

TEACHER'S GUIDE

DNA *to Darwin Case Study*

A question of taste

Version 1.2

Anne Fischer

Formerly of the Max Planck
Institute for Evolutionary
Anthropology, Leipzig

Dean Madden

NCBE, University of Reading



Charles Darwin

Introduction

There are two exercises in this *Case Study* in which students investigate the evolution of taste receptors in a range of animals. In the first activity, protein sequence data is used to generate a phylogenetic tree of sweet, umami and bitter taste receptors from six animal species. In the second, more advanced, activity, the evolution of bitter taste receptors by gene duplication is studied using DNA sequence data. The unusual case of the giant panda genome is also introduced.

A sense of taste

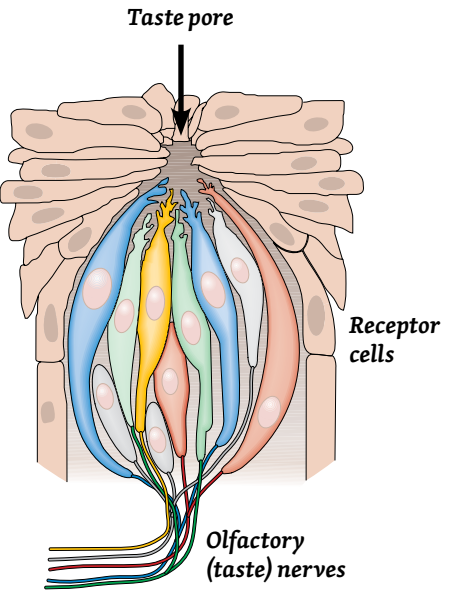
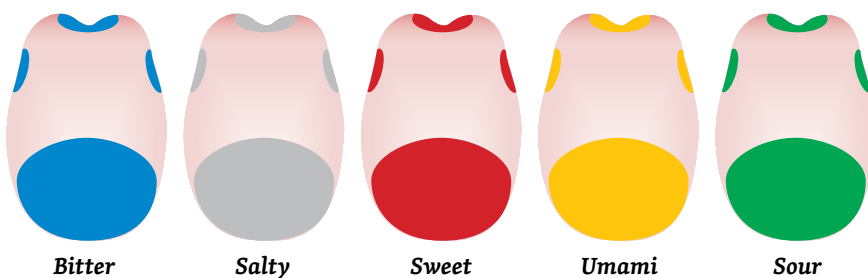
Tell me what you eat: I will tell you what you are

Jean Anthelme Brillat-Savarin, *The Physiology of Taste*, 1825

For many years it was thought that there were only four principal tastes that humans could detect: sweet, sour, salt and bitter. Since the 2000s it has been widely accepted that there is a fifth basic taste, umami, the flavour associated with savoury foods such as meat. Monosodium glutamate and soy sauce have the taste characteristic of umami. All other tastes are thought to arise from the stimulation of these five receptors, in a similar manner to that in which a small range different light receptors in the eye allow sensitivity to the entire visual spectrum.

The five tastes are thought to indicate salt, simple carbohydrate (sweet), acidic flavours associated with ripe fruit, and when highly acidic, decay (sour), toxins (bitter) and protein (umami). The adaptive advantage to animals in being able to detect such compounds is clear.

Contrary to the suggestion found in some textbooks, receptors for the five principal tastes are distributed over the entire tongue (and also the palate and throat) and are not localised to particular areas.



Structure of a single taste bud. After Chandrashekar, J. et al (2006).

Distribution of taste receptors on the human tongue. After Chandrashekar, J. et al (2006).

This is because each taste bud is thought to consist of a bundle of 50–150 different types of receptor cells that are sensitive to each of the main tastes.

As the physiological and molecular mechanisms of taste have become better understood, the existence of other receptors that respond to specific compounds (such as menthol, starches and fat) has been proposed, particularly in non-human animals.

Bitter taste receptors

Bitter taste receptors play an important role in avoiding the ingestion of toxic and poisonous substances (such as plant alkaloids). These receptors are thought to have the simplest structure. They consist of T2R proteins that are anchored in the cell membrane of the taste receptor cells. The proteins are encoded by *TAS2R* genes.

