Published in final edited form as:
J Invest Dermatol. 2009 July ; 129(7): 1599–1601. doi:10.1038/jid.2009.81.

A mouse by any other name…

John P. Sundberg and **Paul N. Schofield**
The Jackson Laboratory, Bar Harbor, ME and Department of Physiology Development and Neuroscience, University of Cambridge, Cambridge, UK

An often-heard policy motto from the UK's Blair Government was "Education, Education, Education". An appropriate exhortation for the biomedical sciences should be "Standardization, Standardization, Standardization". Inevitably the two go hand in hand, and the current challenge we face is how to encourage researchers to comply with existing or emerging standard terminologies and nomenclatures. This is both an educational and a regulatory task, but one in which it is vital to succeed if we are to efficiently and accurately share and use the huge volume of data now emerging in the biosciences.

We communicate effectively in medicine and biomedical research primarily through use of specialized and defined standard terminology (Becker, 1959; Brown *et al.*, 2007; Friedman, 1955; Jackson, 2001; Porter, 2006; Taylor, 2006). This terminology, while fluid over time to accommodate new discoveries and technological innovations, remains the key to advances in science. Within the field of dermatology, a committee has functioned for decades to standardize nomenclature (Becker, 1959). Reviewers for the *Journal of Investigative Dermatology* and other publications in the field need to insist on strict adherence to the most current dermatological nomenclature in order to maintain the high esteem of members in the field.

Similar efforts for standardization are currently underway with the pathology of the laboratory mouse, now the pre-eminent model system for human disease (Rosenthal and Brown, 2007). While more complicated than dermatology, because this approach covers all organ systems and merges veterinary and human medical terminology, various panels have been formed to address this issue. For cancer, the National Cancer Institute's Mouse Models for Human Cancer Consortium created panels of specialists to review mouse models for human cancer by organ system to develop a consensus nomenclature (http://emice.nci.nih.gov/emice) (Cardiff *et al.*, 2000; Kogan *et al.*, 2002; Nikitin *et al.*, 2004; Shappell *et al.*, 2004). A more extensive website, the Mouse Tumor Biology Database (MTB; http://tumor.informatics.jax.org) incorporates the mouse genetic literature with images of all types of cancer arising either spontaneously in mice of inbred strains or as a consequence of genetically engineering (Bult *et al.*, 2006; Naf *et al.*, 2002). For general mouse pathology, an international consortium was formed to develop MPATH, an evolving and expanding ontology of mouse pathology terms. This is linked to a large image database (http://www.Pathbase.net). These on-line resources are supplemented by highly specialized residential training courses and internship programs (Sundberg *et al.*, 2007), but even with these opportunities there still remains a significant gap between demand and availability of appropriately trained pathologists (Schofield *et al.*, 2009).

Recently the second CASIMIR (Coordination And Sustainability of International Mouse Informatics Resources (http://www.casimir.org.uk) annual meeting, held at the Nobel

Corresponding Author: John P. Sundberg, DVM, PhD, The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609-1500, john.sundberg@jax.org, FAX: 207-288-6073, Phone: 207-288-6410.

Forum, Stockholm, Sweden, 2–3 December 2008, focused on the topic “One Medicine, One Pathology”, with the goal of coordinating data collection, nomenclature, and comparative pathology among various disciplines (Sundberg and Schofield, in press). These approaches refine existing nomenclature systems developed over the previous two centuries with which all medically trained scientists are familiar. We should all strive to utilize these online resources to double check our interpretations as well as a way to provide standardization to our publications. To that end databases are now available that provide a “virtual second opinion” for mouse pathology nomenclature with links to photomicrographs (http://research.jax.org/faculty/sundberg/index.html) (Sundberg *et al.*, in press; Sundberg *et al.*, 2008).

A larger and far more serious nomenclature issue involves genetic terminology. This is an area in which few of us were trained.

Rules for genetic nomenclature were devised in 1919 when the American Society of Naturalists appointed a Committee on Genetic Form and Nomenclature with C.C. Little as chairman (Little, 1921). As applied to the mouse, these were published in 1940 by Dunn, Grueneberg, and Snell (Dunn *et al.*, 1940) and subsequently the International Committee for Standardized Genetic Nomenclature in Mice (Green *et al.*, 1963) was formed in 1963 to standardize inbred, congenic, and recombinant inbred strain nomenclature as well as mutant locus/gene symbols.

