THE ORIGINS, PATTERNS AND IMPLICATIONS OF HUMAN SPONTANEOUS MUTATION

James F. Crow

The germline mutation rate in human males, especially older males, is generally much higher than in females, mainly because in males there are many more germ-cell divisions. However, there are some exceptions and many variations. Base substitutions, insertion–deletions, repeat expansions and chromosomal changes each follow different rules. Evidence from evolutionary sequence data indicates that the overall rate of deleterious mutation may be high enough to have a large effect on human well-being. But there are ways in which the impact of deleterious mutations can be mitigated.

AUTOSOME
A chromosome other than the X or Y.

Genetics Department, University of Wisconsin, Madison, Wisconsin 53706, USA
e-mail: jfcrow@facstaff.wisc.edu
http://www.wisc.edu/genetics/CATG/crow/index.html

> "If a more exact analysis of birth order were indeed to confirm a high incidence in last-born children, this would speak for the formation of the initial predisposition for dwarfism by mutation."
>
> Wilhelm Weinberg

The remarkable statement above was published by Weinberg in 1912 (FIG. 1)[1]. While studying achondroplasia (dominantly inherited, short-limbed dwarfism), he noticed that sporadic cases were most often found among the last-born children of a sibship and inferred a possible mutational origin. Mutation was a vague concept in those days and the correctness of Weinberg's interpretation is only one example of his early insight into human genetics[2]. Weinberg made no distinction between paternal age, maternal age and birth order as possible causes of this bias, but 40 years later Penrose[3] showed that paternal age was the main, if not the sole, cause of Weinberg's observation.

Achondroplasia is only one of several traits for which sporadic cases show a paternal age effect. In 1987, Risch *et al.* reported a strong paternal age effect for twelve syndromes, including Apert, basal cell nevus, Crouzon, Marfan, progeria and Waardenburg[4]. Others have been found since, and several more show a mild increase with paternal age. This widespread phenomenon indicates that mutations may occur disproportionately in males.

The purpose of this review is to examine the evidence for sex- and age-specific patterns in human spontaneous mutations. I also review our current understanding of the absolute rate of deleterious mutation and consider the implications of these findings for human health and welfare.

Male and female mutation rates

In early genetic studies it was not possible to determine whether a new AUTOSOMAL mutant gene in an affected child came from the mother or the father. The inheritance of the X chromosome, however, offered the possibility of making this distinction. The first person to take advantage of this was Haldane[5]. He noted that males with X-linked haemophilia came much more often from heterozygous, carrier mothers than from homozygous, normal mothers, showing that the mutation had occurred in an earlier generation. He inferred that the male mutation rate is roughly ten times that of the female, although his estimates were uncertain, first because of heterogeneity (the different types of haemophilia were not distinguished) and, second, because of uncertainty in heterozygote detection. Nevertheless, Haldane's conclusion that the male mutation rate is much higher than the female rate has stood the test of time for both haemophilia A[6] and haemophilia B[7,8]. The general conclusion is supported

by data on other X-linked traits, including Lesch–Nyhan syndrome, a severe defect in purine metabolism[9], and ornithine transcarbamylase (OTC) deficiency[10].

There are 13 dominant X-linked diseases that are lethal or sterilizing in females. Surprisingly, there is a nearly complete absence of affected males. The traditional explanation has been that in males the trait is so severe that it causes prenatal death, and there is evidence for this in some cases. But it seems unlikely that all 13 diseases have this property. In contrast, Thomas[11] has pointed out that this bias is what would be expected if the male mutation rate were higher than the female rate. In this case, affected males would almost always come from heterozygous mothers. But if affected females do not reproduce there would be no affected sons. This is a plausible alternative to the gestational death hypothesis and, if correct in whole or in part, supports the idea that the high male:female mutation rate ratio is a general phenomenon.

With the advent of molecular markers in a densely mapped genome it is now often possible to distinguish between maternal and paternal origin of mutations by examining markers linked to the genes of interest. For Weinberg's classic trait, achondroplasia, this technique shows that essentially all mutations occur in males. Wilkin *et al*. report 40 sporadic cases, of which all the mutations were paternal[12]. In 57 cases of Apert syndrome (acrocephalosyndactyly)[13], 25 of multiple endocrine neoplasia 2B (MEN 2B)[14], ten of MEN 2A (REF. 15), and 22 of Crouzon and Pfeiffer syndromes[16], all new mutations were paternal. In these six conditions, which involve three different genes — *FGFR3*, *FGFR2* and *RET* — 154 new mutations have been analysed and all have a paternal origin.

The mutations just discussed are single base substitutions. The most striking is achondroplasia, in which 153 of 154 analysed cases are due to a glycine to arginine substitution at codon 1,138. The mutations are in the transmembrane domain of the fibroblast growth factor receptor 3 (FGFR3). Of the 153 mutations, 150 were guanine to adenine transitions and three were guanine to cytosine transversions of the same nucleotide[17]. This means that all the cases of achondroplasia are due to changes in one nucleotide — a nucleotide with the highest known mutation rate (about 10^{-5} per generation). There are mutations at other sites in this gene, but the phenotypes are different[18].

Box 1 | Estimating the number of male germ-cell divisions

We can estimate the number of germ-cell divisions in a male of age A as follows. There are an estimated 30 cell divisions before puberty and then one stem cell division every 16 days, or 23 per year. Then, before sperm formation there are four mitotic and two meiotic divisions (one chromosome replication). Letting N_A be the number of germline chromosome replications at age A, N_p the number at puberty and A_p the age at puberty, taken to be 15 years,
$N_A = N_p + 23(A - A_p) + 4 + 1 = 35 + 23(A - 15)$.
This calculation gives the following results.

