5. Watts, D. J. & Strogatz, S. H. *Nature* **393,** 440–442 (1998).
6. Bunde, A. & Havlin, S. *Fractals in Science* (Springer, Berlin, 1994).
7. Barthélémy, M. & Amaral, L. A. N. *Phys. Rev. Lett.* **82,** 3180–3183 (1999).
8. Barabási, A.-L., Albert, R. & Jeong, H. <http://www.nd.edu/~networks>.

Internet

Growth dynamics of the World-Wide Web

The exponential growth of the World-Wide Web has transformed it into an ecology of knowledge in which highly diverse information is linked in an extremely complex and arbitrary manner. But even so, as we show here, there is order hidden in the web. We find that web pages are distributed among sites according to a universal power law: many sites have only a few pages, whereas very few sites have hundreds of thousands of pages. This universal distribution can be explained by using a simple stochastic dynamical growth model.

The existence of a power law in the growth of the web not only implies the lack of any length scale for the web, but also allows the expected number of sites of any given size to be determined without exhaustively crawling the web. The distribution of site sizes for crawls by Alexa and Infoseek is shown in Fig. 1. Both data sets display a power law over several orders of magnitude, so on a log–log scale the distribution of the number of pages per site appears as a straight line. This distribution should not be confused with Zipf's like distributions[1,2], where a power law arises from rank ordering the variables[3].

In order to describe the growth process underlying this distribution[4], we assume that the day-to-day fluctuations in site size are proportional to the size of the site. One would not be surprised to find that a site with a million pages has lost or gained a few hundred pages on any given day. On the other hand, finding an additional hundred pages on a site with just ten pages within a day would be unusual. So we assume that the number of pages on the site, n, on a given day, is equal to the number of pages on that site on the previous day plus or minus a random fraction of n.

If a set of sites is allowed to grow with the same average growth rate but with individual random daily fluctuations in the number of pages added, their sizes will be distributed log-normally after a sufficiently long period of time[5]. A log-normal distribution gives high probability to small sizes and small, but significant, probability to very large sizes. But although it is skewed and has a long tail, the log-normal distribution is not a power-law one.

Two additional factors that determine the growth of the web need to be considered: sites appear at different times and grow at different rates. The number of web sites has been growing exponentially since its inception, which means that there are many more young sites than old ones. Once the age of the site is factored in to the multiplicative growth process, $P(n)$, the probability of finding a site of size n, is a power law, that is, it is proportional to $n^{-\beta}$. Similarly, considering sites with a wide range of distributions in growth rates yields the same result: a power-law distribution in site size. The simple assumption of stochastic multiplicative growth, combined with the fact that sites appear at different times and/or grow at different rates, therefore leads to an explanation of the observed power-law behaviour.

The existence of this universal power law, which is yet another example of the strong regularities[6,7] revealed by studies of the web, also has practical consequences. The expected number of sites of any arbitrary size can be estimated, even if a site of that size has not yet been observed. This can be achieved by extrapolating the power law to any large n; for example, $P(n_2) = P(n_1) \times (n_2/n_1)^{-\beta}$. The expected number of sites of size n_2 in a crawl of N sites would be $NP(n_2)$. For instance, from the Alexa data we can infer that, if data were collected from 250,000 sites, the probability of finding a site with a million pages would be 10^{-4}. This information is not readily available from the crawl alone, as it stops at 10^5 pages per site.

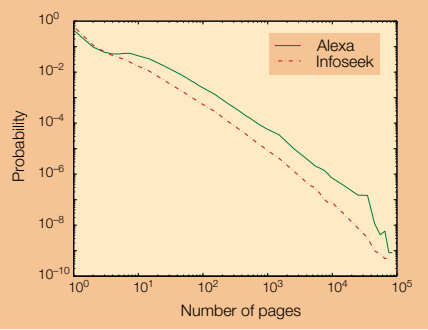

Figure 1 Log–log plot of the distribution of pages in sites for Alexa and Infoseek crawls, which covered 259,794 and 525,882 sites, respectively. There is a drop-off at approximately 10^5 pages because server limitations mean that search engines do not systematically collect more pages per site than this. A linear regression on the variables log(number of sites) and log(number of pages) yielded [1.647, 1.853] as the 95% confidence interval for the exponent β in the Alexa crawl, and [1.775, 1.909] for the Infoseek crawl. These estimates for the power-law slope are consistent across the two data sets and with the model, which predicts that β is greater than 1.

