Chapter 1

Defining Important Concepts

1.1 KEY CONCEPTS IN MOLECULAR BIOLOGY FOR THE STUDY OF HUMAN NUTRITION

Until very recently, the study of human nutrition and molecular biology were considered to be mutually exclusive domains within the biological sciences. This is simply no longer the case. Today, the leading edge of our endeavor to explain the very nature of mankind, and our ascent to planetary dominance blends both nutrition and molecular biology into the fields of nutritional genetics and nutrigenomics. These new disciplines exploit our knowledge of the human genome and its variability to explain how nutrients, their dependent proteins, and encoding genes conspire to forge and maintain our species. These interactions not only help explain the etiology of many diseases, but also they provide a framework for gaining a better understanding of the likely evolution of our species. Human evolution was forged out of our ancestors obligate need to forage for chemical nutrients that varied in their abundance according to habitat and season. This forced early humans to find and compete for limited resources; humans that foraged optimally and competed most successfully for those resources were fitter and more able to reproduce and, hence, could pass on their genetic material to their progeny. In other words, they were selected for. This process of evolution is characterized by a change in gene frequency over time, but what are genes, and how do they lead to the expression of traits, the summation of which produces the state of "being human?" To understand this process, we need to examine the building blocks of our genetic code.

1.1.1 Molecular Structure of DNA

Polymeric DNA is composed of four different nucleotides. Each nucleotide consists of a 2′-deoxyribose sugar, purine or pyrimidine base, and phosphate moiety. Purine bases are either adenine or guanine, whereas pyrimidine bases are either thymine or cytosine. When a base is linked to the 1′ carbon of the deoxyribose sugar, it is referred to as a nucleoside.

Molecular Nutrition and Genomics: Nutrition and the Ascent of Humankind, Edited by Mark Lucock

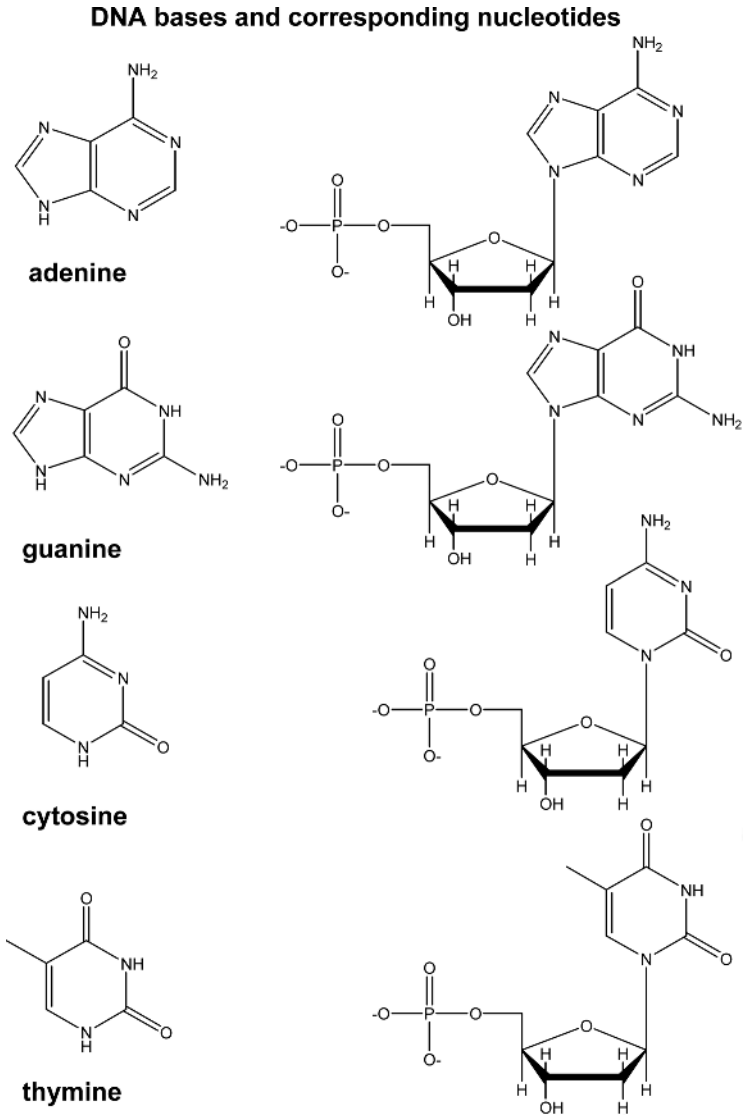

Figure 1.1. *Bases adenine, guanine, cytosine, and thymine along with their corresponding nucleotides that form the building blocks of DNA.*

When, in addition, phosphate moieties are attached to the sugar, the structure is referred to as a nucleotide.

Nucleotide triphosphates (Figure 1.1) of adenine (A), guanine (G), cytosine (C), and thymine (T) are polymerized to form DNA via phosphodiester bond formation between the 5′ phosphate of one nucleotide and the 3′ hydroxyl group of the next nucleotide. The sequence of bases is what encodes the genetic blueprint for life. It can be read in the $5' \rightarrow 3'$ or the $3' \rightarrow 5'$ direction.

The primary sequence of DNA permits a three-dimensional structure to form, which is represented by a double helix. The sugar–phosphate linkage forms the molecular backbone

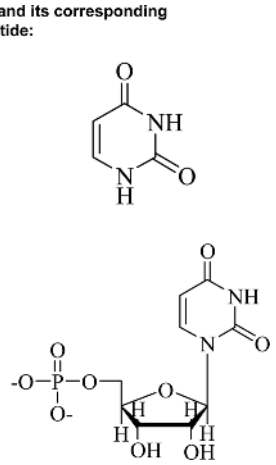

Figure 1.2. *RNA is the same as DNA except RNA contains uracil, whereas DNA contains thymine. Additionally, in RNA, ribose replaces DNA's 2-deoxyribose.*

of this structure. The bases face inward and stabilize the double helix via hydrogen bonds between adjacent T and A bases, and again between adjacent G and C bases. This base pairing is specific, and purine always interacts with pyrimidine, a phenomenon referred to as "complementary base pairing." The double helix is right-handed with a turn every 10 bases. Examination of the structure reveals a major molecular groove, which facilitates protein interactions.