Age	Chromosome replications
15	35
20	150
30	380
40	610
50	840

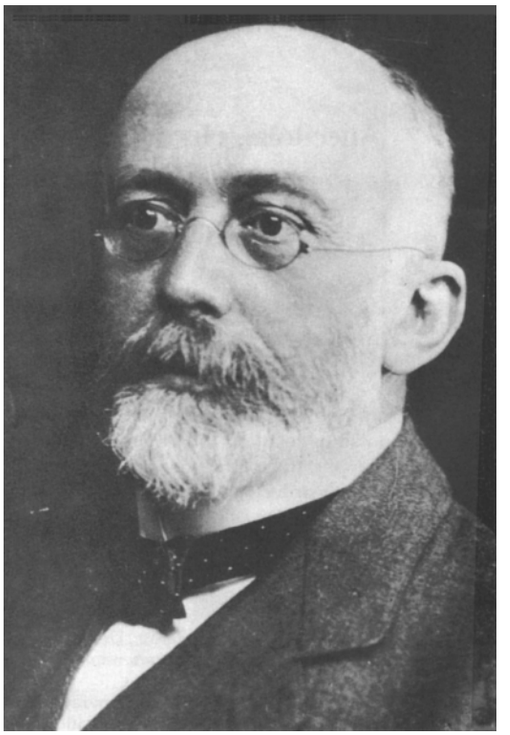


Figure 1 | **Wilhelm Weinberg.** Photograph taken from REF. 56. © Genetics Society of America.

The achondroplasia mutations are all at a CpG nucleotide pair, known to be a mutation hot-spot. In fact it is generally true that a disproportionate number of base substitutions occur at such sites. However, CpG mutability is not responsible for the higher male mutation rate at other loci, as this also occurs at non-CpG sites. Other studies have discovered a few maternally derived mutations[19], but male-derived mutations greatly predominate. Mutations at *FGFR2*, *FGFR3* and *RET* represent the most extreme examples of male mutation bias, and I shall return to them later, but first, how can we explain such a striking sex difference?

Germ-cell divisions in males and females

One marked difference between the human male and female is that there are many more germline cell divisions in the life history of a sperm relative to that of an egg. Furthermore, the difference increases with the age at which the sperm is produced. This suggests a possible explanation for the sex difference and the paternal age effect.

The number of cell divisions preceding the production of the mature ovum or sperm has been calculated by Vogel and Rathenberg[20], and Drost and Lee[21]. FIGURE 2 illustrates this process[22]. In the female there are 22 cell divisions before meiosis and two during meiosis, giving 23 chromosome replications in total, because only one replication occurs during the two meiotic divisions. As all the cell divisions are completed before birth, there is no increase with postnatal age.

Spermatogenesis is quite different. Sperm are produced continuously throughout reproductive life, so the

TRISOMY
Having three copies of a chromosome.

ANEUPLOID
Having an unbalanced chromosome number. An example is trisomy.

number of cell divisions and chromosome replications that have occurred increases with age. As shown in BOX 1 (previous page), a sperm produced by a man of age 40 has gone through 610/23, or more than 25 times as many chromosome replications as an egg. Conversely, for a man of age 20, this number is only about 7 times as many. The ratio of germ-cell divisions between males and females is high, but it is not sufficient to account for the high male:female mutation ratios observed for the two *FGFR* genes and *RET*.

The acceleration in mutation rate with paternal age is borne out by the curves relating disease incidence to paternal age. FIGURE 3 shows data for achondroplasia and Apert syndrome. The relationship is clearly nonlinear and rises sharply as age increases, as is true for some other dominant mutations[4]. An accelerating rate with age could be the result of reduced fidelity of DNA replication and inefficiency of repair mechanisms, which are expected to deteriorate with age. Another possibility is cell death at old ages in the germ line, compensated for by an increased number of cell divisions. Still another factor might be the accumulation of mutagens, from either external or internal sources, the effects of which should certainly increase with age, although not necessarily in a non-linear manner.

