Background scientific knowledge

Table of contents

Background to the workshop

Introduction

The purpose of this document is to provide staff responsible for delivering *A Question of Taste* workshops with the necessary background information required to deliver the workshop successfully. It has been written for a general scientific audience who may have little previous experience of genetics or evolution. However, parts of the following sections do assume a little knowledge of DNA; readers who are unfamiliar with DNA might like to read:

1. *A brief introduction to genetics and molecular biology*, which starts on page 10
2. *A brief introduction to evolution*, which starts on page 12
3. *A brief introduction to taste*, which starts on page 13

A little history

Arthur Fox was a chemist who, in 1931, was using a powder called phenylthiocarbamide, or 'PTC'. Some of the powder blew into the air, prompting one of Fox's colleagues to complain about an unpleasant bitter taste. Strangely, Fox could not taste anything himself, so he decided to investigate further...

A Question of Taste

A Question of Taste is a workshop developed as part of a programme of educational projects celebrating the bicentenary of Darwin's birth, and 150 years since the publication of his most famous work, *On the Origin of Species by Means of Natural Selection*. Nowgen, At-Bristol and the Centre for Life originally developed the workshop in the UK, on behalf of the Wellcome Trust.

The *A Question of Taste* workshop uses molecular biology techniques to investigate the process by which humans and chimpanzees have both evolved the ability to *not* taste a bitter chemical. For decades scientists believed that some humans and chimpanzees share this trait as they had inherited it from common ancestors before the two species diverged. However, recent genetic evidence has shown that both humans and chimpanzees have undergone independent evolutionary processes that have resulted in this shared trait. This new evidence exemplifies Darwin's theory; selection pressures have acted on both species, and each has adapted accordingly. In this case, the adaptation (an inability to taste) has occurred via independent routes through a process known as convergent evolution.

Hands-on DNA: Exploring Evolution

Hands-on DNA: Exploring Evolution is a project led by The Association for Science and Discovery Centres (ASDC), in collaboration with the original project partners (Nowgen, At-Bristol and Centre for Life), again supported by the Wellcome Trust. The project's vision is to give students across the UK the opportunity to explore evolution through providing access to high-

quality engaging molecular biology experiences. *Hands-on DNA: Exploring Evolution* will make this possible by providing training and equipment to enable new organisations to deliver either the *A Question of Taste* PCR workshop, or for those organisations with less experience of molecular biology, a simpler two and a half hour DNA workshop.

Overview

The *A Question of Taste* workshop is based on the findings of Stephen Wooding *et al.*, which were published in Nature in 2006[1]. Before this paper was published it was known that both humans and chimpanzees are either 'tasters' or 'non-tasters' of a bitter tasting chemical, phenylthiocarbamide (PTC). More recently, the (in)ability of humans to taste PTC has been attributed to differences in the *TAS2R38* gene, which are now well understood[2]. The *TAS2R38* gene encodes a taste receptor, which is located on the tongue and is responsible for bitter taste perception in both humans and chimpanzees.

As humans and chimpanzees share this trait (being able or unable to taste PTC) and evolved from common ancestors, for decades scientists assumed that the trait had been inherited from a common ancestor before the divergence of the two species. As a logical conclusion, this went untested until 2006 when Wooding *et al.* set out to identify the differences in the DNA sequence of *TAS2R38* in chimpanzees that cause the (in)ability to taste PTC. The group found that the genetic differences resulting in an inability to taste PTC in the chimpanzee *TAS2R38* gene were different to those resulting in an inability to taste PTC in the human *TAS2R38* gene. Therefore, this research demonstrated that the previously held hypothesis was wrong; humans and chimpanzees did **not** share an inability to taste PTC when the two species diverged in evolution. Rather, the two species diverged from a common ancestor (that could taste PTC) and, at a later time, independently evolved the inability to taste PTC through unrelated changes to the *TAS2R38* taste receptor gene. This was one of the first papers to provide genetic evidence for a form of convergent evolution; independent evolutionary events which result in the same end effect.

This workshop sets out to allow students to investigate their own *TAS2R38* gene and see if their results correlate with their ability to taste PTC. This practical work can then be extended, through bioinformatics, to understand evolutionary processes by analysing the genetic similarities and differences between humans and our closest relative, the chimpanzee.

[1] Wooding S, Bufe B, Grassi C, Howard MT, Stone AC, Vazquez M, Dunn DM, Meyerhof W, Weiss RB, Bamshad MJ. Independent evolution of bitter-taste sensitivity in humans and chimpanzees. *Nature.* 2006; 440(7086):930-4.
[2] Kim UK, Jorgenson E, Coon H, Leppert M, Risch N, Drayna D. Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science.* 2003; 299(5610):1221-5.

What causes differences in bitter taste perception?