Complimentary base pairing ensures that the sequence of one DNA strand predicts the base sequence of the other. This simple fact is what permits the fidelity of the genetic blueprint to be preserved during replication of DNA as part of cell division, and during the expression of genes.

Expression of DNA, which is the conversion of the base sequence blueprint into an amino acid sequence within a functional protein, requires as a first step, the transcription of the DNA sequence into an RNA transcript. RNA is the same as DNA except RNA contains uracil, whereas DNA contains thymine (Figure 1.2). Additionally, in RNA, ribose replaces DNA's 2-deoxyribose. The RNA transcript is referred to as messenger RNA (mRNA). mRNA is then translated into a protein on the ribosome—transfer RNAs (tRNA) are small molecules that coordinate individual amino acids to form proteins that have been specified by the mRNA sequence.

This phenomenon of gene expression in which the biological data encoded by a gene is made available in terms of a functional protein is referred to as "the central dogma." That is, information is passed from DNA to RNA to protein.

Humans contain around 23,000 genes on 23 chromosomes. These genes are separated by intergenic (noncoding) DNA. Although a gene is the fundamental unit of information in that a single gene codes for a single polypeptide, higher organisms such as man also have multigene families. In their simplest form, a gene family contains more than one copy of a gene where its expression product is required in large amounts. Complex multigene families also exist. These yield similar, but distinct, proteins with related function, for example, the globin polypeptides.

To orchestrate gene regulation according to cellular need, gene promoter regions exist upstream from the coding region of a gene. Promoter sites bind the enzyme for synthesizing the RNA transcript (RNA polymerase II) and any associated transcription factors that are required to initiate mRNA synthesis. Promoter regions usually contain a TATA box around 25 base pairs upstream from the site at which transcription commences. Transcription factors bind DNA around the TATA box and orchestrate the binding of RNA polymerase II. RNA polymerases I and III are associated with transcription of ribosomal RNAs and genes encoding tRNAs, respectively.

Transcription factors can be considered as modular molecules that contain DNA binding, dimerization, and transactivation modalities. These regulatory factors exhibit characteristic structural motifs. The DNA binding modality contains three potential motifs: zinc fingers, basic domains, and helix-turn-helix motifs. Dimerization modalities contain two motifs: leucine zippers and helix-loop-helix structural motifs. The formation of homo- and heterodimers leads to transcription factor variation and, hence, a diversity of function. Transcription factors can act to both initiate and repress transcription.

Genes do not contain a continuous code; rather they are split into coding regions known as exons and noncoding regions known as introns. Introns are removed from the RNA transcript by a process referred to as splicing. This process occurs before protein synthesis.

Some genes have accumulated nonsense errors in their base sequence and no longer function. These archaic genes are referred to as pseudogenes.

1.1.2 Molecular Encryption

The base sequence of DNA encodes the amino acid sequence of a polypeptide via the intermediate polymer—RNA. Amino acids are encrypted by 64 triplets; each triplet represents a sequence of three DNA bases and is known as a codon. Within a gene, each set of codons that builds up to form a genetic unit of information is referred to as a reading frame. The reading frame is determined by "initiation" and "stop" codons. In between these initiation and stop codons, one has what is referred to as an "open reading frame."

As the four nucleic acid bases can combine to form 64 permutations of codon (Table 1.1), but only 20 amino acids exist in proteins, all amino acids save tryptophan and methionine are encrypted by more than one codon. This fact is why the genetic code is often referred to as having built-in degeneracy or redundancy. Sixty-one codons encode amino acids, and three are used to terminate protein synthesis (UAA, UGA, UAG). The codon for methionine (AUG) encodes initiation of protein expression. Clearly, all nascent polypeptides therefore start with methionine.

1.1.3 Organizing the Human Genome

DNA is organized into cellular structures called chromosomes that are only visible after they have replicated during the cell cycle. Unique structures found at the end of the chromosome are known as telomeres. Telomeres consist of short repetitive DNA sequences. What is of interest in regard to telomeres is the fact that the number of repeat sequences declines with age in somatic cells, but in cancer and germ cells, the enzyme telomerase maintains telomere length (see later). Telomeres are purposeful as they prevent recombination of the chromosomes.

Table 1.1. Matrix showing how amino acids are encrypted by specific three base codons within RNA.

Amino acid/signal encrypted by codon—the genetic code					
Initial base at 5′ end	Middle base				Third base at 3′ end
	U	C	A	G	
U	Phe UUU	Ser UCU	Tyr UAU	Cys UGU	U
U	Phe UUC	Ser UCC	Tyr UAC	Cys UGC	C
U	Leu UUA	Ser UCA	Stop UAA	Stop UGA	A
U	Leu UUG	Ser UCG	Stop UAG	Trp UGG	G
C	Leu CUU	Pro CCU	His CAU	Arg CGU	U
C	Leu CUC	Pro CCC	His CAC	Arg CGC	C
C	Leu CUA	Pro CCA	Gln CAA	Arg CGA	A
C	Leu CUG	Pro CCG	Gln CAG	Arg CGG	G
A	Ile AUU	Thr ACU	Asn AAU	Ser AGU	U
A	Ile AUC	Thr ACC	Asn AAC	Ser AGC	C
A	Ile AUA	Thr ACA	Lys AAA	Arg AGA	A
A	Met AUG	Thr ACG	Lys AAG	Arg AGG	G
G	Val GUU	Ala GCU	Asp GAU	Gly GGU	U
G	Val GUC	Ala GCC	Asp GAC	Gly GGC	C
G	Val GUA	Ala GCA	Glu GAA	Gly GGA	A
G	Val GUG	Ala GCG	Glu GAG	Gly GGG	G

Chromosomes are actually an aggregation of proteins and DNA. This material is referred to as chromatin. Chromatin that is inactive is known as heterochromatin, whereas active chromatin that permits RNA transcription is known as euchromatin (Figure 1.3). Human gametes are haploid and contain 23 chromosomes, whereas non-sex cells (somatic cells) are diploid and contain 46 chromosomes.

It has been estimated that the entire human genome comprises around 3 billion base pairs. However, the 23,000 human genes account for only a fraction of our entire cellular DNA—the rest is extragenic or "junk" DNA.

As part of the cell cycle, the cell will divide. This entails that chromosomes are replicated. The DNA is copied in the $5' \rightarrow 3'$ direction by the enzyme DNA polymerase using single-stranded DNA as a template.