Bernardo A. Huberman, Lada A. Adamic
Xerox Palo Alto Research Center,
3333 Coyote Hill,
Palo Alto, California 94304, USA
e-mail: ladamic@parc.xerox.com

1. Zipf, G. K. *Human Behavior and the Principle of Least Effort* (Addison-Wesley, Cambridge, Massachusetts, 1949).
2. Mantegna, R. N. *et al. Phys. Rev. E* **52,** 2939–2950 (1995).
3. Gunther, R. *et al. Int. J. Theor. Phys.* **35,** 395–417 (1996).
4. http://www.parc.xerox.com/iea/www/growth.html
5. Crow, E. L & Shimizu, K. *Lognormal Distributions: Theory and Applications* (Dekker, New York, 1988).
6. Huberman, B. A., Pirolli, P., Pitkow, J. & Lukose, R. M. *Science* **280,** 95–97 (1998).
7. Huberman, B. A. & Lukose, R. M. *Science* **277,** 535–538 (1997).

Genome evolution

Global methylation in eutherian hybrids

O'Neill *et al.* propose that epigenetic processes help to drive karyotypic evolution in marsupials[1]. Here we present evidence that global methylation patterns do not undergo dramatic changes in interspecific hybrids among three orders of placental mammals, indicating that the mechanisms underlying genome evolution may be different in placental mammals and marsupials.

Interspecific hybridization in mammals frequently results in male sterility[2], abnormal growth[2] and placental dysplasia[3–5], which together may cause post-meiotic reproductive isolation. It has been proposed that incompatibility between rapidly evolving genes that interact normally in the intraspecific context[6] and genomic rearrangements[7] may explain interspecific hybrid defects.

O'Neill *et al.* have given a striking example for the latter mechanism in an interspecific hybrid of the marsupials *Macropus eugenii* and *M. bicolor*[1]. This first-generation (F_1) hybrid exhibited genome-wide demethylation, retrotransposon amplification and centromere expansion on the autosomes derived from *M. eugenii*. Undermethylation of F_1 genomes compared with those of the parental species was also seen in two hybrids of other species within the genus *Petrogale*[1]. These findings were taken to indicate that retrotransposon amplification and chromosome expansion secondary to genome-wide undermethylation could be a frequent phenomenon in mammalian hybrids, leading to rapid karyotypic evolution and finally to reproductive isolation[1].

We have analysed genome-wide methylation in interspecific hybrids in the placental mammalian families of three orders, Equidae (Perissodactyla), Muridae (Rodentia) and Camelidae (Artiodactyla), by following the digestion of genomic DNA with the methylation-sensitive and methylation-insensitive enzymes *Hpa*II and *Msp*I, respectively, and Southern blotting the digest. This analysis included hybrids between horse and donkey, three species of mouse (*Mus musculus*, *Mus spretus* and *Mus macedonicus*), and llama (*Lama glama*) and dromedary (*Camelus dromedarius*). This analysis gave no indication for any changes in genome-wide methylation in any of the F_1 hybrids when compared with parental animals (Fig. 1).