The *TAS2R38* gene encodes a taste receptor, which is a type of protein found on the surface of some tongue cells. Differences in this gene affect whether people can taste a chemical called phenylthiocarbamide (PTC), as well as other, similar chemicals, such as propylthiouracil (PROP). These different genetic make-ups are described as *genotypes*. The ability to taste or not taste PTC is described as a *phenotype*, an observable effect of the genotype. To people who can taste it, PTC is very bitter.

There are different forms of the human *TAS2R38* gene. A number of single nucleotide polymorphisms (SNPs; a point in DNA where a single nucleotide can vary) have been found in tasters and non-tasters, but the key SNP that is investigated in this workshop is at nucleotide position 145 in the human *TAS2R38* gene. SNPs can be analysed to genotype individuals using a variety of methods, including PCR, a technique that allows DNA to be copied millions of times. It is important to remember that each individual has two copies of every gene, one inherited from their mother and one inherited from their father. Therefore if we look at an individual's DNA, at any one point in a gene people can have two copies of the same letter (*homozygous*) or have two different letters (*heterozygous*). Table 1 shows the possible genotypes resulting from the single nucleotide polymorphism at nucleotide position 145 in the human *TAS2R38* gene:

Genotype		Individual's phenotype
Maternal DNA	**Paternal DNA**	
C	C	homozygous 'taster'
C	G	heterozygous 'taster'
G	C	heterozygous 'taster'
G	G	homozygous 'non-taster'

Table 1 Potential genotypes and their corresponding phenotypes for an individual

The *TAS2R38* gene accounts for up to 85% of variation in the ability to taste PTC, and is therefore a good predictor of an individual's (in)ability to taste it. Heterozygotes can taste PTC, but with less sensitivity than homozygous 'tasters'. However, ability to taste PTC can be affected by other factors, such as other genes involved in taste, age and smoking habits, amongst others.

Another important factor contributing to an individual's (in)ability to taste is the number of taste buds that are present on the tongue. Some taste buds are contained within fungiform papillae, the visible red dots found mainly on the tip of the tongue. Individuals who are very sensitive to a variety of tastes have a greater density of fungiform papillae than others. These individuals are often termed 'supertasters'. It is important to stress that the number and density of fungiform papillae can affect an individual's response to PTC. As the density of fungiform papillae

increases, the ability to taste bitter tastes increases for all 'non-taster' and 'taster' genotypes[3]. This effect would only be observable in the taste test; there is no genetic link between an individual's *TAS2R38* genotype and their fungiform papillae density.

What is convergent evolution?

Convergent evolution describes the acquisition of the same biological trait in species that have evolved through separate evolutionary pathways. In this example, the trait is the inability to taste the chemical PTC. This trait has arisen from small changes in the DNA sequence of the taste receptor gene, *TAS2R38*. Humans, chimpanzees and gorillas all have evolved from common ancestors. All gorillas tested can taste PTC; there is no 'non-taster' phenotype, which suggests that our common ancestor was also a taster. Accordingly, the DNA sequence of the gorilla's *TAS2R38* gene is used as a known starting point before the evolutionary changes occurred. Humans and chimpanzees have since evolved so that some individuals cannot taste PTC. As the changes to the gene are distinct between humans and chimpanzees, but they cause the same 'non-taster' phenotype, the evolutionary pathway is described as convergent.

What is the evidence for convergent evolution between humans and chimpanzees?

Observational work has shown that humans and chimpanzees can be either 'tasters' or 'non-tasters' of a bitter tasting chemical, PTC. Specific changes to the DNA sequence of *TAS2R38* result in an inability to taste PTC in both humans and chimpanzees. There are several SNPs that occur within the DNA sequence of *TAS2R38* that can cause differences in ability to taste PTC.

Table 2 shows that nearly all human 'non-tasters' have a guanine (G) at position 145 and nearly all human 'tasters' have a cytosine (C) at this position. The gorilla DNA sequence is identical to the most common 'taster' DNA sequence in humans and chimpanzees, suggesting that at some point in history, some humans and chimpanzees evolved from 'tasters' to 'non-tasters' and this trait was propagated through the generations.

Species	TASTER genotype	NON-TASTER genotype
Gorilla	C	N/A
Chimpanzee	C	C
Human	C	G

Table 2 Substitution of a guanine (G) nucleotide at position 145 results in non-tasting in humans, but is unchanged in non-tasting chimpanzees.

[3] Hayes JE, Bartoshuk LM, Kidd JR, Duffy VB. Supertasting and PROP bitterness depends on more than the TAS2R38 gene. Chem Senses. 2008;33(3):255-65.

The SNP at position 145 in human *TAS2R38* changes the DNA sequence such that it encodes a different amino acid. A cytosine (C) at nucleotide 145 leads to proline being encoded, whereas a guanine (G) at this position leads to an alanine. This amino acid substitution changes the taste receptor and reduces its affinity for the PTC chemical.