1.1.4 DNA Variation: The Provision of Biological Diversity

Errors in the fidelity of DNA replication along with physical and chemical agents all potentially induce mutations in the DNA sequence. If they affect coding sequences, this may influence the function of any expressed protein. That is, the "phenotype" may alter. The types of mutation include missense, nonsense, and frameshift mutations. All are classified as point mutations. The latter two point mutations have the most serious consequences for the expressed proteins function.

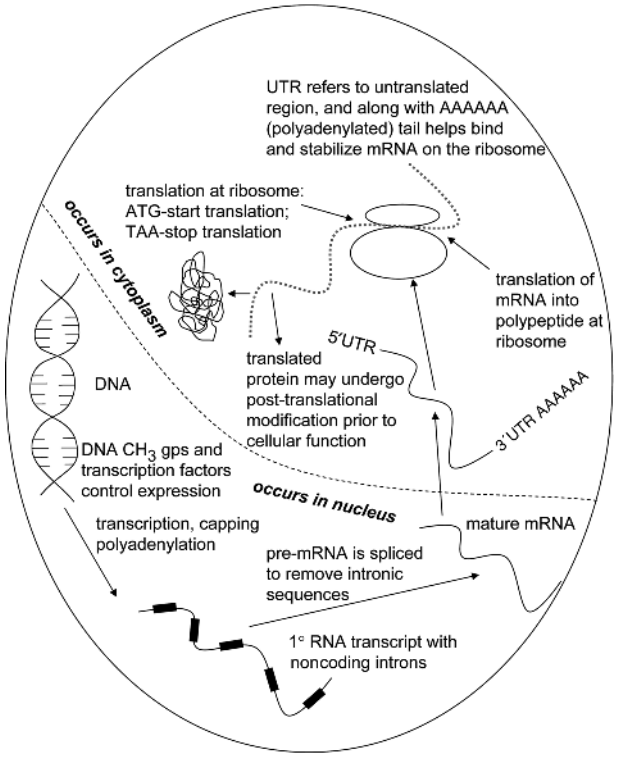

Figure 1.3. *Simplified schematic shows the process of gene expression.*

As living organisms are exposed to so many mutagens, life has evolved elaborate DNA repair mechanisms as a counter-measure. The mechanisms include excision-, direct-, and mismatch repair, and they are discussed at length later. This is one area where as an example, antioxidant nutrients prove useful, although they are only one form of defense in this cellular war that is continuously waged within every one of us.

Not all mutations are necessarily bad. A gene that has, for example, an A where previously there was a G, may, under the influence of evolution, become more frequent in successive generations. That is, it is advantageous to possess this mutation in a given environment because it improves reproductive efficiency. Perhaps the protein change provides a selective advantage. As a hypothetical example, maybe the mutated protein in question leads to a more efficient form of an intestinal binding protein specific for a trace nutrient that is important in sperm motility. This provides an easy visualization of how a beneficial trait will be selected for by nature.

Many people use the term *mutation*, but as I have said, not all mutations are deleterious, so the term *polymorphism* is more appropriate to use and simply means variant.

If you examine the genetic code within any population, you will find an enormous amount of variation. This stems from mutations and provides the fodder for the process of natural selection first described by Charles Darwin. Of course, although Darwin made his deductions from an examination of whole organisms, we are examining the same phenomenon, but from a molecular perspective. Maintaining population variation by natural selection alone is unlikely, because much of the variation within a population is selectively neutral,

and subject to random change or what evolutionary biologists refer to as "drift." Drift is interesting because it can promote or eradicate extremely rare traits, particularly in small populations, which relates to the founder effect described earlier. In North America, the Anabaptist Amish and Hutterite communities give recent human examples of small culturally isolated populations that grew in size, and that now have a unique genetic signature with unrepresentative gene frequencies. The Amish grew from a founder population of around 200 and the Hutterites from 443 people. Both communities were closed to immigration. As a further example, Dutch immigrants arrived in South Africa during the seventeenth century, and although they were a small group, they were interesting in that they carried several rare genetic disorders that were not representative of the parent population from which they were drawn. The Dutch Afrikaner population grew rapidly and maintained the high frequency of these abnormal genetic traits. For example, a single couple of émigrés from Holland in the 1680s is now responsible for around 30,000 Afrikaners carrying the trait for porphyria variegata.

In the new synthesis of neo-Darwinian evolution, selection is examined in the context of how it acts on the fundamental genetic unit—the allele. We inherit a copy of any given gene from each of our parents. If neither copy (allele) contains, for example, an A where there is normally a G, then the genotype is wildtype. If one allele contains an A and the other allele a G, the genotype is referred to as heterozygous. If both alleles contain the abnormal (mutant) A, the genotype is homozygous recessive. By considering the frequency of polymorphic alleles, we can look at genetic evolution in a quantitative manner. For example, it is possible to work out how many generations it would take for a given level of selection pressure to substitute one allele for another. This is different to the view many people have of natural selection, because we are looking at the selection of molecular rather than phenotypic traits. As a consequence, scientists are now very interested in the relatively new idea of "selfish genes." Selfish genes and not phenotypes or genotypes span the generations. Consider that phenotypes senesce and die, whereas genotypes are determined as a function of meiosis—only the allele is immortal.

There is considerable debate as to the relative contribution of the following three phenomena as drivers of human evolution: (1) mutational induction of new alleles, (2) drift leading to selectively neutral random changes in allele frequency, and (3) natural selection forcing directional allele change. To put the importance of these evolutionary mechanisms into perspective, what makes us unique as individuals is the subtle, yet extensive variation in our genetic codes. There are in fact several alleles for any given gene in the human genome, emphasizing the seemingly infinite number of possibilities for individuality.

When wildtype and homozygous recessive genotypes are less fit than heterozygotes, then both wildtype and mutant alleles will be maintained in a population. This is known as a heterozygote advantage or balanced selection. The example that is always given to demonstrate this phenomenon describes how a valine substitution for glutamic acid in the hemoglobin molecule can protect individuals from sickle cell anemia. The "mutant" Hb*S* allele is particularly common where malaria is endemic because heterozygosity (Hb*A*Hb*S*) for this trait protects against this life-threatening parasitic infection. Although wildtype (Hb*A*Hb*A*) individuals are less able to contend with *falcoparium* malaria, homozygous recessive individuals (Hb*S*Hb*S*) suffer from overt sickle cell anemia, a debilitating and often lethal condition. Despite this awful condition, the frequency of Hb*S*Hb*S* individuals in parts of Africa within the malaria belt can reach 4% of the population. Clearly, the advantages of maintaining heterozygosity for this trait within the population are high. Another example of the

heterozygote advantage is given by Tay–Sachs disease in which heterozygosity may confer a degree of protection against tuberculosis despite the recessive genotype being fatal by age 4. However, one of the most interesting and perhaps bizarre examples of a putative heterozygote advantage is given later in a discussion of human prion disease and cannibalism (see Chapter 7).