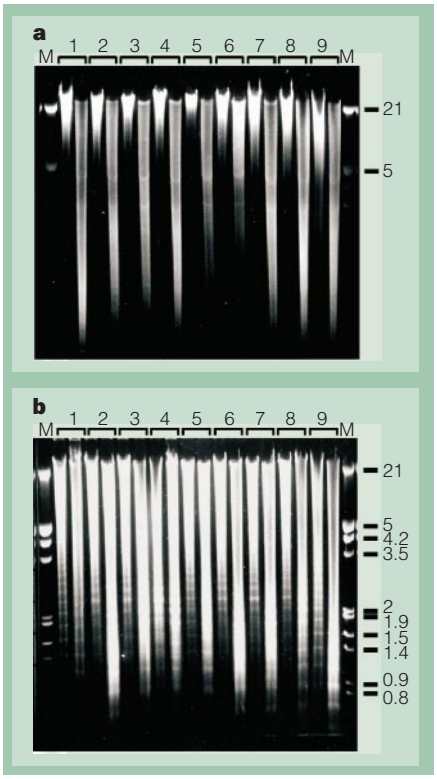

Figure 1 Genome methylation analysis in interspecific hybrids by *Hpa*II (left lanes) and *Msp*I (right lanes) digestion. **a,** Hybridization between *Mus spretus* and *M. musculus*. Lane 1, male *M. spretus*; 2, male *M. musculus* × *M. spretus* F_1; 3, male *M. musculus* × *M. spretus* F_1; 4, male *M. musculus* × *M. spretus* F_1; 5, male *M. musculus* × *M. musculus*; 6, female *M. musculus*; 7, female *M. musculus* × *M. spretus* F_1; 8, female *M. musculus* × *M. musculus*; 9, female *M. spretus*. **b,** Hybridization between horse and donkey (mule) and donkey and horse (hinny). Lane 1, female horse; 2, female mule; 3, female mule 2; 4, male donkey; 5, male horse; 6, female hinny; 7, female donkey; 8, female hinny; 9, male horse. In all matings, the female is indicated first. Each lane corresponds to an individual animal. M, marker.

In addition, we assessed cytosine methylation by *in situ* nick-translation on the metaphase chromosomes of mouse F_1 hybrids and parental mice. For this analysis, mice of the *Mus musculus* strain Cremona were used. Because this strain has an aberrant chromosome number ($2N = 22$ instead of 40), it is possible to discriminate between the chromosomes derived from *Mus musculus* Cremona and *Mus spretus* in F_1 hybrids (Fig. 2). This analysis indicated that no major change had occurred in genome-wide demethylation or in centromere expansion. Finally, we were unable to detect amplification of the L1 retrotransposon by Southern-blot analysis (results not shown).

These results do not appear to agree with those obtained by O'Neill *et al.*[1] in their analysis of interspecific hybrids between several different species of macropodid marsupials. However, their findings are unequivocal and supported by earlier cytogenetic investigations of hybrids between additional macropodid species[8]. Genome-wide undermethylation, retrotransposon

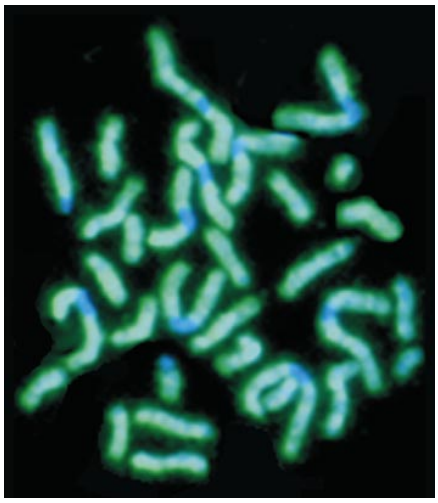

Figure 2 Methylation analysis of *M. musculus* Cremona × *M. spretus* chromosomes. Metaphase spread from bone marrow was digested with *Hpa*II followed by *in situ* nick-translation. Staining patterns were identical with *Msp*I. There was no detectable difference in labelling intensity between *M. spretus* and *M. musculus* euchromatin. The faint labelling of *M. musculus* compared with *M. spretus* centromeric heterochromatin reflects the low frequency of *Msp*I/*Hpa*II sites in *M. musculus* major satellite DNA.

activation and chromosome extension may therefore be specific to interspecific hybridization in marsupials, or perhaps they occur only in macropodid marsupials.

In any case, our results argue against the idea that such profound alterations in genome organization of interspecific hybrids are common events in placental mammals. Marsupials and placental mammals diverged about 130 million years ago[9], so the functional role of methylation may have changed between the two subclasses. For example, marsupial and placental mammals show pronounced differences in their processes of X-chromosome inactivation[9], in which the role of methylation is thought to be important[10].

