHMAGAYGEDMVMAE-AVGI------APATFVDPGKRF----------
#IBbu ------KVHTAD-YVTTDDGTGIVHTAPAFGEDDYRILKK-HTNV-----DIIDPLDAEGRF----------
#ICme -------VLMGI-HVSDKNGETGIVHSSPAFGIDDFYLCQ-KNKL-------KEVLSIDEKGFV----------
#ISPn -------VLLGN-HVLNSBGTGLVHTAPLYGLDFYLCQ-NNKL------NEALVSLDEKGFV----------
#ISyn -------IVIGGD-YVTTESGTGIVHMAPAGGEDDYIVGQ-KYEL-----PVISPIDKGFV----------
#IPfl -------VYLAD-YVELGSGTGIVHCSAYGVVDVDFVICK-KYGL------VNDDIINPVQSN----------
#ISmt -------VLMGI-HVSDKNGETGIVHSSPAFGIDDFYLCQ-KNKL-------KEVLSIDEKGFV----------
#ICme -------VLMGI-HVSDKNGETGIVHSSPAFGIDDFYLCQ-KNKL-------KEVLSIDEKGFV----------
#IHnt --------PLLPPNVMAKGTGVTGGLVHTAPAHGMEDYGVAS-QHNL------PMDCLVDEDFV----------
#IHIn -------VILGD-HVTLDGGTGLVHTAPDHGLDFIVGVQ-YQDL------FMAAGLVSNDNKGF----------
#IEco -------AILGD-HVTLDGGTGLVHTAPDHGLDFIVGVQ-YQDL------FMAAGLVSNDNKGF----------
#IIGo -------MLNGE-HVTKDAGGTGIVHTAPAGLEDYAVCN-KYGI------EHNVFNAEGKY----------
#ISce -------ETARFVISDD-YVTSDEGTGVHNAPAGGEGEDNQAAC-KNGVISESNVTNPPNADGGRF---------
#ITHh -------DCFLRILAGD-FVTAADGTGIVHCPGFGDEYKSV-AANNKPDPPV-PVDENHGF---------
#IHas ------NGAFVTLVDA-VYKEEEETGVHQPAGPYFEGEDRVCN-MFINRDKLSPCVPDAGSFCSG------
#ICel -------ERANFRVLLNTFVTSDEGTGVHQPAGPYFGEIDFQVCV-AVPNVIADQCMICRPSDGK--------
#INLo ------FGFRVVHYAD-FVSEIDGTVHCPGGFEDDYAMT-MPLIKENLPSDLENGRF----------
#IIMh -------DMEHRVILGD-HVTLTEGTGCVHTAPAGHIPDEFEIGK-EYGL------PVFPCPDAGAVF----------
#IMja ------LENAHTVILGE-FVTLLEGTGCVHTAPAGHIPDEFEVKG-KYNL------PIYSPIDEGKY----------
#IPfz -------KEWHRVLGD-HVTLLEGTGCVHTAPAGHIPDEFEQG-KYLG------PVYSPIDQEGKY----------
#IAfu ------NWKHQVFFA-DFVTAENTGCVHIAPGHGEDYELGY-EKGL-------EVFNPVDAGGKY----------
#ISac KINHHKVDGGEAVTHQEGTGLVHSAPGHzDFEIGK-KYDM------PVVMLVNDKGF----------
#ITma --TEEAGKY--KGMFIE--DANEVIEDLKRK---------GILVHASSITHSYPHCWRCKGPVIFRATEQWFIS
#IApy --VEPAFEPFIRGVRVF--DANKLIIALLKEK---------GLYVHEARIHRSYPHCWRCKNKFVIFRATFQWFFIG
#ICje --FSHLLEEFI--GLHIF--KANEKILELLGEK---------LLHSKFIYHPSFCWRTHKPVYIRATKQWFIL
#IHp --DES--QGEHVF--AQKRRIQVLQGDS---------GILVHASSITHSYPHCWRCKGPVIFRATEQWFIS
#ISau --TEEGGFQ--EGMFYD--KANKAVDILTEK---------GALLLDFITHSYPHDWTKKPVIFRATFQWFIS
#ISaR --NDRFPEL--VGNKAK--NSDIEIIKLLSKK---------GILVHASSITHSYPHCWRCKGPVIFRATEQWFIS
#ITpa --TAEVADY--QGLFVK--KADPKILEYMKEN---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IBbu --TNQVKDF-VGEYIK-SADKGIARRLKNE---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ICtr --TAEVKDF-VGEYIK-SADKGIARRLKNE---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IMge --NNLNDKELCENFYKLKANDLIINRLQKN---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IMpn --NDTLNDFVLGTFLFYKLKANDVIEELQKH---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IMtu --DVTVDY--QQHVF--DANAQIVRDLLKQGDS---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IBbu --TNQVKDF-VGEYIK-SADKGIARRLKNE---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IIsaw --TSEAPGF--EGMFYD--KANKAVDILTEK---------GALLLDFITHSYPHDWTKKPVIFRATFQWFIS