1.1.5 Population Genetics and the Hardy–Weinberg Equilibrium

If we want to examine allelic frequency within a population, and the forces that impact upon and change either the frequency of gene alleles or the genotypes, we can. The Hardy–Weinberg equilibrium permits us to calculate the expected genotype frequency from the allele frequency within the same population and the allele frequency from the known genotype. To accomplish this, we make certain assumptions: Mating occurs at random; reproductive efficiency is constant; no mutations are occurring; there is no effect on the population and its genotypes through selection pressure; and there is no effect on the population and its genotypes through inward or outward migration.

If we apply the Hardy–Weinberg equation, and the population we are studying does not fit Hardy–Weinberg predictions, then we have substantial evidence that some force like natural selection is acting on the population.

Hardy–Weinberg equation:

$$p^2 + 2pq + q^2 = 1$$

As a first step to see whether a population fits the Hardy–Weinberg equation, we need to calculate the allele frequencies. Let's look at this with some real data generated in the author's laboratory. 5,10-methylenetetrahydrofolate reductase (5,10MTHFR) is a folic acid-dependent enzyme that exists in polymorphic form. It is discussed extensively later in this book because it exhibits an important nutrient–gene interaction that impacts upon occlusive vascular disease, cancer, and birth defects. 5,10MTHFR helps regulate both DNA and homocysteine metabolism. The gene encoding 5,10MTHFR exhibits a common C-to-T substitution at nucleotide 677 (this is often written as 677C $\rightarrow$ T MTHFR or C677T-MTHFR). The C-to-T substitution at nucleotide 677 converts an alanine to a valine residue in the functional protein. This kind of polymorphism is often referred to as a single nucleotide polymorphism or SNP.

The possible genotypes are therefore wildtype—CC; heterozygote—CT; and homozygote recessive—TT. In a population of control patients recruited into a study to examine how this gene influenced vascular disease, we counted 41 CC, 46 CT, and 14 TT individuals. We can measure the allele frequency easily. Simply add the number of copies of each allele in the control population, and express it as a frequency. Remember that the population is diploid, and therefore, individuals have $2N$ alleles; the heterozygote has, as an example, one C allele and one T allele. Therefore, the frequency of the C allele is given by

$$(n_{CT} + 2n_{CC})/2N$$

Therefore, in our control population, $46 + 82/202 = 0.63$.

The frequency of the wildtype MTHFR-677C allele is 0.63, and by default, the frequency of the mutant MTHFR-677T allele is 0.37.

The frequency we obtain for the wildtype C allele is referred to as p, whereas the corresponding non-p allele frequency is termed q. As I have shown above, $p + q =$ unity. We can use this information to work out the expected genotype frequencies as predicted by the Hardy–Weinberg equation. If we examine the two alleles C and T that have frequencies of p and q, respectively, then we can expect a CC wildtype frequency of p^2, a CT heterozygote frequency of $2pq$, and a TT recessive homozygote frequency of q^2. Thus, $p^2 + 2pq + q^2 = 1(0.63^2 + 2(0.63 \times 0.37) + 0.37^2 = 1$.

This equation shows that when the frequency of a mutant allele is very low, the occurrence of the recessive homozygous genotype is extremely low, as in many rare genetic diseases. In the case of such rare genetic diseases, the mutant alleles tend to be concealed within heterozygotes where they are not expressed, so selection pressures cannot act against them. Consider this in the context of allele immortality as alluded to earlier.

As mentioned, nature acts to distort the idealized frequencies that are predicted by the Hardy–Weinberg equation. Some causes of this include:

- Ingress of migrants with a different allele frequency
- Natural selection against fertility or against survival to reproductive age of a certain genotype
- Subpopulation mating—in extreme situations, inbreeding
- Mutations creating new alleles
- Drift

The usual way to compare an observed genotype frequency with an expected one, assuming the Hardy–Weinberg equilibrium holds, is to perform a chi-square test for goodness of fit.

1.2 THE INHERITANCE OF GENETIC PACKETS OF INFORMATION

When alleles are juxtaposed on the DNA molecule, they are usually inherited together and do not segregate. The typical packet of genetic information that is inherited as a consequence of meiotic recombination might typically contain in excess of 20,000 base pairs.

Any given packet of genetic information will contain many polymorphisms. These SNPs are considered to be in linkage disequilibrium (LD). That is they are nonrandomly associated with nearby alleles. LD is associated with the physical distance on the DNA molecule between the loci of alleles, and it is under the variable influence of recombination.

A single packet of genetic information is referred to as a haplotype. Haplotype size within a population varies according to meiotic recombination, such that where ancestral human populations that are large in number, and have remained so for a significant period, will in all probability have smaller haplotypes (shorter DNA packets) and hence a lower LD. This stems from the greater number of genetic influences (mutations and recombinations) that have occurred in such populations and the effect that these events have on LD decay.

In the context of what follows on the ascent of man, African populations exhibit a larger number of haplotypes and more diverse LD patterns than non-African humans, who have

evolved from small founder groups into new environments that differ significantly from the ancestral one. This greater genetic diversity among African populations is consistent with the view that modern man emerged out of an African evolutionary crucible.

Scientists also often refer to the "molecular clock" when investigating the evolutionary past and its various processes. To establish molecular dates, it is necessary to quantify the genetic distance between species, and then use a calibration rate such as the number of genetic changes expected per unit time. This permits one to convert genetic distance to time. Sophisticated models for achieving this include maximum likelihood (4,5) and Bayesian approaches (6). At the end of the day, the reliability of all molecular clock methods and their ability to provide information on the mechanisms that drive molecular evolution depends on the accuracy of the estimated genetic distance and the appropriateness of the calibration rate. See the panel on mitochondrial DNA (mtDNA) and elucidating "Eve."