Mammals live in very different environments and consequently they are adapted to different diets and have evolved appropriate sensory repertoires. In the human genome, 25 bitter taste receptor genes have been identified; in the mouse genome there are 33 genes.

Gene duplication, new genes and pseudogenes

Each of the wide range of bitter taste receptors is encoded by a slightly different gene that has arisen through gene duplication followed by mutation. This is one of the main ways in which new genes arise. The duplication process is not easy to explain. It can occur through several mechanisms and it is still unclear exactly how all of these work.

Initially, the two copies of the gene are identical and share the same function. As mutations accumulate on the copy of the gene it can lead to the emergence of a new gene with a new function. Most of the time, however, mutations will disrupt the function of one of the two copies. This process leads to the emergence of a *pseudogene*, an inactive version of the gene. In humans, ten bitter taste receptor pseudogenes have been identified; in the mouse, only three. Pseudogenes are useful in evolutionary studies, as they form a 'fossil record' in the genome.

Whatever the fate of the new copy of the gene, it will evolve faster. This is the result of the relaxation of selective constraints acting on the gene which is reflected by a greater number of mutations accumulating in the duplicated gene.

In the second exercise in this *Case Study*, students compare two bitter taste receptors from seven primates (including humans) and try to determine which of them is the copy (that is, the one with the greatest number of mutations).

Sweet taste receptors

Most animals seek foods such as fruit that taste sweet to humans. These foods are a valuable source of carbohydrate and the adaptive value of being able to detect such foods, like sensitivity to bitter poisons, is obvious.

In humans, the sweet taste receptors can be 'fooled' as they also respond to artificial non-carbohydrate sweeteners such as aspartame and acesulfame-K, although this is not true of all animals.

The receptors that detect sweet tastes they are made of two parts. The protein T1R3 is linked to another, similar, protein called T1R2. T1R3 is encoded by a protein called *TAS1R3* and T1R2 is encoded by *TAS1R2*.

It has been found that variants (polymorphisms) of the gene *TAS1R3* lead to different sensitivities to sweetness. This could be the reason for individual dietary preferences.

Some animals, such as cats and chickens, cannot detect sweet foods. This is because, in cats, the gene encoding T1R2 has accumulated mutations and become a non-functional pseudogene. In chickens, the gene is missing.

Umami taste receptors

Amino acids and small peptides, indicative of protein in food, are also detected by two-part receptor. Here, however, the T1R3 protein is linked to a T1R1 protein (encoded by *TAS1R1*) instead of a T1R2. This is the basis of the 'umami' or savoury taste.

Giant pandas

In giant pandas, the gene encoding T1R1 has gained two mutations, a two base-pair insertion ('GG') and a four base-pair deletion ('GTGT'). These mutations make the gene non-functional and like the *TAS1R2* gene in cats it has become a pseudogene. Consequently, pandas, unlike other bears, cannot taste meaty flavours.

In the first exercise, students use protein sequence data to generate a phylogenetic tree of sweet, umami and bitter taste receptors from eight animal species. In so doing they can speculate about the order in which the different genes evolved and the tasting abilities of the various species.

General reading

The making of the fittest. DNA and the ultimate forensic record of evolution by Sean B. Carroll (2009) Quercus Books (Paperback) ISBN: 978 1847247247.
A popular lay account of some of the molecular evidence for evolution.

Reading the story in DNA: A beginner's guide to molecular evolution by Lindell Bromham (2008) Oxford University Press (Paperback) ISBN: 978 0199290918.

An engaging textbook on molecular evolution, which assumes no specialist mathematical knowledge and takes the reader from first principles. Although it's aimed at undergraduates, this superb book contains sufficient detail for PhD students, yet parts will appeal equally to 16–19 year-olds.

The paper on which the first activity in this Case Study is based (Shi and Zhang, 2006) is also presented as a Case Study in this book.

Scientific papers

All of these documents except for Chandrashekar *et al* can be accessed free-of-charge, online.