#ICac --VDEVEWP--KGLFVK--KADPKILEYMKEN---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ISPn --MKNAGPEL--VGNKAK--NSDIEIIKLLSKK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ISyn --TEEAGKF--AGLNVLNDANEIAINALLQDQ---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IPfl --GVYVPSLEFFGGQFIFKADQPIIEKLREV---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ISmt ELPQSVRSIVRDEGDLKQGRVWLDAETAKIPCKLSLDNLLYKSHTEITHSYPHDWRSKKPVIRATFQWFIS
#ICmt -----LGHD--LDGKEV-LGEGKIALRLNN---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IHmt --TDVAGPE-LQNKAV--LEEGTDLVVKMLQ---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IHln --ISTTEFF--AGKGVF-EANPLVIEKLQEV---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IEco --LPGTYFTL--DGYNVF--KANDIVVALLQEK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#INgo --ISETFPR--AGMSVW-EANPVILEQWPEET---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ISce --TKDVPDF-EGVYVK-DADKLIIKYLTNT---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ITth --TNVSDF--AGYIKF--EADKLIRNKLENER---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IHsa --TTEVTDF--AGQYVK--DADKSIIRTLKEQ---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ICel --TSEVPDY-QQYVK--DADLIIKRLKEM---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#INlo --NTEEVEY--KGMHVQ--DADSSIIKDLGKK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IMth --TEDAGKY--RGLFVK--DADSKIDDLLRSK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IMja --VE--GKW--KGVFYV--DADAEIIIETLKNK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IPfu --VE--GKW--KGVFYV--EADPQIIIEHLKEK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#IAfu --TEEAGKY--AGKHKV-EANVDIIIDLKRYK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
#ISac --TQDSGY--AGKYVR--SASEEIIISLKRK---------GILVHASSITHSYPHCWRCKGPVIFRATSEQWFIS
| #ITma     | ------VDHNNLRKQEEIDKVKWPEW------GRNIRSMVMSRDWCISQRVWGTPAPIAVKC |
| #IApy     | MDIEYEGKTLGSES-LEEIEKVWIPEY-----GRNIRKSMVENRDPDCISQRVWGVPITVFYC |
| #ICje     | MDIFKQGKTLRREC-AKEQLLLTFYQP-----SGVKRIGSMVENRDPDCISQRDWTIPAFFRD |
| #IHp y    | MDFQIQNGQKQK-LREVALDAIEKVFVPSGKNRLKTMNMRDWCISQRVWGVPILAFFID |
| #ISau     | I------SKVRQDI-LDAIENTNFKVNW-----GKTRIYNMRDGEVISRQVWGVPILVFYA |
| #ISar     | I------TTHEKNEI-INNNNIEFWSH-----IKEGRTMGLENMVWDNWGNFNYGPTLNVWIC |
| #ITpa     | VEIKIDKLMBAACNASCWQFSH-----IRDGRFGKWLVCARDWASRDGYGNFLILPWIC |