1.3 A BRIEF OVERVIEW OF EVOLUTIONARY BIOLOGY AND THE ASCENT OF MAN

How can one briefly overview such a topic when it is possible to write volumes on the subject? In an excellent and fairly concise review of the "Genetics and making of *Homo sapiens*," which appeared in the journal *Nature* (7), the author, Sean Carroll, cites a passage from Shakespeare:

> What is man,
> If his chief good and the market of his time
> Be but to sleep and feed? A beast, no more.
> Sure, he that made us with such large discourse,
> Looking before and after, gave us not
> That capability and god-like reason
> To fust in us unused
>
> —W. Shakespeare, Hamlet IV:iv

We recognize that all human races presently on Earth are part of the same species, and that around 4 million years ago, a hominoid ape-like ancestor evolved out into three lineages—chimpanzees, gorillas, and early humans. Perhaps the best-known artifact from this time was discovered at Hadar, Ethiopia, and has been affectionately named "Lucy." Lucy is almost 4 million years old, and although she seems to be built in a robust ape-like manner, she was bipedal and walked upright on two legs as we do today.

It seems likely that bipedalism evolved early as a mechanism to free hands for the dexterous manipulation of tools and weaponry. Many of the attributes that man evolved such as increased intellect and brain size are discussed later in this book in the context of nutrition. Some of the oldest stone tools date back 2.5 million years and are associated with the fossils of our bipedal ancestor, *Homo habilis*. A million years later, the early human brain had enlarged and permitted the development of more highly refined tools.

These evolved characteristics are associated with *Homo erectus*. This species began a migration out of Africa about three quarters of a million years ago. However, within Africa, *Homo erectus* continued to evolve into modern man (*Homo sapiens*). This process was

complete by around 100,000 to 200,000 years ago. *Homo sapiens* then migrated out from Africa and eventually supplanted *Homo erectus*. This simple view ignores the possibility that subspecies may have existed.

The cold climate that prevailed during the quaternary ice age in Eurasia probably gave rise to the Neanderthals (*Homo neanderthalensis*). These stoutly built people had heavy brow ridges above their eyes and were well evolved to survive the cold. They lived from 120,000 to 35,000 years ago and are considered to be *Homo sapiens*. Although they had extremely large brains, and well-evolved cultural practices, they eventually gave way to Cro-Magnon man who had appeared right across Europe by 35,000 years ago. This is a parallel time frame to the colonization of Asia and Australasia by what one would consider to be an anatomically modern form of *Homo sapiens* (Figures 1.4 and 1.5).

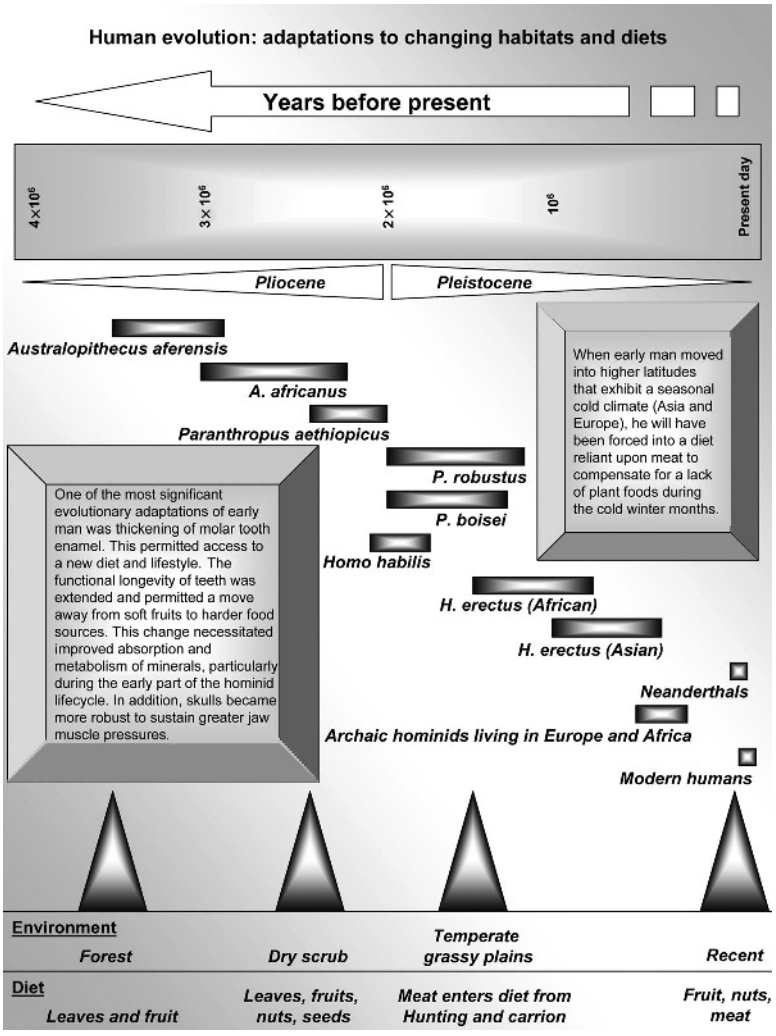

Figure 1.4. *The exposure of ancestral man to changing habitats and hence diets over the past 4 million years has played a role in our evolution as a species.*