Chimpanzee non-tasters have a different SNP at position 2 (see Table 3). This SNP is in the start codon of the *TAS2R38* gene; therefore, chimpanzees with this mutation do not make a the TAS2R38 protein.

Species	TASTER genotype	NON-TASTER genotype
Gorilla	T	N/A
Chimpanzee	T	G
Human	T	T

Table 3 Substitution of a guanine (G) nucleotide at position 2 results in non-tasting in chimpanzees, but is unchanged in non-tasting humans.

From the information above, it is evident that the SNPs which result in chimpanzees and humans (in)ability to taste PTC are different. Therefore, humans and chimpanzees have adapted to become 'non-tasters' via independent evolutionary processes, but share the same phenotype. This is an example of convergent evolution and is illustrated in Figure 1.

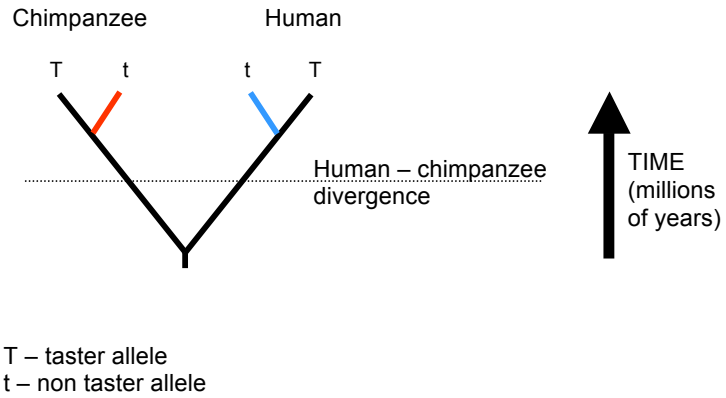

Figure 1 A diagram to show the mechanism of convergent (or independent) evolution for the PTC tasting phenotype (adapted from Wooding et al, 2006)

Why might bitter taste perception be evolutionarily important?

These findings have led to a discussion about why some humans and chimps have evolved not to taste PTC. It is unclear what selection pressure has led to, and maintained, this evolutionary change. Some scientists have proposed that differences in ability to taste may be advantageous in different geographical locations. *Brassica* vegetables such as broccoli and Brussels sprouts contain isothiocyanates which activate the TAS2R38 receptor. There is evidence that these vegetables have potent anti-cancer effects, and are therefore beneficial to eat. However, in

geographical regions of low iodine, over-ingestion of isothiocyanates has been associated with thyroid disease and goitre. Therefore, differences in the *TAS2R38* gene may be beneficial in preventing, or allowing the ingestion of isothiocyanates without experiencing an unpleasant taste [reviewed in 4].

To date there has been an assumption that changes to the *TAS2R38* gene that result in an inability to taste PTC mean that the resulting taste receptor is non-functional or broken. However, it may be that the receptor has changed to become responsive to a different chemical. Two studies have shown that fruits of the plant *Antidesma bunius* taste bitter to PTC 'non-tasters', but sweet to PTC 'tasters', which suggests that this plant contains a chemical which may act upon the TAS2R38 receptor. If the version of the *TAS2R38* gene that results in an *inability* to taste PTC, also results in an *ability* to taste a different chemical, this could indicate that both alleles are advantageous.

Both the examples above describe advantages associated with each *TAS2R38* allele, implying that it is advantageous to maintain both 'taster' and 'non-taster' SNPs within the *TAS2R38* gene of humans and chimpanzees. These advantages prevent one trait being 'selected out' (removed) over generations by means of natural selection and rather, the traits are maintained, or steady, within the populations. This process is called balancing selection.

How does this build on Darwin's work?

Darwin proposed that all species have evolved from common ancestors through a process termed natural selection. In his most famous work *On the Origin of Species by Means of Natural Selection* the only diagram is a tree of life, which shows the relationships between species that have evolved, in response to selection pressures, from common descent. This workshop builds on the principle that humans and chimpanzees have evolved from common ancestors, and investigates the process by which this evolution has occurred.

Darwin's work has been supported by fossil evidence and comparative studies in anatomy, physiology and biochemistry. Genetic information provides further supporting evidence for his theories. As DNA sequencing techniques and computing capability continue to advance, the scientific community has access to an increasing body of DNA sequence data with which to study evolution. Projects are underway to collect genetic information from all of the world's species which can then be compared using internet-based bioinformatics software. Experiments like these have shown humans and chimpanzees to be equally evolutionarily divergent from the gorilla, with a DNA sequence similarity of over 95%.