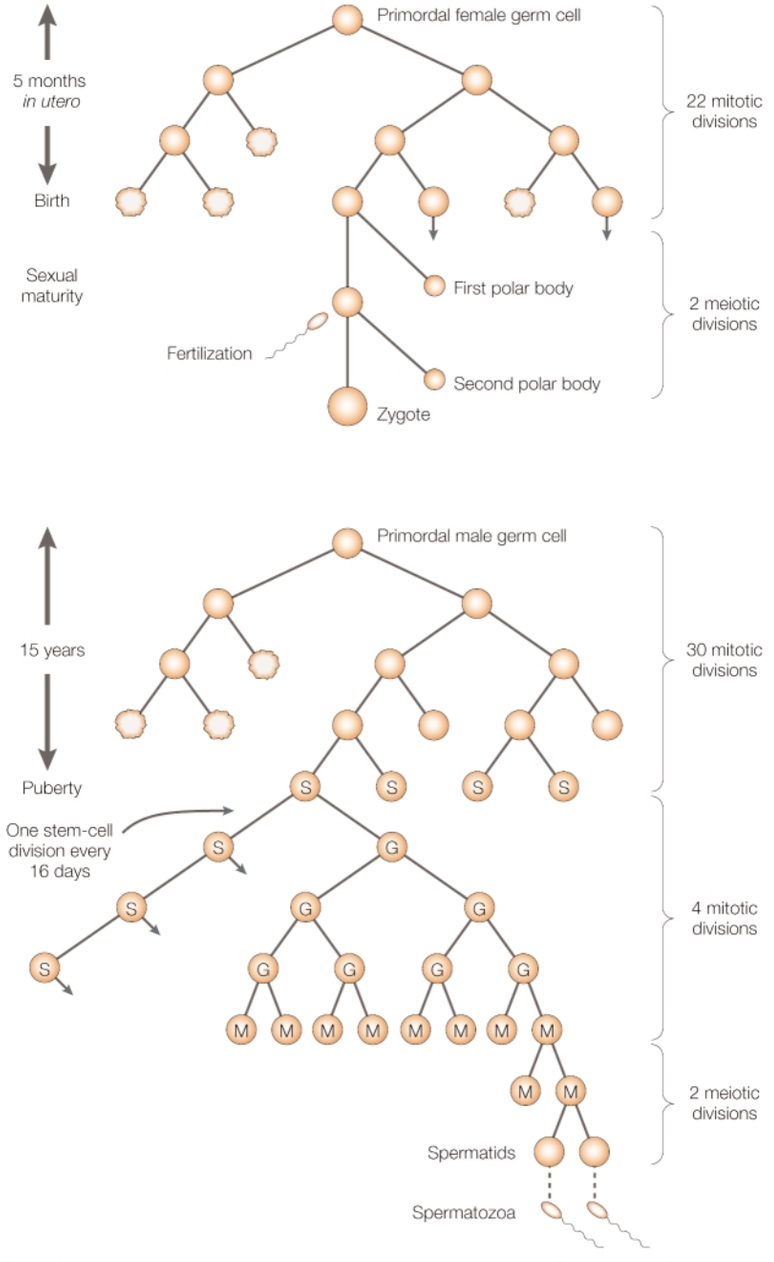


Figure 2 | **Cell divisions during oogenesis and spermatogenesis.** S, stem cells; G, gonial cells; M, meiotic cells. The total number of cell divisions in the life history of an egg is 24. In males this depends on the number of stem-cell divisions, which is greater in older males. (Figure adapted from REF. 22.)

Exceptions to the paternal age effect

There are some exceptions — single gene traits that show only a slight paternal bias in the male:female mutation ratio and a small paternal age effect. Two such traits are neurofibromatosis type 1, an autosomal dominant condition (also called Von Recklinghausen disease) (FIG. 4), and Duchenne muscular dystrophy, an X-linked recessive trait.

The genes for both of these diseases are enormous. That for Duchenne muscular dystrophy has 55 exons[23] and that for neurofibromatosis has 59 (REF. 24). The neurofibromatosis gene has one of the highest mutation rates — 10^{-4} per gene per generation[25] — and the Duchenne rate is similar. Many of the mutations are intragenic deletions, which are more common in larger genes. This is not because deletions are produced less often in smaller genes, but rather because deletions removing a part of these large genes could, in smaller genes, delete the whole gene and more. This could lead to early lethality, which would not be detected.

The following hypothesis arises: whereas base substitutions occur primarily in males and are age-dependent, small chromosomal changes (mainly intragenic deletions) are not age-dependent because they occur by different mechanisms. The data indicate that the rate of occurrence of deletions is actually higher in females than in males. For Duchenne muscular dystrophy, 93% of point mutations were from sperm, whereas 87% of deletions were maternal[23]. The data from neurofibromatosis are consistent with this view, although the numbers are small — 16 out of 21 deletions were maternal, whereas 9 out of 11 point mutations were paternal[24]. Therefore, only a small paternal age effect is expected for diseases like neurofibromatosis, because only part of the mutations are base substitutions (FIG. 4). Retinoblastoma[26] and Wilms tumour[27] are two further examples of diseases where only a small paternal age effect is observed, because a significant fraction of the new mutations are not base substitutions.

The best known example of a maternal age effect occurs in Down syndrome. It has long been known that transmission of an extra chromosome leading to TRISOMY is much more common from females than from males[28]. About 0.3% of liveborns are ANEUPLOID, with the most common being trisomy 21 leading to Down syndrome. The trisomy rate at conception is, of course,

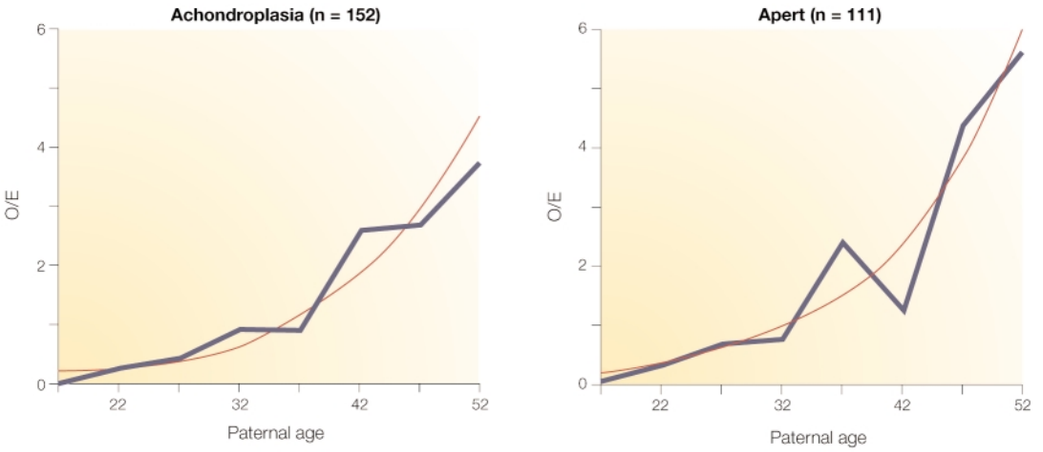


Figure 3 | **Relative frequency of *de novo* achondroplasia and Apert syndrome for different paternal ages.** The ordinate is the ratio of the observed number of mutations (O) to the number expected (E), if all paternal ages are associated with the same frequency of mutation. The blue line gives the actual data; the red line is the best-fitting exponential curve. (Figure adapted from REF. 4.)

much higher (at least 10%), with most dying very early in the prenatal period. For trisomy 21, 93% of 436 informative cases were of maternal origin. For the other trisomies, the numbers were similar, ranging from 81 to 100%. The exception is XXY, which occurs at roughly the same rate in males and females[28].

The maternal age effect is striking, even more so than the paternal age effect for base substitutions (FIG. 5). In contrast to the paternal age effect, where an obvious explanation is the excess of male cell divisions, there is no such explanation of the extreme maternal age effect for trisomy. As there is no increase in the number of premeiotic cell divisions with maternal age, perhaps the length of time for which the chromosomes are 'suspended' in meiosis is somehow responsible.

Heterogeneity in the sex effect

The three loci discussed above (*FGFR2*, *FGFR3* and *RET*) stand out. Essentially all the mutations occur in males and the paternal age dependence is strong. Other traits are less extreme. Reported ratios for the male:female base-substitution rate for X-linked traits are mostly uncertain because of small sample numbers: about 50 for OTC[10], 10 for Lesch–Nyhan[9], 5–10 for haemophilia A[6], 4–9 for haemophilia B[7,8] and 40 for Duchenne muscular dystrophy[23]. Most of these are underestimates as the data include some deletions. With the use of the base-substitution fraction of other, autosomal genes, two values for the male:female ratio are 4.5 (neurofibromatosis)[25] and 6 (retinoblastoma)[29]. The inconsistencies are not surprising in view of ascertainment problems and the statistical uncertainties of small numbers. Mosaicism might be another factor diluting the paternal age effect.