Mitochondrial DNA—The Search for Eve

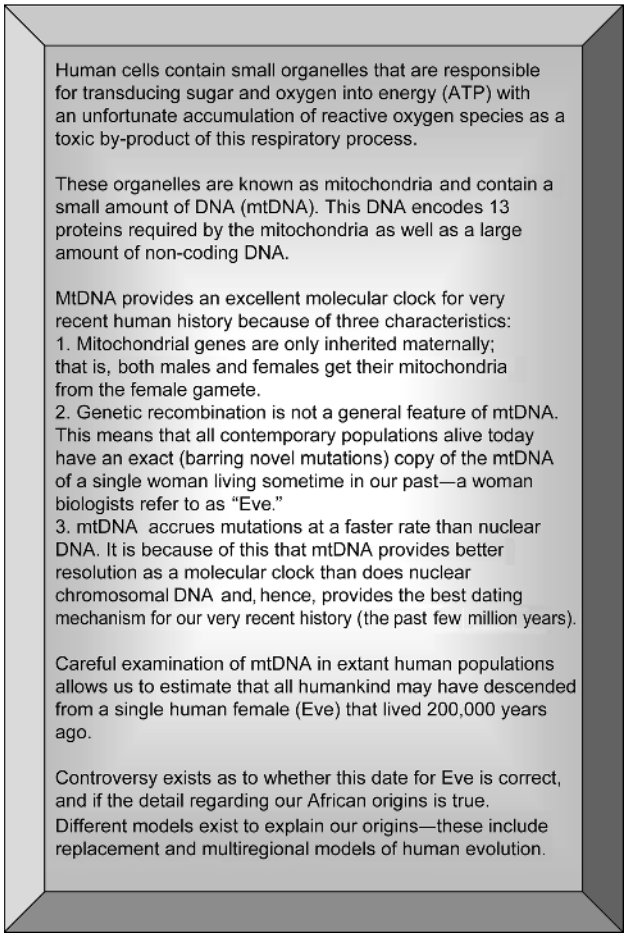

Figure 1.5. *The concept of mitochondrial Eve is based on the molecular clock inherent in the maternal mitochondrial genome. The clock allows us to trace the female lineage back to the original ancestor of modern man.*

We will never know the complete story of our recent past, but there is consensus that as our brains grew, so to did our ability to produce and use tools and weapons. The skills to do this are necessarily learned. The ability to pass on and acquire such important information for survival probably acted as a driving force for the natural selection of intelligence, effective communication, and hence language. It is interesting to note, however, that the left–right asymmetry in Broca's area of the frontal lobe of the neo-cortex, an area that is associated with language ability, occurs in chimpanzees, bonobos, and gorillas, as well as in humans. This means the neuro-anatomical substrate of left-hemisphere dominance for speech was in place before the origin of hominins (7,8). Wernicke's posterior receptive language area in the temporal lobe is responsible for speech and gesture,

as well as for musical talent, and again shows left-hemisphere dominance. Evidence from *Homo erectus* and *Homo neanderthalensis* endocasts as well as from chimpanzees show the presence of this shared asymmetry, again indicating its presence before the divergence of hominins.

From a physical viewpoint, the trend in the evolution of modern man was toward larger body mass, larger brains, longer legs relative to trunk, and smaller dentition. At a subtler, molecular level, the genetics of human evolution are of tremendous interest and yet, at the same time, are extraordinarily complex. With technological advances, however, we are now able to gain a far better idea of exactly what we are and how we came about (Figures 1.6–1.8).

1.4 THE –*OMICS* REVOLUTION

The new technologies that embrace the term *–omics* have evolved to address increasingly complex biological questions arising out of the postgenomics era. I describe some of these advanced techniques toward the end of this book. Briefly they encompass techniques like DNA microarray technology, real-time polymerase chain reaction, denaturing hplc, two-dimensional (2-D) protein electrophoresis coupled with matrix-assisted laser desorption/ionization–time-of-flight (MALDI–TOF) mass spectrometry, and *in silico* bioinformatics. These state-of-the-art techniques permit us to venture into the world of proteomics, transcriptomics, metabolomics, nutrigenomics, methylomics, and perhaps at the ultimate level to understand the "interactome." The interactome is defined as the sum of all protein interactions in the cell. A graphical representation of a typical "interaction map" looks like a massive aggregated collection of hairy dandelion seeds and is hugely complex (Figure 1.9). Such interactomes are often simplified into "functional interaction maps" in which proteins are allocated to functional categories (i.e., protein degradation, carbohydrate metabolism, and signal transduction). This provides a simpler three-dimensional (3-D) rendering of the network of cellular functions.

At the leading edge of scientific endeavor, it is becoming increasingly difficult to pigeonhole one's research interest. This book is a prime example of how interests in food, nutrition, genetics, molecular biology, clinical medicine, evolutionary theory, and anthropology come together to address the most fundamental of all human questions: "What does being human mean, and how did the condition arise?" Essentially, what is the meaning of life?

As an educator within our university system, I became frustrated by the notion that human nutrition is simply all about food, its constituents, and how they prevent disease or contribute, to it. As this book proves, nutrition is a far more diverse and philosophically deep subject than many students (and educators) think, and one that has never been more relevant than it is today. The two novel subdisciplines within nutrition that are now increasingly important are nutrigenomics and nutritional genetics. Peter Gillies (9) has defined these terms as follows: "Nutrigenomics refers to the prospective analysis of differences among nutrients with regard to the regulation of gene expression. In this context, nutrigenomics is a discovery science driven by the paradigms of molecular biology, enabled by microarray technology, and integrated on an informatics platform" (10,11). Gillies goes on to define nutrigenetics, or what many people refer to as nutritional genetics, as "the retrospective analysis of genetic variations among individuals with regard to their clinical response to specific nutrients. In this context, nutrigenetics is an applied science driven by the paradigms of

The Rise of Modern Man: "Out of Africa Replacement" or a "Multiregional" Evolution?

The Concepts:

The "Eve" hypothesis to support an "out of Africa" model for human evolution has much to commend it; this "out of Africa replacement" scenario is based on the molecular relatedness between African groups from within the continent, and between groups from Africa and other regions. The "out of Africa replacement model" assumes a human population originated in Africa 150,000 years ago with Eve as its source, and then radiated out, supplanting other human groups en route. However, this is only one possible option to account for our origins. An alternative model proposes that modern man evolved slowly from ancestral humans in many different areas of the world. This is the "multiregional model". This paradigm is based on populations gradually evolving into modern humans in many different locations. Implicit in this model is that differences in physiognomy (skin and hair color, build, etc.) between geographically distinct human populations have a longer adaptive chronology than would be implied by an "out of Africa" model. A third paradigm exists—"the assimilation model" contends that our origins were in fact African, but regional groups of archaic humans like the Neanderthals made a substantive contribution to our existing gene pool.