[4] Bufe B, Breslin PA, Kuhn C, Reed DR, Tharp CD, Slack JP, Kim UK, Drayna D, Meyerhof W. The molecular basis of individual differences in phenylthiocarbamide and propylthiouracil bitterness perception. Curr Biol. 2005;15(4):322-7.

A practical that bridges evolutionary research and molecular biological techniques

The work of Wooding and others provides an opportunity to extend the scope of PCR-based workshops that are traditionally offered to post-16 students in centres throughout the UK. The practical reflects current scientific research, and shows how scientific theories change over time.

The practical component of the workshop allows students to perform DNA extraction, PCR and gel electrophoresis and is therefore attractive to teachers in that it enables their students to cover these aspects of the curriculum. The practical will also focus on the analysis of results from a restriction digestion, which is included in most A-level biology curricula, and currently omitted from the majority of existing PCR-based workshops aimed at this audience. A further marked difference is that students investigate their own phenotype and genotype for PTC tasting using a taste test and analysis of their own DNA, which they find very exciting. The protocol has been considered by a leading UK geneticist who has stated that it is highly unlikely to reveal any health-related genetic information to the students, nor does it raise any other ethical issues.

Finally, and perhaps most importantly, the practical workshop demonstrates how a molecular technique is applied to a major scientific question. Students will learn the theory of the techniques, and apply them practically to investigate their own DNA, and to compare differences to their classmates, other participating students, and their evolutionary ancestors. This work brings together the history of evolution and cutting edge genetics research. Surely there is no better way to bring evolution to life than to look inside your cells and understand the evolutionary history of part of your own DNA?

How does the practical work?

Students use a range of techniques to investigate their genotype and phenotype for the ability to taste PTC.

Students investigate their phenotype by tasting strips of paper that have and have not been impregnated with PTC. They assess their own ability to taste PTC by comparing the taste of the PTC-impregnated strips with the taste of the control strips.

To investigate their genotype students extract their DNA and amplify the portion of the *TAS2R38* gene that contains the SNP at position 145 using polymerase chain reaction (PCR). The students perform a restriction digest on the PCR product, which cuts ‘taster DNA’ containing a cytosine (C) at nucleotide position 145, but not ‘non-taster DNA’ which contains a guanine (G). Finally, students separate the products of the restriction digest using gel electrophoresis and interpret the results in order to work out their *TAS2R38* genotype.

Generally, there is a high correlation between genotype and phenotype in this experiment, but it is important to stress that this correlation is not absolute. The study that first discovered that

TAS2R38 affects PTC sensitivity found that 50-85% of the (in)ability to taste PTC is due to polymorphisms in the gene. This is an important learning point within the practical that ensures students understand that a combination of several genes and environmental conditions are often determinants of a single trait.

A brief introduction to genetics and molecular biology

Genetics is the study of genes, variation and inheritance, while molecular biology is concerned with the molecular basis of biological processes. Both disciplines overlap considerably.

DNA

DNA is ‘the molecule of inheritance’ and its unique structure allows it to encode the information for other chemicals that help your body to grow, develop and function. Famously, the structure of DNA is that of a ‘double helix’ (Figure 2), which is akin to a twisted ladder.

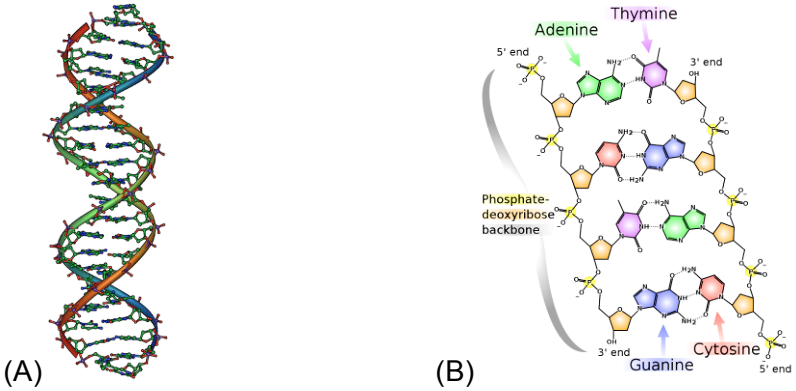

Figure 2 (A) DNA in a double helix. You can clearly see its backbone and the bases projecting inwards. (B) Simplified view of DNA, illustrating its constituent parts and showing the backbone, the four bases and how they link together. Both pictures are from Wikipedia.