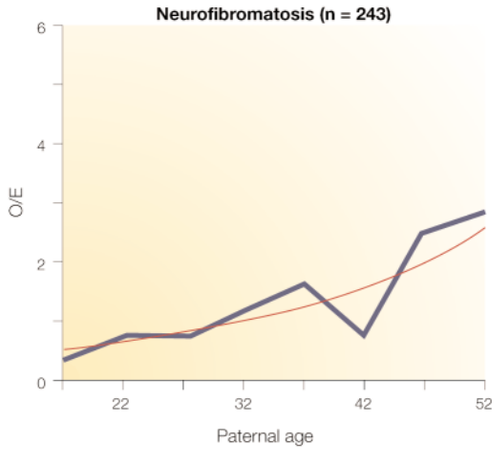


Figure 4 | **Relative frequency of *de novo* neurofibromatosis for different paternal ages.** The ordinate is the ratio of the observed number of mutations (O) to the number expected (E), if all paternal ages are associated with the same frequency of mutation. The blue line gives the actual data; the red line is the best-fitting exponential curve. (Figure adapted from REF. 4.)

The geometric mean of these ratios is about ten — probably an underestimate. Nevertheless, the paternal mutation bias seen in most genes is clearly less than that seen in *FGFR2*, *FGFR3* and *RET*, the three loci in which virtually all the mutations originate in males. The mutations in these three genes are all gain-of-function mutations affecting proteins belonging to a single protein family. It is not expected that the phenotype of the mutations would affect their rate of occurrence, but it might affect their frequency of recovery. Along these lines, it has been suggested that mutations at the three loci somehow confer a selective advantage during spermatogenesis[13,30], but there is no direct evidence for this. Alternatively, if we are to invoke germline selection, there could be selection against the mutant alleles during oogenesis. Evidence for germline selection might be obtained from segregation ratios in matings involving affected people.

From an evolutionary standpoint, an increase in mutation rate at later reproductive ages is not surprising. In our remote ancestry, probably very few males lived to reproduce in their 40s. So, there would be very little selection pressure to reduce the harmful mutation rates late in the reproductive period. The luxury of reproducing at older ages is a bonus that contemporary men receive from higher living standards, medical advances and other environmental improvements.

The causes of the discrepancy between the three extreme loci and the others are unknown and will have to await the results of future research. It is clear, even though the magnitude is in doubt, that the base substitution rate is much higher in males than in females and that the difference increases with paternal (or grandpaternal) age. This supports the view that base substitutions are associated with DNA replication in mitotic cells.

Complex traits

The traits discussed so far are all caused by single gene changes. The reason for studying simply inherited traits is of course that they are easier to observe. So it is not surprising that these have dominated the literature. There is every reason to believe that the patterns in the origins of spontaneous mutations can be extended to more COMPLEX TRAITS. In particular, there is no reason to think that the mutation rate should depend on the magnitude of the phenotypic effect.

Olshan *et al.* reported a slight paternal age effect for congenital heart defects, including ventricular and atrial septal defects, and patent ductus[31]. The relative risk increased by two- to threefold at a paternal age of 50 (or older) relative to age 25–29. Any paternal age effect, not accompanied by an equal maternal age effect, shouts 'mutation'. Therefore, it seems reasonable that a small fraction of these heart defects is due to one or more paternal gene mutations, among the undoubtedly complex array of genetic and environmental causes.

A comparable paternal age effect is reported for prostate cancer[32]. For four paternal age groups (<27, 27–32, 32–38 and >38) the relative incidences are 1.0, 1.2, 1.3 and 1.7, respectively. The age trend is significant (P = 0.049) and there is no maternal age effect, after adjusting for paternal age. Likewise, there is a slight paternal age effect for the CHARGE syndrome, a combination of congenital defects[33]. Some forms of cerebral palsy also show a possible paternal age effect[34] as do non-hereditary Alzheimer disease[35] and schizophrenia (S. Harlap, E. Susser & D. Malaspina, personal communication).

These observations suggest a research strategy — to identify children with such defects whose fathers are exceptionally old. This should provide a greatly enriched population for discovering new mutations. Then a search for gene differences or gene-expression differences might lead to much-sought-after genes for multifactorial traits. I think it likely that inclusion of the age of the father will become more common in future epidemiological studies and may provide leads to causal factors.

Paternal excess in other mutations

Mutation types other than base substitutions have also been identified in which a paternal increase is seen, but whether they will be as numerous as single base changes is doubtful. Haemophilia provides one example[36]. Around 40–45% of severe haemophilia (factor VIII deficiency) cases are caused by an X-chromosome inversion in the factor VIII region that arises almost exclusively in the male. Essentially all the affected males come from carrier mothers and, of 70 mutations in which the source could be identified, 69 were from the maternal grandfather. Despite this extreme sex bias, there was no grandpaternal age effect, so this effect clearly has a different cause from that producing point mutations. The suggested mechanism involves pairing between the factor VIII locus at intron 22 and nearby repeated sequences, leading to an interruption of the gene. This probably happens during meiosis, which accounts for the absence of an age effect. Presumably the reason this occurs only in males is that the male X chromosome does not have a synaptic partner, so intrachromosomal mispairing occurs.

This inversion accounts for a large fraction of severe haemophilia cases. When mild cases are included the results are quite different[6]. Among 126 chromosomes analysed , 37% showed the factor VIII inversion, 32% were point mutations, 9% small deletions, 5% large deletions and 1% small insertions (the others could not be analysed). Other than the factor VIII inversion, these followed the familiar age and sex pattern; point mutations had a large male excess, whereas deletions showed a small female excess.