- Fischer, A. *et al* (2005) Evolution of bitter taste receptors in humans and apes. *Molecular Biology and Evolution*, 22 (3) 432–436. doi: 10.1093/molbev/msi027. *This is the paper on which the second activity is based.*
- Li, X. *et al* (2005) Pseudogenization of a sweet-receptor gene accounts for cats' indifference toward sugar. *PLoS Genetics*, 1, 27–35. doi:10.1371/journal.pgen.0010003. *Why cats can't taste sweet things.*
- Shi, P. and Zhang, J. (2006) Contrasting modes of evolution between vertebrate sweet/umami receptor genes. *Molecular Biology and Evolution*, 23 (2) 292–300. doi: 10.1093/molbev/msj028. *This is the paper on which the first activity is based.*
- Chandrashekar, J. *et al* (2006) The receptors and cells for mammalian taste. *Nature*, 444, 288–294. doi:10.1038/nature05401. *General review article (paid access).*
- Breelin, P. and Spector, S. (2008) Mammalian taste perception. *Current Biology*, 18 (4) R148–R155. doi:10.1016/j.cub.2007.12.017. *A primer on how mammals perceive taste.*
- Ruiquiang, L. *et al* (2010) The sequence and *de novo* assembly of the giant panda genome. *Nature*, 463, 311–317. doi:10.1038/nature08696. *Why giant pandas can't taste savoury things.*

Requirements

DNA sequence data

The software required, *Geneious*, can be downloaded free-of-charge from: www.geneious.com. The software is available for Windows, Macintosh and Linux operating systems. Only the free, 'basic' version of the software is required for this activity.

DNA sequence data

For the first activity, students will need the *Geneious* document containing 18 amino acid sequences for T1R1, T1R2 and T1R3 taste receptors: **T1Rproteins.geneious**

For the second exercise, students will need the *Geneious* document containing 14 ready-aligned DNA sequences from bitter taste receptors: **Bittertastereceptors.geneious**.

Students' worksheets

Students will require copies of worksheet pages 5–9 for the first activity.

The second exercise is described on worksheet pages 11–21.

Page 10 describes the giant panda's insensitivity to savoury tastes and can be used to promote discussion about evolution or as an extension activity or homework.

Presentations

Teachers may find the *QuickTime* animation and *PowerPoint* or *Keynote* presentations helpful for introducing this exercise. *QuickTime* may be downloaded free-of-charge from the Apple web site: www.apple.com/quicktime

Educational aims

The activity reinforces students' understanding of protein and DNA structure and mutation. It introduces the principle of alignment. Several evolutionary concepts are introduced.

Prerequisite knowledge

Before starting either exercise, students would benefit from carrying out one or more of the *Introductory activities* which are described in separate documents. These cover DNA structure, the principle of alignment, building evolutionary trees and using the *Geneious* software.

Students will need to understand the basic structure of DNA and proteins. They will need to be taught the principle of sequence alignment. Knowledge of cell membrane structure would be useful.

Answers to the questions on the worksheets

Page 4

- a. *Sweetness* usually indicates readily-available carbohydrate.
Sourness (acidity) shows, for example, that fruit are ripe or when very acidic, that the food is bad (although some sour, fermented foods are more nutritious or easier to digest).
Bitterness can indicate the presence of toxins, and thus that a particular substance should not be eaten.
Saltiness indicates sodium ions (or other ions) which are essential for survival.
Umami indicates the presence of protein.
- b. Refer to the table below:

TASTE	GENES INVOLVED	PROTEINS ENCODED
Umami	TAS1R1 + TAS1R3	T1R1 + T1R3
Sweet	TAS1R2 + TAS1R3	T1R2 + T1R3
Bitter	TAS2R	T2R