New names for mutant mouse strains and stocks (and later for genes) were proposed to this committee by investigators, and unique names and symbols were assigned to prevent ambiguity. Unfortunately, journal editors have been slow to require authors to adhere to this system, creating major problems. Today, with the advent of genetic engineering and large-scale mutagenesis projects, multiple allelic mutations (both spontaneous or chemical/radiation-induced “remutations”, and multiple constructs of targeted mutations involving the same gene), often with very different phenotypes, are available. Strain and mutation symbols, when used correctly, are critical parts of the materials and methods section of any manuscript and help reviewers determine which allelic mutation is under investigation to enable them to determine the validity of the work being reported.

Mouse (International Committee on Standardized Genetic Nomenclature for Mice: http://www.informatics.jax.org/mgihome/nomen/strains.shtml), human (HUGO Gene Nomenclature Committee: http://www.genenames.org/), and rat (Rat Gene Nomenclature Committee: http://rgnc.gen.gu.se/RGNChem.html) nomenclature rules are available online. For laboratory mice, inbred strain names are all in capital letters. After a forward slash (/) following the strain name, and equally important, are the laboratory (investigator) and institutional codes that designate substrains. For example, NOD/ShiLtSzJ designates a subline of the Non-Obese Diabetic (NOD) strain originally inbred at Shionogi and Company, Ltd., (Shi) in Japan, and later maintained by Dr. Edward Leiter (Lt) from whose colony a subline was initiated by Dr. Leonard Shultz (Sz). The strain is maintained and distributed by The Jackson Laboratory (J). Abbreviations for commonly used inbred strains are also standardized. C57BL/6J is abbreviated B6, which is also used to refer to mixed or unknown/unspecified C57BL/6 substains. B6ByJ refers to C57BL/6ByJ and B6EiJ to C57BL/6JEi, which, like many other substrains carry unique mutations. By contrast, the BALB/cJ inbred strain is abbreviated C and BALB/cByJ mice are CBy.

Mixed inbred or incipient congenic strains, in which a mutated gene is being transferred from one strain background onto another strain, are designated with a semicolon between the strain abbreviations, e.g. B6;129, followed by a hyphen and the mutant gene symbol. This nomenclature, indicating a segregating background, is in sharp contrast to congenic

strain names, in which the semicolon is replaced by a period to indicate the congenic procedure has been completed (10 backcrosses, N10, onto the new strain; e.g., B6.129). Six backcross generations (N6; incipient congenic) are commonly accepted by many journals to be adequate, and many mouse distributors use congenic nomenclature at N5; however, speed congenic technology has shown that this is not optimal (Markel *et al.*, 1997).

Mouse gene symbols are printed in italics with the first letter capitalized and subsequent letters in lower case. Symbols for dominant or semi-dominant spontaneous or chemical/radiation-induced mutations of unidentified genes are written in the same manner as gene symbols. Recessive allelic mutations are written all in lower case. Once the previously unknown gene is identified, the allele (mutation) symbol is superscripted immediately, after the gene symbol. For example, the mouse hairless and rhino Jackson mutations are written Hr^{hr} and Hr^{rh7J}, respectively. To differentiate these from human genes, human gene symbols are written entirely in capital letters; the human hairless gene symbol is *HR*. For both mouse and human, gene and allele names (as opposed to symbols) are written entirely in lower case unless they include a proper noun (e.g. Alstrom syndrome I). Whereas gene symbols are italicized, symbols for their respective proteins are not. Both mouse and human protein symbols are printed entirely in capital letters.

Specific nomenclature guidance for strains, genes, alleles/mutations, and chromosomal aberrations can be found on the Mouse Genome Informatics website via links from the Nomenclature Home Page (http://www.informatics.jax.org/mgihome/nomen/gene.shtml). Strict adherence to these nomenclature standards will allow work to be fairly compared and, more importantly, reviewed accurately.

The power of informatics to integrate and analyze phenotype and genotype data within and across species is increasing all the time, although it is still outstripped by the volume of data emerging, particularly from the analysis of mouse mutants. It is essential that the way in which alleles are expressed and the disease descriptions captured are semantically unambiguous and standardized to allow computational analysis. This is a serious barrier to the analysis of large historical datasets where local nomenclature and data structure are idiosyncratic, but it is becoming a rate-limiting step in the analysis of new data, particularly those published only in the printed literature and not uploaded to databases, as failure to use correct terminology results in ambiguity and inaccuracy which is very difficult to deal with, for example using text mining tools. This results in the need for laborious and expensive extraction of the data by professional curators.

Acknowledgments

We thank Beverly Richards-Smith for her constructive comments and verification of nomenclature issues raised in this manuscript. This work was supported by grants from the National Institutes of Health (CA34196, CA089713, and RR17436) and the Commission of the European Community, Framework Programme 6 contract no. LSHG-CT-2006-037811, CASIMIR.