Some gross chromosomal changes show a paternal effect. Partial loss of the long arm of chromosome 18

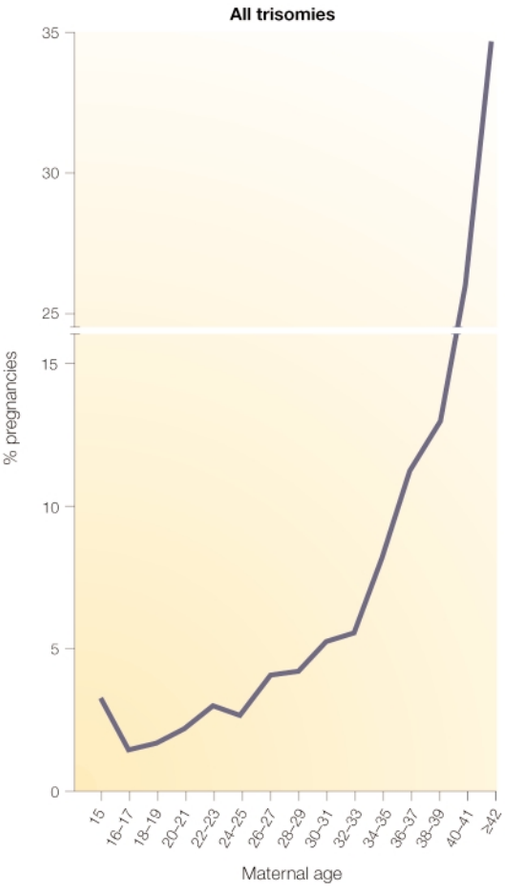


Figure 5 | **Relative frequency of all trisomies for different maternal ages.** The ordinate is the percentage of trisomy occurring among recognized pregnancies. (Figure adapted from REF. 28.)

COMPLEX TRAIT
A trait determined by many genes, almost always interacting with environmental influences.

NONSYNONYMOUS
A nucleotide change that alters the coded amino acid.

NEUTRAL MUTATION
A mutation that is selectively equivalent to the allele from which it arose.

INDEL
Insertion or deletion in a chromosome.

FITNESS
A measure of the capacity to survive and reproduce.

GENETIC DEATH
A pre-reproductive death or failure to reproduce.

QUASI-TRUNCATION SELECTION
Approximate or inexact truncation selection — selection in which all individuals below a certain threshold survive and reproduce equally; the others are eliminated.

occurs disproportionately in the male (29 out of 34 cases)[37]. The deletion in the short arm of chromosome 4 that causes Wolf–Hirschhorn syndrome is similar (24 out of 29 cases)[38], as is the deletion of the short arm of chromosome 5 that causes the cri-du-chat syndrome (20 out of 25 cases)[39]. These large deletions follow a set of rules different from the small intrachromosome deletions responsible for some cases of neurofibromatosis and Duchenne muscular dystrophy.

Still another type of mutation with a sex bias is expansion in the number of units in repeated sequences. In particular, the expansion of trinucleotide repeats is the cause of about 20 diseases, including fragile X syndrome and myotonic dystrophy. One well-characterized example is Huntington disease in which the sequence CAG is repeated a variable number of times in the *huntingtin* gene[40]. The severity of the disease is strongly correlated with the number of repeats, 11–34 being normal and 37–84 producing the disease. Paternally derived loci have more repeats than maternally derived ones. But different trinucleotide-repeat diseases behave differently — in some, most changes are in females.

Perhaps the best place for an overall assessment of parental effects is oligonucleotide repeats that do not cause a disease. These seem to show a three- to fivefold paternal excess[41–43], not as much as would be expected on the basis of a simple relationship with the number of male cell divisions.

Absolute mutation rates

The mutational spectrum covers a wide range of effects, but mutations with severe, conspicuous effects have been studied preferentially. These are not necessarily the most important — simply the easiest to study. Therefore, it has not been possible to extrapolate from these specific examples to the entire genome, which is essential if we are to understand the effects of deleterious mutation on human welfare.

Single base changes constitute the most frequent class of mutations. The effects of these mutations range from conspicuous, highly deleterious phenotypes, to those with very mild effects or none. Beneficial mutations also occur, but these have been difficult to study quantitatively. At present the best estimate of genomic rates comes from studies of molecular evolution.

Eyre-Walker and Keightley[44] measured amino-acid changes in 46 proteins in the human ancestral line, since its divergence from the chimpanzee. Among 41,471 amino acids, they found 143 NONSYNONYMOUS substitutions. Because NEUTRAL MUTATIONS accumulate at a rate equal to the mutation rate[45], they used non coding data to calculate that 231 neutral mutations would have been expected. Deleterious mutations occurred during this time interval, but they have left no record. So the difference between the expected number of neutral mutations and the observed number of mutations that have persisted is the number that were lost. These are the deleterious mutations, and their number is estimated to be 231 – 143 = 88. In other words, 88/231 or about 38% of the mutations have been eliminated by natural selection. The estimated rate of deleterious mutation is 88/41,471 or 0.00212 per amino acid per six million years, the estimated time of divergence between chimpanzees and humans. The genes studied have an average of 1,523 amino-acid codons and the estimated number of genes per diploid genome was taken as 120,000, giving 387,451 deleterious mutations per diploid genome per six million years. Using 25 years as the average age of reproduction gives 1.6 ± 0.8 deleterious mutations per diploid genome per generation.

An independent estimate of the overall mutation rate on the basis of the factor IX locus gives 1.3 per zygote[46]. But there are many reasons to question the accuracy of the estimates. Eyre-Walker and Keightley[44] suggest from other comparisons that their estimate of the fraction of deleterious mutations, 38%, is too low, and an independent estimate of the number of mutations now being eliminated is about 60% (REF. 47). There is also uncertainty about the time since our split from the chimpanzee and the number of genes in the genome[48]. Finally, mutations outside the coding region and INDELS were not included.

Eyre-Walker and Keightley propose the best estimate for the deleterious mutation rate as three new deleterious mutations per zygote. More research is urgently needed to test and extend these important conclusions.