DNA is made from two long chains of repeating subunits called nucleotides. A nucleotide has three parts:

1. a phosphate group
2. a sugar (deoxyribose)
3. a base

The sugar and phosphate groups link together to form DNA’s sugar-phosphate backbone, while the bases are connected to the sugar and point inwards, where they bind via hydrogen bonds with a complementary base on the second strand, forming a base pair. DNA contains four bases:

1. Adenine
2. Cytosine
3. Guanine
4. Thymine

For ease these bases are usually just referred to by their first letter; so, ‘A’, ‘C’, ‘G’ and ‘T’. Accordingly, the bases are often called the ‘letters’ in DNA’s ‘alphabet’. Importantly, a base in one strand of DNA always has a complementary base in the other strand. ‘A’ always complements ‘T’ (and vice versa), while ‘C’ always complements ‘G’ (and vice versa).

The chemistry of DNA is such that each strand can be described as having distinct ends, which are called the 5' (said, 'five prime') and 3' ('three prime') ends. The strands are 'anti-parallel', which means that one strand runs 5' to 3', while its complementary strand runs 3' to 5' (see Figure 2B). The bases that ultimately form the code of DNA can be read along a length of DNA; by convention, and to avoid confusion, DNA is always read 5' to 3'.

Genes and proteins

Genes can be described as discrete portions of DNA which encode another chemical, usually a protein. Proteins are the workhorses of a cell; they 'make things happen'. Like DNA, proteins are polymers, but rather than being made of nucleotides they are made of amino acids. The sequence of amino acids in a protein is determined by the sequence of bases in a gene. In the DNA, three bases form a codon which encode a single amino acid. Note that one codon only encodes one amino acid, but some amino acids are coded for by more than one codon; this is because there are only 20 amino acids, but 64 possible codons ($4^3 = 64$).

Because DNA determines the sequence of amino acids in a protein, it follows that a change in the DNA could result in a change to a protein. It is changes to proteins that usually cause an observable change to an organism; that is, a change in phenotype.

To make a protein from a gene, the gene is first copied to an intermediate molecule called messenger RNA (mRNA) in a process called transcription. Like DNA, mRNA is also made of nucleotides that contain bases, but in RNA the sugar is slightly different. Also, in RNA the base 'T' is replaced by another base, uracil ('U'). Because of the nature of DNA, the RNA copy is complementary to the template DNA strand. Following transcription, mRNA is transported from the nucleus (where the DNA resides) and translated into proteins by molecular machines called ribosomes. Ribosomes work by sequentially reading the codons in the mRNA and incorporating the appropriate amino acid for each codon. There are four special codons which act as punctuation, one is the start codon (ATG in DNA; AUG in RNA), the other three are stop codons. The start codon tells the ribosome where to start building the amino acid chain, a stop codon tells the ribosome where to stop.

If you want to find out more about genetics, you might like to look at:

- The DNA Learning Center's, *DNA from the beginning*
  http://www.dnaftb.org
- The University of Utah's, *Genetic Science Learning Centre*
  http://learn.genetics.utah.edu/

A brief introduction to evolution

Evolution describes the phenomenon of organisms' inherited characteristics changing over time. Such changes are typically small and happen gradually but, in combination with natural selection and speciation, evolution led to simple, single-celled organisms giving rise to millions of different, weird and wonderful complex forms of life.

Evolution relies on changes to DNA. Alterations to DNA lead to differences in an organism and these differences being passed on from one generation to another. Natural selection acts to preferentially maintain those organisms (or individuals) that have characteristics suited to their environment, while removing organisms/ individuals with less well-suited characteristics. Beneficial changes to an organism will therefore spread throughout its population, while harmful changes will not.

Speciation occurs when two populations of an organism are isolated in some way and are therefore able to develop independent changes, which do not spread between populations. Eventually, the two populations will develop characteristics so distinct (and, by implication, have such different genetics) that they can no longer interbreed and can be considered separate species. (Note, other, hotly debated mechanisms of speciation exist, but they are beyond the scope of discussion here.)

In order for evolution to occur it is important that DNA is not immutable. Were DNA unable to change (mutate) there would be no way that the phenotype could alter. DNA is able to change in a variety of different ways, including spontaneous, small mutations, called 'point mutations' where just a single base pair is changed. A mutation can be:

- **neutral** and have no observable effect;
- **positive** and give the individual with that mutation a reproductive advantage (and therefore more likely to pass on their genes and the mutation);
- **negative** and put the individual with that mutation at a disadvantage (and therefore less likely to pass on their genes and the mutation).

It is important to bear in mind that a mutation that is initially neutral might, at some point in the future, become either positive or negative.

Natural selection acts upon the phenotype of an organism, and increases or decreases the likelihood of an individual passing on their genetic information to the next generation, depending on the effect of the phenotype. Because the phenotype is the manifestation of the genotype, natural selection, by definition, affects the distribution of genes, and therefore alleles, in a population.