| #IBbu     | VEIKIKTKL-LEVNEKINWPAMAH----LKKRGFKWLENAKDAISRNNFVWPNIPWIC |
| #IIje     | VEIKIKTKL-LEVNEKINWPAMAH----LKKRGFKWLENAKDAISRNNFVWPNIPWIC |

**Supplementary material**
***************
#ITma FVCPCGKRSEFKMLDTLWVIDGSGASFYIY--------TKREDHPFPLDMYLEGDQHRG
#Iapy YKCPKCGGTSFRKEEDILDVWFDSGCSHSAV--------IRPLGFEDALYLEGSDQHRG
#ICje --SKYNAENLEKVYDILDVWFDSGTFNAVL-------NSGLYDAGEKRAMYLEGSDQHRG
#Ihyv --YQDEAKHEKIMHILDVWFDSGTFAKAVL-------EDYHKEQKPSDVILEGDQHRG
#ISau FTHGSNGTFTKEDIMDWFDGSGRHGVKL-------ETRPELSFADMYLEGDQHRG
#ISar TRLCFNCNKDSREVEVLDVWFDSGMPFAQQHYPDF--NQKIFNQHFADFIAEVDQTRG
#ITpa VITPCQGSGVRLDVDFESGAMPYAGQHYPFE-HATDFERYFAPAFISGOLDQTRG
#IBbu ITWPSKDGFKIFRTSEVLDFCWFESGAMPYASNHYPFE-NEINFKINIFADFAIEGLDQTRG
#ICtr -IEINQGKFSIRYPVDSCWFESGAMPYANQHYPFE-RAEETACPFADIFIAEGLDQTRG
#IMge KPGDKKWKYKHEIDTLDVWFDFSGSSNYVL--------EINKYGIADLYIEGDQYRG
#IMpn NPKKIQPGVEYKKECTDLEVDFDSGSTYNVL--------ISNKLNFADLYLEGSDQYRG
#IMtu NPDDFTGRSTMRRIPFVLDVWFDSGSMPYAVQHYPFE-LNDWFQHYPDFIVEYIQTG
#IBau FTHGSPNGTFTKEQDFIMDVDWDFSGSHQAVL-------EERRDLVRFADLYLEGSDQYRG
#ICac KLTCPKCGKPMTRTEEVLDVWFDSGSMPFAQLHYPFE-NKEYFENTFAQFISEAVDQTRG
#ISpn FTHGSPNGEKFKEIDTMVDWDFSGSSNGVY--------VNRPELTYPADLYLEGSDQYRG
#ISyn --YRNNGRTYRKEDTDMVDWDFSGSSWAAYA--------NAKNNRPKVFDMYLEGDQHRG
#IPlf -----GDEAPLDKVEYSLTLDVWFDSGTTHWHVL----RSNHPGMDPRLDLYLEGSDQHRG
#ISmt PEKHYHDVEYCRSQMDTMVDFDSGSWSVHKDFYKLKLSLKLPSPLYQVCLEGSDQHRG
#ICmt --------EKNTDMDWVDLSGLNWAAR----DNDEHEVADVLEGDQYRG
#IHmt SEVGPDALEYPQGIDILIDIFDSGTWSYV--------LPGDPQADLYLEGQKLG
#IHin -----GADAETYRKVPDFLDVFDSGTSYYSV--------ANHLEPQGIDMYLEGDQHRG
#IEco ------GDEAPLYDKISDTLDVWFDSGTTHWHVL----DVRPEQHAAADDMLEGSDQHRG
#INgo -----SADCEYKDKLDPDTMDVDFDSGSTHYSV-------KQREELWFDLADLYLEGSDQHRG
#ISce TIPSCQGSGDLKRIEVEFDFCWFESGMPYAQHYPFE-NTEKDFERVFANIFESLEQTRG
#ITth TIPSKGKGVLRIDEVFDFCWFESGMPYQQHYPFSMNSEEQSKRFADFIDEGIDQTRG
#ITHa TIPSCQGSGDKLRSEVFDFCWFESGMPYAVQHYPFE-NKREEFADFAIDFIAEIDQTRG
#ICel TIPSGQGSGDLKRIEVEFDFCWFESGMPYAQHYPFE-NKREEDNFADFAIDGIDQTRG
#INlo VICKDGEK--YRIEIEFDFCWFESGMPYQARHPWE-----CDNLCLDFAVADYDQTRG
#ITnh ILECRGCRGRMKRTPDVLWVIDGSGVAGWALHYPRE--KELFSEWFPYDITEHGDQTRG
#IMja KLRCE-CGEMGKPRILDFVDFDSGLAPYASIVG--------KELKADFITEGHDQYVTK
#IPfu IIKVN--GKEMRRVDFVDFVDFDSGIAWALGYPYR--NKLFEKLFADFFIVEGQVYTK
#IAfu TFSCQ-CGVMRVPDFVDVFVDFDSGVASWGSIAAYPL----RKKDFEELWFADITEGHDQTRG
#ISac VLDVKNVEQARRISDADVDFVDSGVAFFASLGQDW--RKR-WSELGFVBYTELEQDQYRG
#ITma WFHSSIFLAVKRGSAFYPKEVL--HGFIKDEQGR-----------------------------
#IApv WFQASLLESVGYSAGAYRSVLT--HGFIKDEQGR-----------------------------
#ICje WFQSSLSEGVTAINESTYESILT--HGFTTDEKQG-----------------------------
#IHpv WFQSSLGCVLNNQPKFVVIT--HGFIKDEKGE-----------------------------