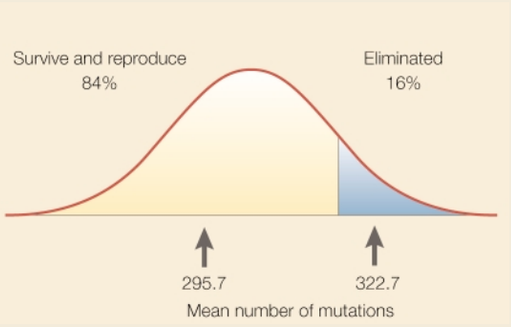
Assessing the impact of a high mutation rate

As first shown by Haldane[49], mildly deleterious mutations, because they persist in the population much longer, must in the long run cause a reduction of FITNESS as large as those with a marked effect. In a population of stable size, each deleterious mutation must ultimately be extinguished by what Muller called a GENETIC DEATH[50]. With three deleterious mutations per generation, why are we not dead three times over? More likely, fitness is multiplicative rather than additive, so the probability of surviving and reproducing is e^{-3}, or about 0.05. Even if the mutation rate were one rather than three, fitness would be reduced by 63%. With such low fitness, why has the human species not become extinct?

There are several answers to this pessimistic question. One is that the estimate of three new deleterious mutations per generation may be wrong. The uncertainty is great and the true value could be considerably less. Second, the human population is almost certainly not at equilibrium, because environmental improvements in much of the world have changed the parameters on which the equilibrium value depends. Third, there is room for considerable early prenatal selection without a substantial social cost. But this seems insufficient to balance a mutation rate higher than about one per generation.

I think the answer, or part of it, lies in the efficiency of sexual reproduction in eliminating deleterious genes from the population. In BOX 2, I present a simplified model, showing how QUASI-TRUNCATION SELECTION can be an effective means of eliminating deleterious mutations. The idea is that most death and failure to repro-

Box 2 | Removal of deleterious mutations by truncation selection



A mildly deleterious mutation can persist in the population for many generations, the number being related to the reduction in fitness caused by the mutation. We have very little idea of what the average persistence of mildly deleterious human mutations is. In *Drosophila*, it is estimated at 50 to several hundred generations. As an example, I will take 100. Three new mutations per generation, each persisting for 100 generations, means that the average person carries 300 mutations. If these are independently inherited, the number per person will have a distribution that is roughly POISSON (actually with a little less spread because of incomplete randomization during meiosis)[54]. I shall assume a standard deviation of 15 (the Poisson value is the square root of the mean, or about 17), that the distribution is normal (entirely reasonable with this many mutations) and that those with a number above one standard deviation from the mean, about 16% of zygotes, fail to survive and reproduce. In the next generation new mutations occur and existing ones are shuffled by recombination so that the original normal distribution is restored.

This is illustrated in the figure. The mean number of mutations per person is 300. With 16% eliminated, the mean number among these is 322.7. The mean number in the 84% not eliminated is 295.7. Three new mutations are not quite enough to bring the number back to 300, so 16% selective elimination is more than enough to balance mutation accumulation. Truncation selection is indeed an efficient method for eliminating harmful mutations.

Of course, nobody thinks that this is a realistic model — nature does not truncate precisely. For this reason, I was once reluctant to regard this as a reasonable possibility. Then I was surprised to find[55] that a very fuzzy approximation to truncation selection (quasi-truncation selection) works nearly as well. All that is required is that the probability of removal increases monotonically with the number of deleterious mutations. Until recently in our evolutionary past, the population was nearly stable and every generation produced more progeny than could survive and reproduce. To some extent those who failed would have been those with the largest number of deleterious mutations. So, I think it quite likely that in the past such quasi-truncation selection has been an efficient means of eliminating deleterious mutations. Without belabouring the specific assumptions, it seems reasonable that elimination of harmful mutants was far more efficient than would have been expected if they were eliminated independently as proposed by Haldane[49].

duce occurs among those with the largest number of mutations. Fertility differences among males were probably important in our remote ancestry.

As explained in BOX 2, this kind of selection requires sexual reproduction. This suggests that a reason that we have been able to evolve a longer life cycle — with greater opportunity to learn — is that sexual reproduction mitigates the effect of the higher mutation rate that would tend to accompany an increase in generation length. So there are ways to explain how we have survived this long and continue to thrive. But what about the effects of relaxed selection in the recent past in those countries with a high living standard?

Effects of relaxed selection

Our high standard of living, improved sanitation and better medical care mean that a number of mutant genes that would have been selectively eliminated in the past are now perpetuated. Whatever equilibrium may have been reached in the past no longer exists. We are certainly accumulating mutations faster than they are being eliminated. Furthermore, the possibility of new kinds of mutagens, external and internal, may increase the imbalance. There is every reason to think that the bulk of mutations, if not neutral, are deleterious. There is also reason, as I have said, for thinking that mild mutations are disproportionately frequent, and that the disproportion increases as they approach neutrality. How serious is the problem, and how soon will it become important?

Most of the mutations have a very small effect and are to a large extent compensated for by environmental improvements. Who worries about having to wear spectacles? But can we continue to improve the environment indefinitely? Will a time come, especially if there is some sort of catatastrophe (war, epidemic or famine), when we are forced to return to the life of our ancestors? Under those circumstances we would surely see an increase in human misery, for all the mutations that have accumulated would be expressed in full force.

But there are grounds for optimism. The brave new world of molecular genetics will provide ways of detecting and eliminating important mutant genes with little human or social cost. As for genes with very minor effects, the accumulation rate is very slow, while environmental improvement is rapid. I am thinking of dozens or more generations, far longer than we are able to foresee what kind of environment we shall have. Mutation accumulation is a process that may or may not ultimately be important, but one thing is certain: the time scale is very long. We have time to learn more.

Future directions

In earlier reviews, I have regarded the *RET*, *FGFR2* and *FGFR3* loci, for which the data are clearest, as typical[51,52]. But in the absence of additional examples, there is the possibility that these are exceptional. We know that there is a greater number of base substitutions in males, but whether these are more frequent than can be accounted for by the number of stem-cell divisions remains to be determined. The role of DNA methylation and CpG sites also needs clarification. One place to look for mechanisms is in somatic cells, which are more amenable to study, and there are already some data on mutation and age[53].