The Evidence:

The debate between replacement (perhaps more logically referred to as uniregional) and multiregional models remains to be resolved: In a 2002 review, Satta and Takahata [S1] weigh up the evidence for both models and conclude that the uniregional model is the most likely option to have occurred. Similarly, a recent 2005 paper by Ray and colleagues [S2] describes the use of multilocus genetic data to infer the geographic origin of humans and distinguish between uni- and multiregional models. Using 377 genetic markers, they claim that East Africa is the most likely place of origin for modern humans and the source of human expansion into the Old World. However, Eswaran and colleagues [S3] prefer a model in which the modern human phenotype

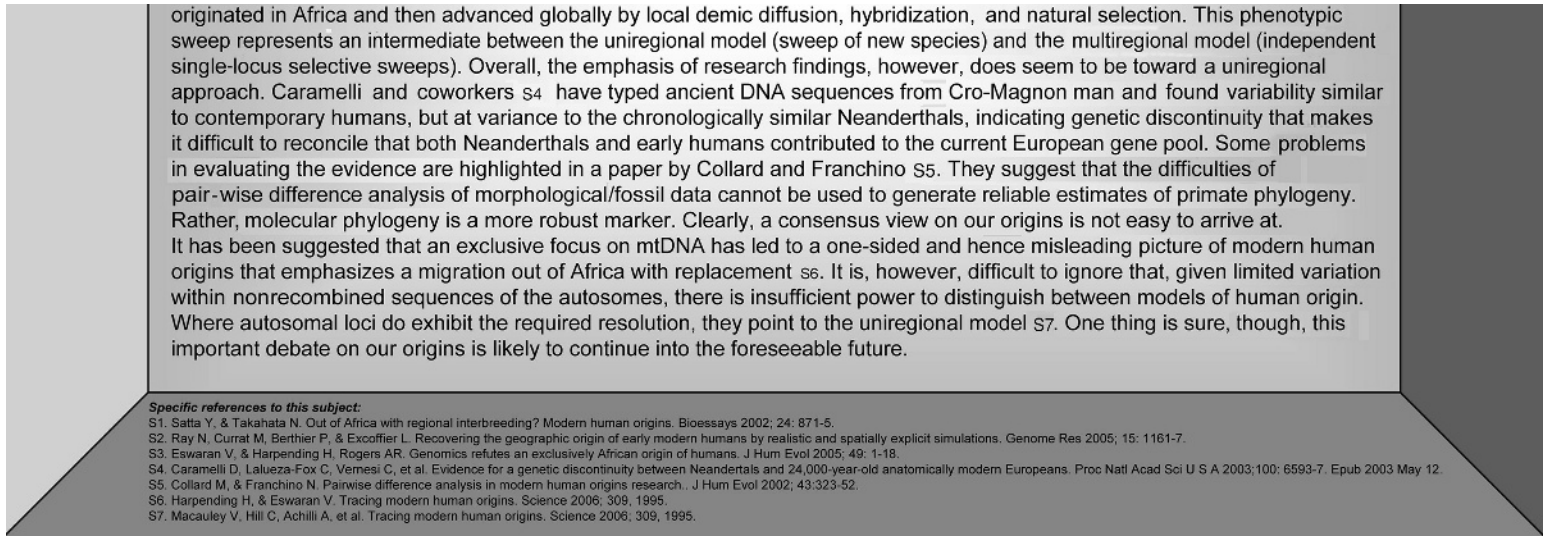

Figure 1.6. *Contemporary theories to explain the recent ascent of humankind are based on two models: an "out of Africa replacement model" in which we evolved as a species in Africa, and radiated out to colonize the planet, and a model in which "multiregional evolution" of our species occurred.*

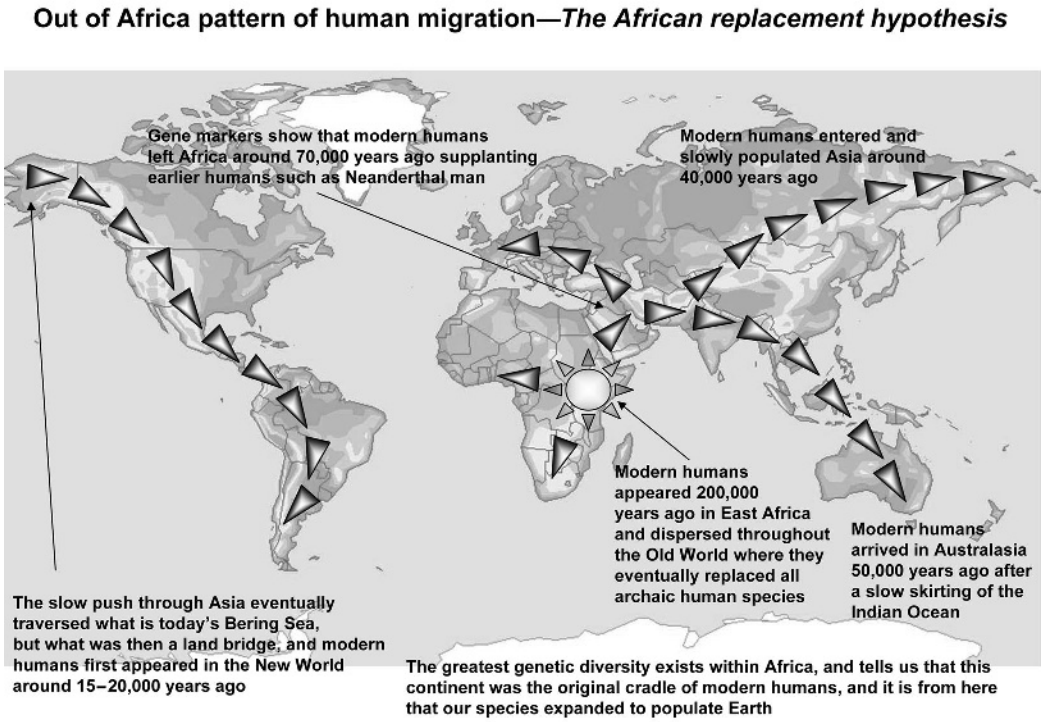

Figure 1.7. *This figure shows the pattern of migratory radiation that our species took when leaving Africa based on the "out of Africa replacement model," which is the favored paradigm for our recent evolutionary past.*