More information about evolution is available online from the University of Berkley's Understanding evolution: http://evolution.berkeley.edu/

A brief introduction to taste

Similar to how our eyes detect just three colours, plus shades of grey, so we only taste five distinct classes of taste. These are:

- Sweet
- Sour
- Bitter
- Salty
- Umami

Each of these different types of taste is detected by taste receptors; specialised proteins that sit on the surface of taste receptor cells (TRCs). In turn, a collection of 50-100 TRCs form a taste bud; roughly circular structures that sit on the surface of the tongue in further specialised structures called papillae. This is illustrated in Figure 3.

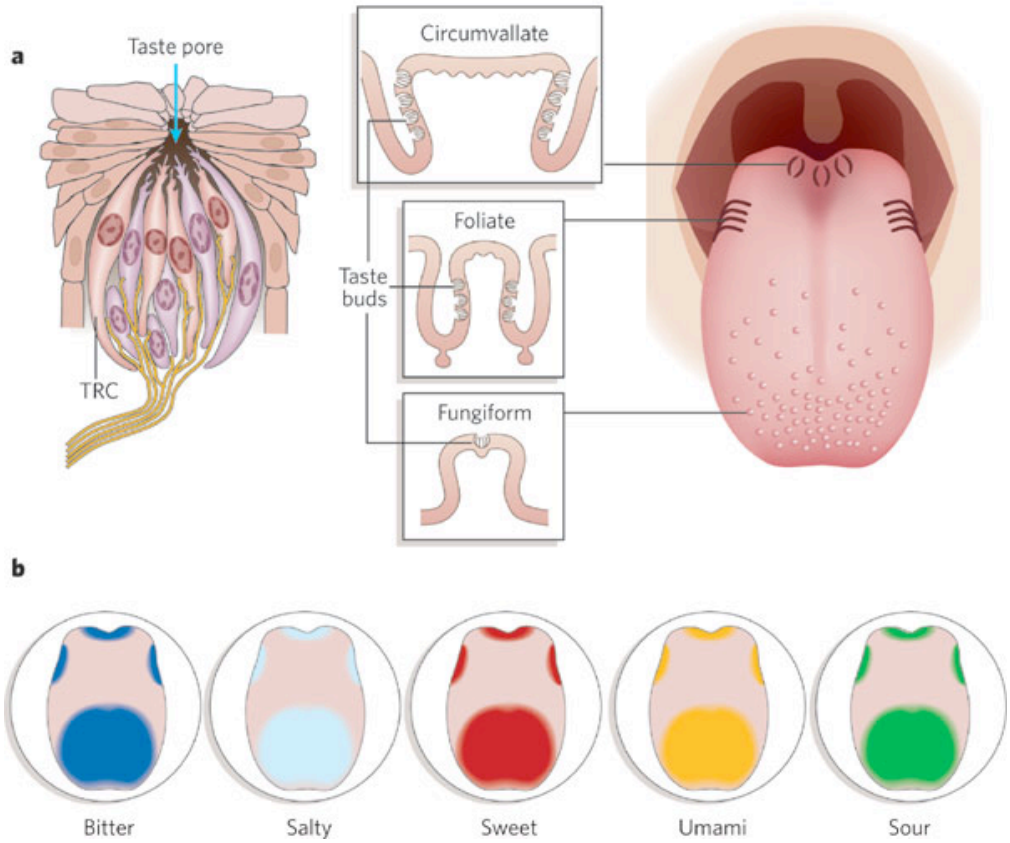

Figure 3 Illustration of the relationship between taste receptor cells (TRCs), taste buds, papillae and the tongue. (a) From the left-most panel: 50-100 TRCs form a taste bud; taste buds are found in papillae, which fall into three categories; different types of papillae are found on different parts of the tongue. (b) Contrary to popular belief, all regions of the tongue can detect all types of taste. Figure taken from *The receptors and cells for mammalian taste*, Nature 444, 288-294 (16 November 2006).

When food is in the mouth, small molecules (called ‘tastants’) dissolve in the saliva and physically interact with the taste receptors on the TRCs. Each different taste is detected by a different class of taste receptors. Over 30 bitter receptors have been described to date. It is currently thought that TRCs only contain one class of taste receptor; so one TRC will have bitter receptors, while another will have sweet receptors, and so on.

Activation of a given taste receptor is dependent upon a strong physical interaction with a tastant; for example, a single type of taste receptor might be strongly activated by some tastants, but only weakly activated by others. It follows that changes to the nature of either a tastant or a taste receptor can affect the strength of the interaction.