A rich and exciting source of data on male mutation rates will be provided by direct study of spermatozoa (or spermatids). It is now possible to detect small chromosome changes, and it should soon be possible to detect single base changes on a scale that makes quantitative mutation study feasible. The paternal age effect can then be determined precisely, and any uncertainties about whether the diseases I have discussed are a representative sample can be removed.

I fully expect many more studies of paternal age

POISSON
A statistical distribution in which the probability of an individual event is small, but the number of opportunities is large enough that several occur.

Links

DATABASE LINKS acondroplasia | Apert syndrome | basal cell nevus | CHARGE syndrome | cri-du-chat syndrome | Crouzon syndrome | Down syndrome | Duchenne muscular dystrophy | factor VIII locus | factor IX locus | FGFR3 | FGFR2 | fragile X syndrome | haemophilia A | haemophilia B | Huntington disease | Lesch–Nyhan syndrome | Marfan syndrome | MEN2B | MEN2A | myotonic dystrophy | neurofibromatosis type I | OTC deficiency | Pfeiffer syndrome | Progeria syndrome | retinoblastoma | RET | Waardenburg syndrome | Wilm tumour | Wolf–Hirschhorn syndrome

effects in complex diseases. As emphasized earlier, those families with exceptionally old fathers are a promising source for discovery of mutant genes affecting a disease.

Finally, this review has discussed a wide variety of mutation mechanisms. Sorting out the relative importance of these mechanisms and discovering new ones guarantees excitement and progress in the years ahead.

