Incorrect use of the term synteny

The term 'synteny' (or syntenic) refers to gene loci on the same chromosome regardless of whether or not they are genetically linked by classic linkage analysis[1]. This term was introduced in 1971 by John H. Renwick, of the London School of Hygiene and Tropical Medicine, at the 4th Internal Congress of Human Genetics in Paris with one of us (E.P.) in attendance. The need for such a term was suggested to J.H. Renwick by E.A. Murphy, of Johns Hopkins University[2]. It arose as a consequence of the new methods in gene mapping using somatic cell hybrid cells. Human genes located on the same chromosome with a genetic distance that could not be determined by the frequency of recombination lacked a term of reference.

'Synteny' means 'same thread' (or ribbon), a state of being together in location, as synchrony would be together in time. Although several textbooks[3–10] and other reference works[11–15] give a correct definition, the term synteny nowadays is often used to refer to gene loci in different organisms located on a chromosomal region of common evolutionary ancestry. This new usage of the term synteny does not correspond to its original definition and correct language derivation. A survey of 11 articles in *Nature Genetics* since 1992 using the term syntenic or synteny in either the title or the abstract revealed usage incorrect in 8 and ambiguous in 3.

We believe molecular biologists ought to respect the original definition of synteny and its etymological derivation, especially as this term is still needed to refer to genes located on the same chromosome. We recognize the need to refer to gene loci of common ancestry. Correct terms exist: 'paralogous' for genes that arose from a common ancestor gene within one species and 'orthologous' for the same gene in different species.

Eberhard Passarge[1], Bernhard Horsthemke[1] & Rosann A. Farber[2]
[1]Institut für Humangenetik, Universitätsklinikum Essen, Essen, Germany. [2]Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. Correspondence should be addressed to E.P. (e-mail: eberhard.passarge@uni-essen.de).

1. Renwick, J.H. *Annu. Rev. Genet.* **5**, 81–120 (1972).
2. Renwick, J.H. in *Human Genetics. Proceedings of the Fourth International Congress of Human Genetics, 6-11 September 1971* (eds de Grouchy, J., Ebling, F.J.G. & Henderson, I.W.) 443–444 (Excerpta Medica, Amsterdam, 1972).
3. Levitan, M. *Textbook of Human Genetics* (Oxford University Press, Oxford, 1988).
4. Sutton, M.E. *An Introduction to Human Genetics* (Harcourt Brace Jovanovich, San Diego, 1988).
5. Thompson, M.W., McInnes, R.R. & Willard, H.F. *Genetics in Medicine* (W.B. Saunders, Philadelphia, 1991).
6. Rothwell, N.V. *Understanding Genetics. A Molecular Approach* (Wiley-Liss, New York, 1993).
7. Connor, J.M. & Ferguson-Smith, M.A. *Essential Medical Genetics* (Blackwell Scientific, Oxford, 1993).
8. Mueller, R.F. & Young, I.D. *Emery's Elements of Medical Genetics* (Churchill Livingstone, Edinburgh, 1995).
9. Jorde, L.B., Carey, J.C. & White, R.L. *Medical Genetics* (Mosby, St. Louis, 1995).
10. Gelehrter, T.D., Collins, F.S. & Ginsburg, D. *Principles of Medical Genetics* (Williams & Wilkins, Baltimore, 1998).
11. Murphy, E.A. & Chase, G.E. *Principles of Genetic Counseling* (Year Book Medical, Chicago, 1975).
12. Rieger, R., Michaelis, A. & Green, M.M. *Glossary of Genetics and Cytogenetics* (Springer Verlag Heidelberg, New York, 1976).
13. Ott, J. *Analysis of Human Genetic Linkage* (Johns Hopkins University Press, Baltimore, 1991).
14. Strachan, T. & Read, A.P. *Molecular Human Genetics* (BIOS Scientific, London, 1996).
15. Snustad, D.P., Simmons, M.J. & Jenkins, J.B. *Principles of Genetics* 749 (John Wiley & Sons, New York, 1997).

Analysis of human transcriptomes

How many human genes are expressed ubiquitously, in all human tissues, and how many are expressed in tissue-specific patterns? To answer these fundamental questions in molecular biology, we have analysed 3.5-million transcripts from 19 normal and diseased tissue types. We found that as many as 43,500 genes can be expressed in a single cell type. Only a small fraction of transcripts were exclusively expressed in any individual tissue, whereas nearly 1,000 genes were expressed in all cell types examined. We found 40 genes to be expressed at elevated levels in all cancer tissues but not in normal cells.

Table 1 • Tissues and transcript tags

	Libraries[b]	Total transcripts[c]	Unique genes[d]
Normal tissues[a]			
colon epithelium[5–9]	2	98,089	12,941
keratinocytes[e]	2	83,835	12,598
breast epithelium[e]	2	107,632	13,429
lung epithelium[10]	2	111,848	11,636
melanocytes[e]	2	110,631	14,824
prostate[e]	2	98,010	9,786
monocytes[e]	3	66,673	9,504
kidney epithelium[e]	2	103,836	15,094
chondrocytes[e]	4	88,875	11,628
cardiomyocytes[e]	4	77,374	9,449
brain[9]	3	202,448	23,580
Diseased tissues[a]			
colon cancer[5–9,11,e]	22	1,004,509	56,153
pancreatic cancer[5–8]	4	126,414	17,050
breast cancer[e]	5	226,630	18,685
lung cancer[10]	5	221,302	22,783
melanoma[e]	10	269,332	25,600
polycystic kidney disease[e]	2	112,839	16,280
haemangiopericytoma[e]	5	199,985	31,351
brain cancer[9]	3	186,567	23,108
Total	84	3,496,829	84,103

[a]Source of the RNA analysed. [b]The number of SAGE libraries analysed. [c]The total number of transcripts analysed from each tissue. [d]The number of unique genes observed in each tissue. [e]Unpublished data. See Methods (http://genetics.nature.com/supplementary_info/) for the derivation of a number of unique genes.

Serial analysis of gene expression (SAGE) studies[1,2] of 84 libraries derived from 19 different sources identified 134,135 transcripts from approximately 84,000 different genes (Table 1; data and analysis available at http://genetics.nature.com/supplementary_info/). Expression levels for these genes ranged from 0.3 to 9,417 transcript copies per cell. The transcript tags matched approximately 4,300 known genes and 41,000 genes with unknown functions, whereas the remaining transcript tags (46%) had no matches to existing databases (Table 2, see http://genetics.nature.com/supplementary_info/).

The subset of expression data from colorectal cancer cell lines provided the first relatively complete analysis of the transcripts expressed in a single mammalian cell type. We analysed 643,283 transcripts from colorectal cancer cell lines. As human cells contain approximately 300,000 mRNA molecules, this number was sufficient to provide approximately twofold coverage of the transcriptome, revealing over 83% of transcripts expected to be pre-