#ISau WFNSSITTSATRGSYKFLLS--HGFTVMDGEGK-----------------------------
#ISqG WFYSLLVISTILGKSKYKRLAS--LGHILDSNGK-----------------------------
#ITpa WFYTITILAVALFEPFVNIV--TGLVIAKDSGK-----------------------------
#IBbu WFYTILTVIALTDIFAPKVNVIV--NGLVLSDSNGK-----------------------------
#ICge WFNSSNCIIQNDLIFKSLVS--HGFTDLKQGN-----------------------------
#IMpn WFNSSASCGIQTDQPFKSLIS--HGFTLDEHGN-----------------------------
#IMtu WFYTILHLALFDRPAKCTCA--HGIVLGFDQG-----------------------------
#IBau WFNSLSTAVAVKAPYKVGLS--HGFAIDGEGK-----------------------------
#ICac WFYllaaTSIFDKSFPNIV--LGHVLDKQG-----------------------------
#ISpn WFNSLSISVANHGPQPKQISS--QQFALDKQGE-----------------------------
#ISyn WFQSSLTSVAVNGAAYKTVLT--HGFLDEKQG-----------------------------
#IPfl WFSRSSITGCAIDNHAPRELLT--HGFTVDQGR-----------------------------
#ISnt WFQSSLITKVASNVPVAPYEYETVHGFTDQNGL-----------------------------
#ICnt WFSRSSITGAVKQPKLGKII--HGFCIDENNN-----------------------------
#IHnt WFSRSSITVAKRAKAPYKVTVIV--HGFTLQKQGE-----------------------------
#IHln WFSRSSITLSTADSKAPYKQVLT--HGFTVQKQQR-----------------------------
#ICeo WFSRSSITMSTAKPKAPYQQVLT--HGFTVQKQQR-----------------------------
#INQG WFSRSSITGSAKQPKQQLT--HGFTVQKQQR-----------------------------
#ISce WFSRSSITGQSVPSFKNIV--GSIQLAADDG-----------------------------
#ITth WFYTTNISTALFLSNQPVNLIV--NGVQACDDG-----------------------------
#IHsa WFSYTVVSSLFAPKQNPKQNVIV--NGVQACDDG-----------------------------
#ICel WFSYTVVSSLFAPKQNPKQNVIV--NGVQACDDG-----------------------------
#INo WFSYTMHISTVSTLAPKPFRNCIV--NGVQACDDG-----------------------------
#IMth WFSYSLGCGVIALTEPYRVM--HGFTLDEEGR-----------------------------
#IMja WFSYQHALSAIVNIDPKKCLM--HGFTLDEEGR-----------------------------
#IPfu WFSYQQAASPITYYPFYYAGVAM--HGYYLDEKGD-----------------------------
#IAfu WFSYQLGTSVCFKAPYKAVLM--HGFTLDEQGR-----------------------------
#ISac WFSLRTGVILMDKAPYEAVLIV--HGFMALDEQGREMKKSSSNGYVPSQVSKYQRYVKIMV
******

#ITma -------KMSKSL
#IApy -------KMSKSL
#ICje -------KMSKSK
#IHpy -------KMSKSK
#ISau -------KMSKSL
#ISaR -------KMSKSK
#ITpa -------KMSKAL
#IBbu -------KMSKSF
#ICtr -------KMSKRL
#IMge -------KMSKSL
#IMPn -------KMSKSL
#IMtu -------KMSKSL
#IBsu -------KMSKSI
#ICac -------KMSKHK
#ISpn -------KMSKSL
#ISyn -------KMSKSL
#IPfl -------KMSKSL
#ISmt -------KMSKLV
#ICmt -------KMSKSI
#IHmt -------KMSKSL
#IHin -------KMSKSI
#IEco -------KMSKSI
#INGo -------KMSKSI
#ISce -------KMSKSL
#ITth -------KMSKSK
#IHsa -------KMSKRK
#ICel -------KMSKSK
#INlo -------KMSKRL
#IMth -------KMSKSL
#IMja -------KMSKSL
#IPfu -------KMSKSL
#IAfu -------KMSKSL
#ISac TLEIPHGKMSKSH