1. Weinberg, W. Zur Vererbung des Zwergwuchses. *Arch. Rassen- u. Gesel. Biolog.* **9,** 710–718 (1912).
In this paper, Weinberg first noted the importance of birth order, which led to the idea of the paternal age effect.
2. Crow, J. F. Hardy, Weinberg and language impediments. *Genetics* **152**, 821–825 (1999).
3. Penrose, L. S. Parental age and mutation. *Lancet* **269**, 312–313 (1955).
4. Risch, N., Reigh, E. W., Wishnick, M. W. & McCarthy, J. G. Spontaneous mutation and parental age in humans. *Am. J. Hum. Genet.* **41**, 218–248 (1987).
A detailed review and analysis of the classical literature on the paternal age effect.
5. Haldane, J. B. S. The mutation rate of the gene for hemophilia and its segregation ratios in males and females. *Ann. Eugen.* **13**, 262–271 (1947).
The first measurement of a human mutation rate, as well as evidence for a higher male rate.
6. Becker, J. *et al.* Characterization of the factor VIII defect in 147 patients with sporadic hemophila A: Family studies indicate a mutation type-dependent sex ratio of mutation frequencies. *Am. J. Hum. Genet.* **58**, 657–670 (1996).
The different kinds of mutation events leading to haemophilia and their relative frequency.
7. Ketterling, R. P. *et al.* Germline origins in the human F9 gene: frequent G:C→A:T mosaicism and increased mutations with advanced maternal age. *Hum. Genet.* **105**, 629–640 (1999).
8. Green, P. M. *et al.* Mutation rates in humans. I. Overall and sex-specific rates obtained from a population study of hemophilia B. *Am. J. Hum. Genet.* **65**, 1572–1579 (1999).
9. Francke, U. *et al.* The occurrence of new mutants in the X-linked recessive Lesch–Nyhan disease. *Am. J. Hum. Genet.* **28**, 123–137 (1976).
10. Tuchman, M. *et al.* Poportions of spontaneous mutations in males and females with ornithine transcarbamylase deficiency. *Am. J. Med. Genet.* **55,** 67–70 (1995).
11. Thomas, G. H. High male:female ratio of germ-line mutations: An alternative explanation for postulated gestational lethality in males in X-linked dominant disorders. *Am. J. Hum. Genet.* **58**, 1364–1368 (1996).
12. Wilkin, D. J. *et al.* Mutations in fibroblast growth-factor receptor 3 in sporadic cases of achondroplasia occur exclusively on the paternally derived chromosome. *Am. J. Hum. Genet.* **63**, 711–716 (1998).
A demonstration that achondroplasia mutations all occur in the father and at a particular CpG hot-spot.
13. Moloney, D. M. *et al.* Exclusive paternal origin of new mutations in Apert syndrome. *Nature Genet.* **13**, 48–53 (1996).
14. Carlson, K. M. *et al.* Parent-of-origin effects in multiple endocrine neoplasia type 2B. *Am. J. Hum. Genet.* **55**, 1076–1082 (1994).
15. Schuffenecker, I. *et al.* Prevalence and parental origin of *de novo* RET mutations in multiple endocrine neoplasia type 2A and familial medullary thyroid carcinoma. *Am. J. Hum. Genet.* **60**, 233–237 (1997).
16. Glaser, R. L. *et al.* Paternal origin of FGFR2 mutations in sporadic cases of Crouzon Syndrome and Pfeiffer Syndrome. *Am. J. Hum. Genet.* **66**, 768–777 (2000).
17. Bellus, G. A. *et al.* Achondroplasia is defined by recurrent G380R mutations of FGFR3. *Am. J. Hum. Genet.* **56**, 368–373 (1995).
18. Vajo, Z., Francomano, C. A. & Wilkin, D. J. The molecular and genetic basis of fibroblast growth factor receptor 3 disorders: the achondroplasia family of skeletal dyplasias, Muenke craniosynostosis, and Crouzon syndrome with acanthosis nigricans. *Endocr. Rev.* **21**, 23–39 (2000).
This paper summarizes the molecular events at the three loci with the highest male:female mutation rate ratio.
19. Kitamura, Y. *et al.* Maternally derived missense mutations in the tyrosine kinase domain of the ret protooncogene in a patient with de-novo men-2b. *Hum. Mol. Genet.* **4**, 1987–1988 (1995).
20. Vogel, F. & Rathenberg, R. Spontaneous mutation in man. *Adv. Hum. Genet.* **5**, 223–318 (1975).
21. Drost, J. B. & Lee, W. R. Biological basis of germline mutation: Comparisons of spontaneous germline mutation rates among *Drosophila*, mouse, and human. *Env. Mol. Mut.* **25** (S26), 48–64 (1995).
22. Vogel, F. & Motulsky, A. G. *Human Genetics; Problems and Approaches.* (Springer, Berlin,1997).
An unusually complete textbook of human genetics, with detail on mutation.
23. Grimm, T. G. *et al.* On the origin of deletetions and point mutations in Duchenne muscular dystrophy: most deletions arise in oogenesis and most point mutations result from events in spermatogenesis. *J. Med. Gen.* **31**, 183–186 (1994).
24. Lazaro, C. *et al.* Sex differences in the mutation rate and mutational mechanism in the NF gene in neurofibromatosis type 1 patients. *Hum. Genet.* **98**, 696–699 (1996).
25. Fahsold, R. *et al.* Minor lesion mutational spectrum of the entire NF1 gene does not explain its high mutability but points to a functional domain upstream of the GAP-related domain. *Am. J. Hum. Genet.* **66**, 790–818 (2000).
26. Lohmann, D. R., Brandt, B., Hopping, W., Passarge, E. & Horstemke, B. The spectrum of RB1germ-line mutations in hereditary retinoblastoma. *Am. J. Hum. Genet.* **58**, 940–949 (1996).
27. Olson, J. M., Breslow, N. E. & Beckwith, J. B. Wilm's tumor and parental age: a report from the national Wilm's tumour study. *Br. J. Cancer* **67**, 813–818 (1993).
28. Hassold, T. *et al.* Human aneuploidy: Incidence, origin, and etiology. *Env. Mol. Mutagen.* **28**, 167–175 (1996).
29. Dryja, T. P., Morrow, J. F. & Rapaport, J. M. Quantification of the paternal allele bias for new germline mutations in the retinoblastoma gene. *Hum. Genet.* **100**, 446–449 (1997).
30. Oldridge, M. *et al.* Genotype–phenotype correlation for nucleotide substitutions in the IgII–IgIII linker in FGFR2. *Hum. Mol. Genet.* **6**, 137–143 (1997).
31. Olshan, A. F. *et al.* Paternal age and the risk of congenital heart defects. *Teratology* **50**, 80–84 (1994).
32. Zhang, Y. *et al.* Parental age at child's birth and son's risk of prostate cancer. *Am. J. Epidemiol.* **150**, 1208–1212 (1999).
33. Tellier, A. L. *et al.* CHARGE syndrome: Report of 47 cases and review. *Am. J. Med. Genet.* **76**, 402–409 (1998).
34. Fletcher, N. A. & Foley, J. Parental age, genetic mutation, and cerebral palsy. *J. Med. Genet.* **30**, 44–46 (1993).
35. Bertram, L. *et al.* Paternal age is a risk factor for Alzheimer disease in the absence of a major gene. *Neurogenetics* **1**, 277–280 (1998).
36. Antonarakis, J. P. *et al.* Factor VIII gene inversions in severe hemophilia: Results of an international consortium study. *Blood* **86**, 2206–2212 (1995).
37. Cody, J. D. *et al.* Preferential loss of the paternal alleles in the 18q syndrome. *Am. J. Med. Genet.* **69**, 280–286 (1997).
38. Dallapiccola, B. *et al.* Parental origin of chromosome-4p deletion in Wolf–Hirschhorn syndrome. *Am. J. Med. Genet.* **47**, 921–924 (1993).
39. Overhauser, J. *et al.* Parental origin of chromosome-5 deletions in the Cri-du-chat syndrome. *Am. J. Med. Genet.* **37**, 83–86 (1990).
40. Duyao, M. *et al.* Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nature Genet.* **4**, 387–392 (1993).
41. Kodaira, M., Satoh, C., Hiyama, K. & Toyama, K. Lack of effects of atomic bomb radiation on genetic instability of tandem-repetitive elements in human germ cells. *Am. J. Hum Genet.* **57**, 1275–1283 (1995).
42. Ellegren, H. Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature Genet.* **24**, 400–402 (2000).
43. Xu, S., Peng, M., Fang, Z. & Xu, X. The direction of microsatellite mutations is dependent upon allele length. *Nature Genet.* **24**, 396–399 (2000).
44. Eyre-Walker, A. & Keightley, P. D. High genomic deleterious mutation rates in hominids. *Nature* **397**, 344–347 (1999).
The first estimate of the total deleterious mutation rate for all genes in humans.
45. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, 1983).
46. Gianneli, F. *et al.* Mutation rates in humans. II. Sporadic mutation-specific rates and rate of detrimental human mutations inferred from hemophilia B. *Am. J. Hum. Genet.* **65**, 1580–1587 (1999).
47. Cargill, M. *et al.* Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genet.* **22**, 231–238 (1999).
48. Aparicio, S. A. J. R. How to count. . . human genes. *Nature Genet.* **25**, 129–126 (2000).
49. Haldane, J. B. S. The effect of variation on fitness. *Am. Nat.* **71**, 337–349 (1937).
The original demonstration that mildly deleterious mutations cause the same reduction of fitness in the long run as drastic ones.
50. Muller, H. J. Our load of mutations. *Am. J. Hum. Genet.* **2**, 111–176 (1950).
51. Crow, J. F. The high spontaneous mutation rate: is it a health risk? *Proc. Natl Acad. Sci. USA* **94**, 8380–8386 (1997).
52. Crow, J. F. Spontaneous mutation in man. *Mutat. Res.* **437**, 5–9 (1999).
53. Akiyama, M. *et al.* Mutation frequency in human blood cells increases with age. *Mutat. Res.* **338**, 141–149 (1995).
54. Bulmer, M. G. *The Mathematical Theory of Quantitative Genetics* (Clarendon, Oxford, 1985).
55. Crow, J. F. & Kimura, M. Efficiency of truncation selection. *Proc. Natl Acad. Sci. USA* **76**, 396–399 (1979).
A demonstration that quasi-truncation selection is almost as effective as strict truncation in eliminating deleterious mutant genes from the population.
56. Stern, C. Wilhelm Weinberg, 1862–1937. Genetics **47**, 1–5 (1962).

Acknowledgements

B. Dove has made several useful suggestions.