References

Becker SW. Dermatological nomenclature. Arch Derm. 1959; 80:778–787. [PubMed: 13848978]

Brown SJ, Tilli CMLJ, Jackson B, Avilion AA, MacLeod MC, Maltais LJ, et al. Rodent *Lce* gene clusters; new nomenclature, gene organization, and divergence of human and rodent genes. J Invest Dermatol. 2007; 127:1782–1786. [PubMed: 17410201]

Bult, CJ.; Krupke, DM.; Vincent, MJ.; Allio, T.; Sundberg, JP.; Mikaelian, I., et al. The Mouse Tumor Biology Database: an online resource for mouse models of human cancer. In: Nagl, S., editor. Cancer bioinformatics: from therapy design to treatment. West Sussex, England: John Wiley & Sons, Ltd; 2006. p. 143-153.

Cardiff RD, Moghanaki D, Jensen RA. Genetically engineered mouse models of mammary intraepithelial neoplasia. J Mammary Gland Biol Neoplasia. 2000; 5:421–437. [PubMed: 14973386]

Dunn LC, Grüneberg H, Snell GD. Report on the committee of mouse genetics nomenclature. J Hered. 1940; 31:505–506.

Friedman AA. Medical nomenclature and dermatology. Va Med Mon (1918). 1955; 82:516. [PubMed: 13274729]

Green MC, Grueneberg H, Hertwig P, Heston WE, Lyon MF, Medvedev NN, et al. A Revision of the Standardized Genetic Nomenclature for Mice. J Hered. 1963; 54:159–162. [PubMed: 14057864]

Jackson R. What's in a Name? Comments on the Dermatological Dictionary by Ledier, Rosenblum, and Carter. J Cutan Med Surg. 2001; 5:505–507. [PubMed: 11907860]

Kogan SC, Ward JM, Anver MR, Berman JJ, Brayton C, Cardiff RD, et al. Bethesda proposals for classification of nonlymphoid hematopoietic neoplasms in mice. Blood. 2002; 100:238–245. [PubMed: 12070033]

Little CC. Report of the Committee on Genetic form and Nomenclature. The American Naturalist. 1921; 55:175.

Markel P, Shu P, Ebeling C, Carlson GA, Nagle DL, Smutko JS, et al. Theoretical and empirical issues for marker-assisted breeding of congenic mouse strains. Nat Genet. 1997; 17:280–284. [PubMed: 9354790]

Naf D, Krupke DM, Sundberg JP, Eppig JT, Bult CJ. The mouse tumor biology database: a public resource for cancer genetics and pathology of the mouse. Cancer Res. 2002; 62:1235–1240. [PubMed: 11888882]

Nikitin A, Alcaraz A, Anver MR, Bronson RT, Cardiff RD, Dixon D, et al. Classification of proliferative pulmonary lesions of the mouse: recommendations of the mouse models of human cancers consortium. Cancer Res. 2004; 64:2307–2316. [PubMed: 15059877]

Porter RM. The new keratin nomenclature. J Invest Dermatol. 2006; 126:2366–2368. [PubMed: 17041617]

Rosenthal N, Brown S. The mouse ascending: perspectives for human-disease models. Nature Cell Biol. 2007; 9:993–999. [PubMed: 17762889]

Schofield PN, Brown SD, Sundberg JP, Arends M, Warren MV, Dubus P, et al. PRIME importance of pathology expertise. Nat Biotechnol. 2009; 27:24–25. [PubMed: 19131991]

Shappell SB, Thomas GV, Roberts RL, Herbert R, Ittmann MM, Rubin MA, et al. Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. Cancer Res. 2004; 64:2270–2305. [PubMed: 15026373]

Sundberg BA, Schofield PN, Gruenberger M, Sundberg JP. A data capture tool for mouse pathology phenotyping. Vet Pathol. (in press).

Sundberg JP, Hackman RC, HogenEsch H, Nikitin AY, Ward JM. Training Mouse Pathologists: Five years of Pathology of Mouse Models of Human Disease Workshops. Toxicol Pathol. 2007; 35:447–448. [PubMed: 17474065]

Sundberg JP, Schofield PN. One medicine, one pathology, one health concept. J Am Vet Med Assoc. (in press).

Sundberg JP, Sundberg BA, Schofield PN. Integrating mouse anatomy and pathology ontologies into a diagnostic/phenotyping database: tools for record keeping and teaching. Mammalian Genome. 2008; 19:413–419. [PubMed: 18797968]

Taylor S. As simple as black and white? J Amer Acad Dermatol. 2006; 54:1070–1071. [PubMed: 16713466]