Glossary

Allele	Alternative form of a *gene*. The difference between alleles can vary, from a single *SNP* to larger differences.
Amino acid	The building blocks of *proteins*; humans have 22 amino acids. The sequence of amino acids in a protein is determined by information encoded within *genes* in *DNA*.
Balancing selection	A type of *natural selection* whereby different alleles are actively maintained within the gene pool. This usually happens as a result of *heterozygote advantage*.
Base and base pair	A 'letter' of *DNA* and part of a *nucleotide*. DNA has four bases, adenine (A); thymine (T); cytosine (C) and guanine (G). DNA is double-stranded, so every base has a 'complementary' counterpart and is part of a 'base-pair'. 'A' always complements 'T' and 'C' always complements 'G'. The words 'base' and 'nucleotide' are often used interchangeably.
Bioinformatics	The use of computers and mathematics to analyse biological data. Most frequently used these days in the context of using computers to analyse DNA sequence data.
Cell	The basic unit of almost all life: a fluid-filled sac that is surrounded by a thin, fatty membrane. Some organisms, such as bacteria or yeast, are just a single cell; others are much more complex and made of trillions of cells, such as humans. With very few exceptions, cells contain a *genome* and in 'higher organisms' the genome is organised into *chromosomes*, which are found within the *nucleus* of the cell.
Chromosome	A threadlike structure that contains a discrete portion of *DNA*. Humans have 23 pairs of chromosomes: 22 pairs of non-sex chromosomes ('autosomes') and two sex chromosomes (XX or XY), giving 46 chromosomes in total. One chromosome of each pair comes from an individual's mother, the other comes from their father.
Coding DNA	DNA within a genome which encodes a functional chemical, usually a protein.
Codon	A sequence of three *bases*, which encode a single *amino acid*. Given that a codon is three bases long and there are four possible bases at each position, there are $4^3 = 64$ different codons. See also *start codon* and *stop codon*.

Complementary	Describes the fact that every *base* on one strand of *DNA* has a complementary base on the other strand. 'A' complements 'T' (and vice versa) and 'C' complements 'G' (and vice versa).
Convergent evolution	Where two species independently evolve a similar characteristic or adaptation. For example, the wings of bats and birds; the eyes of octopuses and vertebrates and the fins of dolphins and fish.
Diploid	Refers to cells or organisms that have two sets of *chromosomes*, such as humans, who have one maternal set and one paternal set.
DNA	Deoxyribonucleic acid. A chemical made out of repeating pairs of *nucleotides*, which join up to create a very long, very thin structure that adopts the form of a *double-helix*. DNA is the chemical that carries the genetic information in almost all living things.
Dominant	An *allele* that will always determine the *phenotype* by masking the effect of the *recessive* allele.
Double helix	The structure adopted by a length of double-stranded *DNA*. A double helix is the shape you would get by twisting the ends of a ladder.
Electrophoresis	A technique that separates fragments of *DNA* according to their size. Typically, an electric current is passed through a buffer in which sits an agarose gel containing the DNA to be separated. Large fragments of DNA migrate slowly through the gel; small fragments migrate quickly.
Enzyme	A special type of *protein* that speeds up chemical reactions. Enzymes have very specific functions.
Evolution	The process whereby, over time, new species form through the action of *natural selection* on random *mutations* which lead to favourable adaptations.
Expressed/ expression	Refers to whether a *gene* is 'turned on' and being *transcribed*. This is a very common word in genetics/ molecular biology. Some examples would include: 'The TAS2R38 gene is only expressed in cells of the tongue,' or 'Expression of any gene is carefully regulated by cells'.
Gene	A portion of *DNA* that encodes another chemical, usually a *protein*. The human genome encodes around 23,000 genes. Genes are found at discrete *loci* along the length of a *chromosome*.

Genetic code	The set of rules whereby genetic information is *translated* into a sequence of *amino acids*, which form a *protein*. The genetic code defines how each *codon* is translated.
Genome	An individuals complete set of genetic information. A human's genome includes all of the *DNA* in their 46 *chromosomes* as well as *mitochondrial DNA*. About 2% of a human's DNA is 'coding DNA' the remainder is 'non-coding DNA'.
Genotype	A description of a portion of an individual's *DNA*.
Haploid	Refers to cells or organisms that have a single set of *chromosomes*. Bacteria and viruses are haploid, and so are the sex cells (sperm and eggs) of humans.
Heterozygote advantage	Where an individual that is *heterozygous* is 'fitter' (better suited to their environment) than homozygotes of either the dominant or recessive allele. A good example is sickle cell anaemia in parts of the world affected by malaria: heterozygotes have some protection from malaria while remaining unaffected by sickle cell disease.
Heterozygous	Indicates that for the *genotype* in question, an individual's maternal *DNA* is different to their paternal DNA. For example, an individual might have a taster *allele* for *TAS2R38* from their mother and a non-taster allele from their father.
Homozygous	Indicates that for the *genotype* in question, an individual's maternal *DNA* is the same as their paternal DNA. For example, an individual might have inherited a taster allele from both parents.
Karyotype	The number and appearance of a complete set of *chromosomes*, usually shown in a photograph. A karyotype typically shows chromosomes in their pairs and arranged in order of size from largest (chromosome 1) to smallest (the sex chromosomes).
Locus/ loci	A locus is part of a *chromosome* that contains a *gene* (plural: loci).
Maternal DNA	Refers to the *DNA* in *chromosomes* inherited from an individual's mother.
Messenger RNA	mRNA; an intermediate chemical, which is *transcribed* from *DNA* before being *translated* into a *protein*.