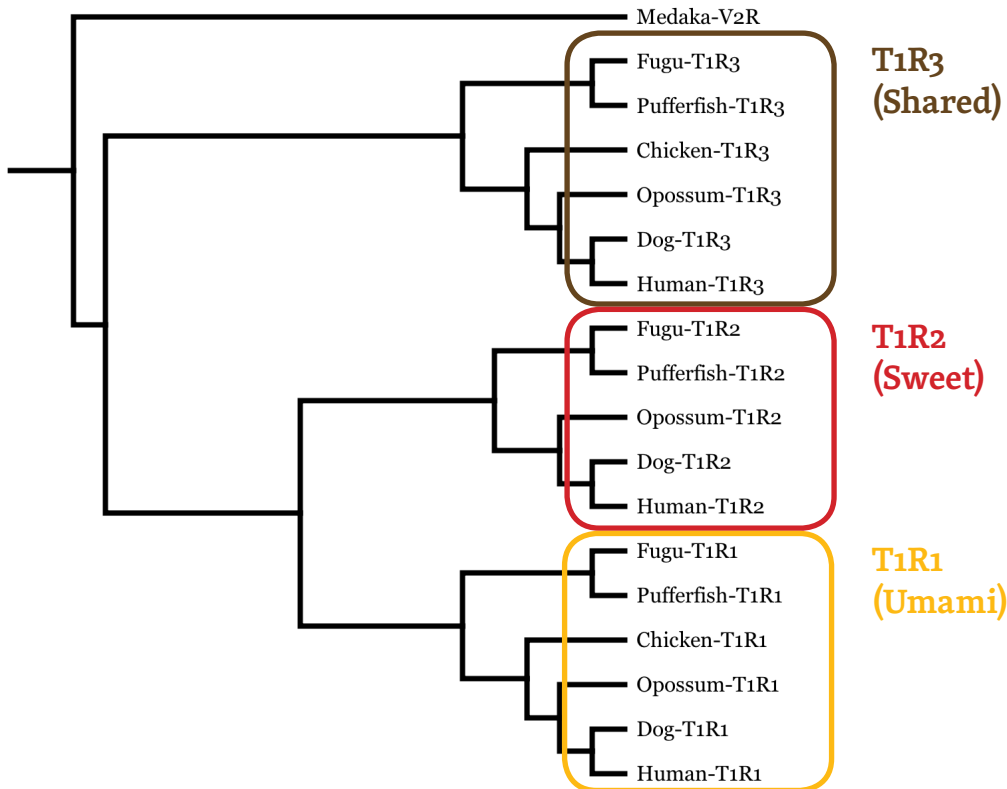
Page 9

- c. The three clusters contain the proteins T1R3, T1R1 and T1R2 (see below).
- d. The common ancestor would be on the extreme left of the tree.
- e. T1R1 and T1R2 are more closely-related to each other than they are to T1R3, as they share a more recent common ancestor with each other than they do with T1R3.
- f. T1R3 appears to have evolved first, followed by T1R2 and T1R1.
- g. The ability to taste sweet things probably evolved first, based on the branch lengths of the tree (students will need to ensure that 'Transform: proportional' is selected in the tree formatting panel).
- h. See the table below:

Organism	T1R1	T1R2	T1R3
Pufferfish	✓	✓	✓
Fugu	✓	✓	✓
Dog	✓	✓	✓
Human	✓	✓	✓
Opossum	✓	✓	✓
Chicken	✓	✗	✓

From the table it would appear that chickens might not be able to taste sugar, because they lack the necessary T1R2 protein.

- i. The data in the table could be misleading if sequences for some species and genes have not been provided or are not available.



Page 14

- Yes, the sequences all look very similar to one another.
- There are 945 bases in *T2R16* and 951 bases in *ps1*.
- Because there are four different bases in DNA, the probability of one base being shared is 1 in 4 (0.25). The probability of two consecutive bases being shared is 0.25×0.25 ... three consecutive bases $0.25 \times 0.25 \times 0.25$... etc. To convert the figure you have calculated to a 'one in x' form, simply divide 1 by it: hence $1 \div 0.25 = 4$; a 1 in 4 chance. The probability of ten consecutive nucleotide bases being identical in two sequences is 1 in 1,048,576. The probability of 20 consecutive bases being the same in two sequences is 1 in 1,099,511,627,776. The probability of all the DNA sequences of the seven primates being the same even for just ten bases is very small indeed. This strongly suggests that the seven primates are related by common descent. (This calculation also explains why the PCR can pick out and amplify specific sections of DNA from an entire genome using primers that are typically only 20 bases in length.)