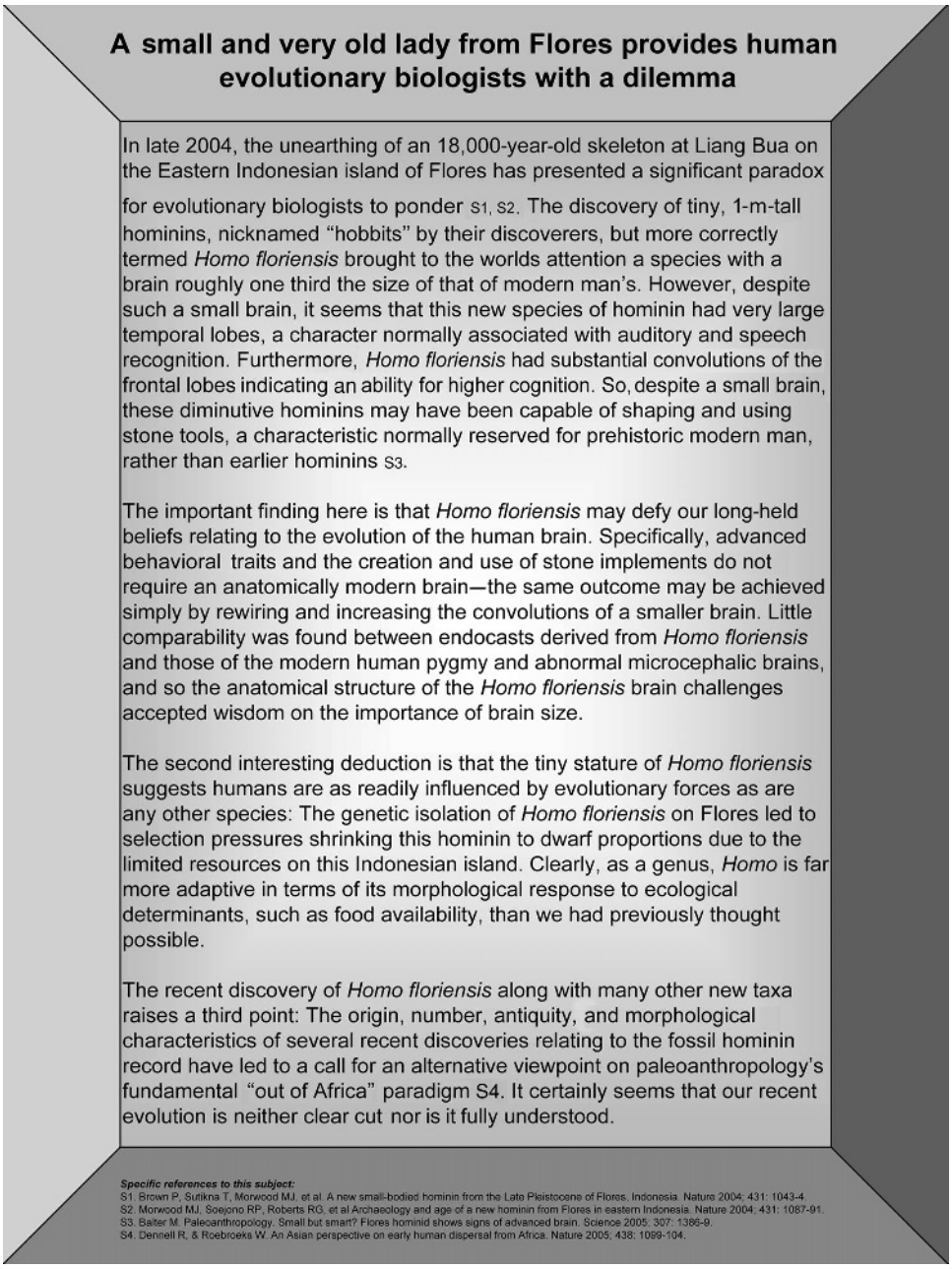

Figure 1.8. *The 2004 discovery of* Homo floriensis *on the Indonesian island of Flores challenges our perceived wisdom relating to man's recent evolutionary past.*

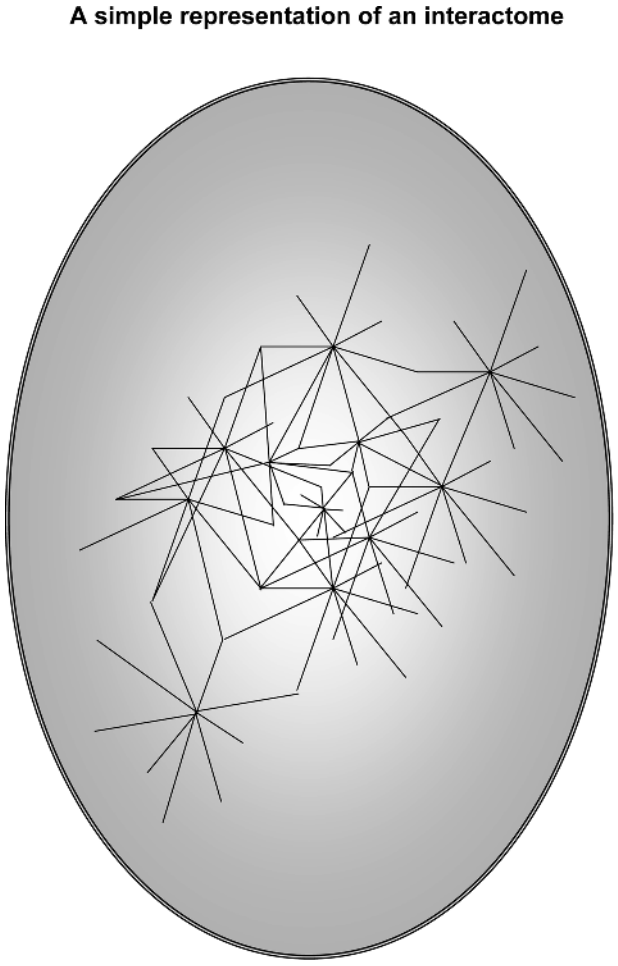

Figure 1.9. *An extremely simple rendering of the interactome; unfortunately, reality is infinitely more complex than can be represented here. Imagine each node as a cluster of proteins with similar cellular function. Each cluster is then linked by an interactive network. As an example, three juxtaposed nodes in the above figure might represent DNA synthesis proteins, DNA repair proteins, and cell-cycle regulatory proteins. Then consider a hypothetical node for proteins involved in protein folding; these are likely to be located at a more distant nexus as they are not closely involved with the former three protein clusters. Now imagine how complex an interactome for humans would be if each protein represented a single node!*

nutritional pharmacology in the context of genetic polymorphisms and clinical experience." These are sound definitions, and worthy of reiteration for all students of the subject.

As our knowledge of the "nutriome" improves and the gaps within the interactome are filled in, it seems likely that the buzzwords of today like nutrigenomics and nutritional genetics will ultimately give way to the unifying field of human molecular nutrition.