Mitochondrial DNA	Mitochondria (singular: mitochondrion) are small structures within cells that provide cells' energy; often referred to as the 'power houses' of a cell. Mitochondria have their own, discrete *DNA*, which is always passed down from the mother.
Mutation	A change in *DNA*. There are many different types of mutation, some which affect single nucleotides, others which affect large stretches of DNA.
Natural selection	The preferential selection of some individuals of a species over others due to slightly different characteristics which are favourable in the selected individuals. Natural selection ensures that individuals that are best adapted to their environment survive and breed, passing on their genes to the next generation.
Non-coding DNA	*DNA* within the *genome* that is not known to encode any other chemical.
Nucleotide	A repeating subunit that makes up *DNA*. Nucleotides include a phosphate group, a sugar (deoxyribose) and a '*base*'. The words 'nucleotide' and 'base' are often used interchangeably.
Nucleus	A discrete structure within a *cell*, which contains the cell's *chromosomes*.
Paternal DNA	Refers to the *DNA* in *chromosomes* inherited from an individual's father.
PCR	Polymerase chain reaction. A method of amplifying a specific region of *DNA* between a pair of *primers*. PCR has three steps, denaturation, annealing and extension, which occur at different temperatures. A *thermal cycler* moves between the temperatures, facilitating the PCR reaction. The DNA is copied by *Taq polymerase*.
Phenotype	The physical manifestation of a *genotype*.
Polymerase	A specialised enzyme that copies nucleic acids. *DNA* polymerases copy DNA; RNA polymerases copy RNA.
Primer	A short (usually around 20 nucleotides), synthetic, single-stranded piece of DNA used in PCR. Also called *oligonucleotide* or *oligo*.
Protein	A complex chemical made up of a string of *amino acids*, folded into a precise three-dimensional shape. Proteins are encoded by *genes* in *DNA*.
Recessive	An *allele* that will only determine the *phenotype* if it is *homozygous* and will otherwise be masked by the *dominant* allele.

Restriction enzyme	A specialised *enzyme* that cuts *DNA* according to a specific sequence. For example, HaeIII will only cut DNA at the sequence CCGG. Restriction enzymes may leave either blunt ends or *sticky ends*, depending on how they cleave the DNA.
RNA	A chemical similar to *DNA*, but with some important differences. RNA is (usually) single-stranded, and instead of a 'T' (thymine), RNA uses 'U' (uracil). There are different forms of RNA, including messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA).
Selective advantage	An adaptation or *mutation* that means that the affected individual is more likely to breed and pass it on to the next generation.
Speciation	The process of a new species forming. This happens when an ancestral population is separated in some way into two distinct populations which can evolve independently of each other. Eventually, they become so distinct, morphologically and/ or genetically that they are separate species.
SNP	Single nucleotide polymorphism, SNP, pronounced 'snip'. A change in a single base pair within *DNA*. A SNP must be reasonably common within a population; otherwise it is considered a *mutation*.
Start codon	A special *codon* sequence that indicates the beginning of a protein-encoding gene. In most organisms there is only one start codon, which is ATG.
Sticky end	A feature of *DNA* that has been cut by a certain type of *restriction enzyme*. One strand of DNA slightly extends (by just a few *nucleotides*, usually 3) beyond the other strand, forming an overhang.
Stop codon	A special *codon* sequence that indicates the end of a protein-encoding gene. There are three stop codons in humans.
Taq	A thermophilic ('heat loving') *DNA polymerase* that can withstand high temperatures. *Taq* is one of the most commonly used polymerases in *PCR*.
Transcription	The process of copying a strand of *DNA* into a molecule of *mRNA*.
Translation	The process of reading the *codons* in an *mRNA* molecule and 'translating' them into *amino acids*, thus creating a string of amino acids which ultimately form a *protein*.

Y chromosome	One of the sex *chromosomes*. In humans, men have one X chromosome and one Y chromosome. The Y chromosome is one of the smallest human chromosomes.
X chromosome	One of the sex *chromosomes*. In humans, women have two X chromosomes. The X chromosome is one of the largest human chromosomes.

Other online glossaries are widely available and may be of use. You might like to look at:

- http://insidedna.org/content/?tag=glossary
- http://www.yourgenome.org/glossary/
- http://www.geneticalliance.org.uk/glossary.htm
- http://www.phgfoundation.org/pages/resources/glossary.htm