Page 15

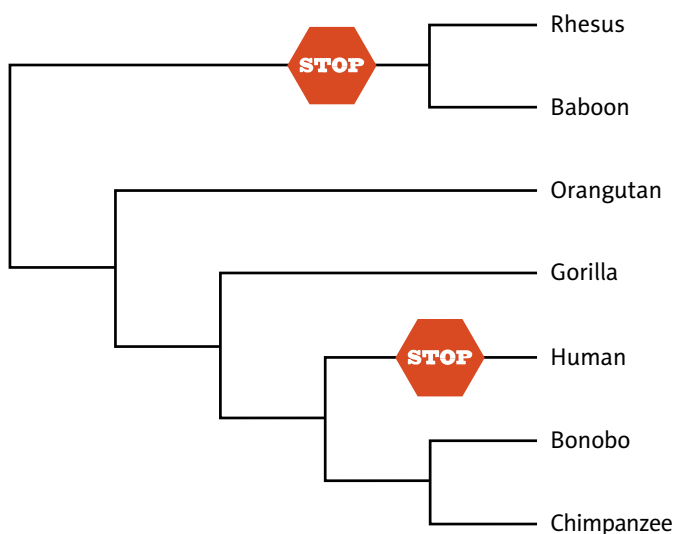
- The DNA sequences look different from each other, but there are sufficient similarities to suggest that they might be related.
- The tools that students will have encountered in *Geneious* are the tree-drawing function (which clusters together similar genes) and the ability to compare pairs (and indeed groups) of DNA sequences, that are shown in the Statistics panel.

Page 17

- The sequences corresponding to each gene cluster together.
- This suggests that the two receptors were present in the ancestor of all these primate species and that the duplication event occurred before the species diverged from each other.
- The DNA sequences of the two sets of genes are different and they therefore produce different phylogenetic trees.

Page 19

See positions of 'Stop' mutations on the tree below:



Three 'Stop' mutations should be visible in the protein sequence data. It is not possible to determine exactly where on the branches of the tree the mutations occurred but as one 'Stop' mutation is shared by the rhesus macaque and the baboon, but is not present in any other species, it must have occurred on the lineage to these two species.

The mutation in the human sequence is unique to the human, therefore it occurred after humans diverged from the chimpanzee and bonobo.

The 'Stop' mutations occur towards the end of the protein for the rhesus and baboon so it is unclear whether the proteins' function will be totally disrupted. The protein is incomplete but it might still be functional.

Pages 20 and 21

Number of sites (bases)

Gene pairs	Receptor T2R16		Receptor ps1	
	Shared	Different	Shared	Different
Human-Chimpanzee	941	4	935	16
Human-Bonobo	941	4	937	14
Human-Gorilla	934	11	938	13
Human-Orangutan	927	18	918	33
Human-Rhesus	900	45	900	51
Human-Baboon	900	45	897	54

Number of sites (bases)

Gene pairs	Receptor T2R16		Receptor ps1	
	% shared	% different	% shared	% different
Human-Chimpanzee	99.6	0.4	98.3	1.7
Human-Bonobo	99.6	0.4	98.5	1.5
Human-Gorilla	98.8	1.2	98.6	1.5
Human-Orangutan	98.1	1.9	96.5	3.5
Human-Rhesus	95.2	4.8	94.6	5.4
Human-Baboon	95.2	4.8	94.3	5.7

In these tables, only the mutations occurring between humans and other primate species are listed. There is a consistent pattern in that the number of mutations occurring on ps1 is always greater than the one on T2R16. This can be explained by a relaxation of constraints.

- j. *ps1* has accumulated the greatest number of mutations.
- k. *ps1* would therefore seem to be under less selection pressure.
- l. It all depends what you're looking for, but *T2R16* is more stable and would probably be a better indicator of evolutionary relationships for many purposes. As *ps1* is under less selection pressure (indeed, it may be a pseudogene lying 'dormant' in the genome), it may be less useful.