Irmgard Roemer*, Frank Grützner*, Heinz Winking†, Thomas Haaf*, Annie Orth‡, Lulu Skidmore§, Doug Antczak¶, Reinald Fundele*

**Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73, 14195 Berlin, Germany*
e-mail: fundele@mpimg-berlin-dahlem.mpg.de
†*Institut für Biologie, Medizinische Universität Lübeck, 23562 Lübeck, Germany*
‡*Laboratoire Génome et Populations, Université de Montpellier II, 34095 Montpellier, France*
§*Camel Reproduction Centre, Dubai, United Arab Emirates*
¶*James A. Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853, USA*

1. O'Neill, R. J. W., O'Neill, M. J. & Graves, J. A. M. *Nature* **393,** 68–72 (1998).
2. Gray, A. P. in *Mammalian Hybrids* (Commonwealth Agricultural Bureau, Edinburgh, 1971).
3. Allen, W. R., Skidmore, J. A., Stewart, F. & Antczak, D. F. *J. Reprod. Fertil.* **9,** 55–60 (1993).
4. Rogers, J. F. & Dawson, W. D. *J. Reprod. Fertil.* **21,** 255–262 (1970).
5. Zechner, U. *et al. Nature Genet.* **12,** 398–403 (1996).
6. Orr, H. A. *Annu. Rev. Ecol. Syst.* **28,** 195–218 (1997).
7. Templeton, A. R. *Annu. Rev. Ecol. Syst.* **12,** 23–48 (1981).
8. Smith, M. J., Hayman, D. L. & Hope, R. M. *Aust. J. Zool.* **27,** 959–972 (1979).
9. Graves, J. A. M. *Trends Genet.* **3,** 252–256 (1987).
10. Migeon, B. R. *Trends Genet.* **10,** 415–417 (1994).

O'Neill et al. *reply* — The absence of global methylation changes in eutherian interspecific hybrids compared with their parents, observed by Roemer *et al.*, sharply contrasts with our own studies of interspecific hybrids between various species of kangaroo. We observed hybrid-specific undermethylation, retroelement activation and genome remodelling[1], and suggested that these events occurring together may bring about rapid karyotypic change.

We agree with Roemer *et al.* that marsupials and eutherians are likely to have diverged from one another in their reliance on and use of the epigenetic information conveyed by DNA methylation. Although these events may be specific to marsupial or even macropod interspecific hybrids, there is evidence that eutherian genomes may be subject to at least some degree of the same sort of hybrid dysgenic perturbations. Interspecific hybrids of the genus *Peromyscus* (deer mice) do not show whole-genome changes in methylation, as determined by digestion with *Msp*I and *Hpa*II (R.J.W.O'N. *et al.*, unpublished observations), yet they exhibit disruptions in imprinted gene expression associated with allele-specific undermethylation[2]. The mechanism underlying the loss of imprinting in these hybrids remains unknown, but there is a subtle change in methylation in this eutherian hybrid cross[2]. Digestion of genomic DNA with *Msp*I and *Hpa*II may be too blunt an instrument to reveal subtle changes (less than 20%) in methylation.

Our main finding is a link between DNA methylation, retroelement activity and genome rearrangement. The dramatic perturbations of methylation and genome structure that we observed in kangaroo hybrids may be an extreme example of dysgenic changes that occur on a broader scale in many organisms.

Rachel J. Waugh O'Neill*, Michael J. O'Neill*, Jennifer A. Marshall Graves†

**Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269, USA*
e-mail: roneill@uconnvm.uconn.edu
†*Department of Genetics and Human Variation, La Trobe University, Bundoora, Victoria 3083, Australia*

1. O'Neill, R. J. W., O'Neill, M. J. & Graves, J. A. M. *Nature* **393,** 68–72 (1998).
2. Vrana, P. B. *et al. Nature Genet.* **20,** 362–365 (1998).