Esperanto for Histones: CENP-A, not CenH3, is the centromeric histone H3 variant

W.C. Earnshaw[1], R.C. Allshire, M. Balasubramanian, B.E. Black, K. Bloom, B.R. Brinkley, W. Brown, I.M. Cheeseman, K.H.A. Choo, G. Copenhaver, J.G. DeLuca, A. Desai, S. Diekmann, S. Erhardt, M. Fitzgerald-Hayes, D. Foltz, T. Fukagawa, R. Gassmann, D.W. Gerlich, D. Glover, G.J. Gorbsky, S.C. Harrison, P. Heun, T. Hirota, L. Jansen, G. Karpen, G.J.P. Kops, K. Luger, M. Lampson, S.M. Lens, A. Losada, H. Maiato, P.S. Maddox, A.D. McAinsh, R. Margolis, H. Masumoto, B.G. Mellone, P. Meraldi, A. Musacchio, K. Oegema, R. O'Neill, E.D. Salmon, K. Scott, P. Sorger, A. Straight, P.T. Stukenberg, B.A. Sullivan, K.F. Sullivan, C.E. Sunkel, J.R. Swedlow, P.E. Warburton, S. Westermann, H.F. Willard, M. Yanagida, T.J. Yen, K. Yoda and D.W. Cleveland[n]

[1]W.C. Earnshaw, Wellcome Trust Centre for Cell Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR, Scotland, U.K.

[n]Ludwig Institute for Cancer Research, Univ. of California at San Diego, La Jolla, CA USA

Since the time of Linnaeus, scientific nomenclature has been based on precedent. Over the past several centuries, the tried and proven path to naming a species (or more recently, a protein) is first to *discover* one, and then name it. In recent years, the rush of scientific progress, with multiple groups often simultaneously discovering and naming the same protein at the same time has stressed the naming convention, and occasionally groups of scientists have stepped in to rationalize the nomenclature.

In 2012, an article entitled "A unified phylogeny-based nomenclature for histone variants" appeared in the journal "Epigenetics and Chromatin" {Talbert, 2012 #1451}. This article had a lengthy list of distinguished authors from the chromatin/epigenetics community, and represents an effort to unify the histone nomenclature. This proposed simplification of naming histone variants follows on the heels of a number of previous distinguished efforts, including the rationalisation of the caspase nomenclature in 1996 {Alnemri, 1996 #1081}, and a proposed standard nomenclature for the kinesin proteins {Lawrence, 2004 #1090}. The caspase proposal was universally adopted almost immediately, as the ten different caspases were known by a host of confusing names at that time. The kinesin article also has been widely influential.

While the proposal to unify the histone nomenclature may have much to recommend it, with respect to the specialized histone H3 variant found at all active centromeres from budding yeast to man, the proposed nomenclature change serves neither scientific fact or scientific traditions well. For the reasons detailed below, we suggest that the molecular cellular biology and genetics communities should maintain the original nomenclature (CENP-A) that has served for 28 years for the centromeric histone H3 variant and avoid the usage of

the misleading name (CenH3) proposed by Talbert et al. {Talbert, 2012 #1451}.

The first known centromeric protein was discovered in human cells and named CENP-A in 1985 {Earnshaw, 1985 #169}. CENP-A was shown to be a histone in 1991 by the late Doug Palmer, working with Bob Margolis {Palmer, 1991 #414}. This conclusion was subsequently confirmed when the protein was cloned by Kevin Sullivan and colleagues {Sullivan, 1994 #530}. CENP-A has been widely referred to by this name over the subsequent 28 years and the CENP nomenclature has now reached as far as CENP-X for well-studied proteins.

It has now been suggested {Talbert, 2012 #1451} that the name CENP-A should be superseded by CenH3 so as to simplify multiple names now in use in multiple species for the histone H3 variant found only at active centromeres. The budding yeast homolog of CENP-A, CSE4, was described in 1995 {Stoler, 1995 #538}, as the product of the *Cse4* gene, which was discovered in a screen for mutations that affected chromosome segregation. A later addition was the *Drosophila* homologue, discovered in 2000 by homology with CENP-A and then named Cid {Henikoff, 2000 #913}. It is an important distinction to *Drosophila* geneticists that Cid was not named because of a pre-existing named mutation (in which case this name would have been retained by tradition). To the contrary, Cid was identified via its sequence similarity to CENP-A. Although the authors knew they had identified the Drosophila variant of CENP-A, they chose instead to give it a new name, Cid. Thus, any confusion in the nomenclature was created by this naming decision. It should be noted that the name CID, while now often used in the Drosophila literature is not mentioned by Talbert et al.

The name proposed in Talbert et al., CenH3, rather than adding clarity to the nomenclature, adds an unnecessary layer of confusion because it is scientifically misleading: its use implies that this protein is *the* centromeric histone H3. This is simply not correct. A range of studies has revealed that regional centromeres contain not only CENP-A, but also lots of canonical histone H3. This canonical centromeric histone H3 is not just a "stuffer" or contaminant of centromere chromatin. Studies ranging from biochemical fractionation {Ando, 2002 #2054;Foltz, 2006 #1569;Hori, 2008 #1866} to high-resolution light microscopy {Blower, 2002 #1128;Sullivan, 2004 #1588;Ribeiro, 2010 #2195} reveal that centromeric canonical H3 nucleosomes are interspersed with CENP-A nucleosomes, and that specific components (e.g., CENP-C and the histone fold-containing CENP-T/W complex {Nishino, 2012 #2300}) that make meaningful contacts with centromeric H3-containing chromatin are important for kinetochore assembly and function {Ohzeki, 2012 #2362}. Recognizing this, the term 'CenH3' would more appropriately refer to centromere associated canonical histone H3 than it does to the centromere-specific CENP-A. Correspondingly, it is inappropriate as a name for the histone H3 variant that is found exclusively at centromeres.

While we appreciate the overall efforts to unify the nomenclature of histones from a phylogenetic perspective, our view is that the proposal by the many authors of the "Epigenetics and Chromatin" article {Talbert, 2012 #1451} to rename CENP-A as CenH3 does not take into account the extensive preceding literature on centromeres or kinetochores. Of high relevance, we note that while the signatories to this Commentary have been primary contributors to the centromere literature and all of us have published (some extensively) on CENP-A, none of us was a co-signatory (or was asked to be a signatory or to comment) on the "Epigenetics and

Chromatin" nomenclature proposal.

Bearing in mind the confusion that will inevitably arise over whether the term CenH3 refers to canonical histone H3 interspersed with CENP-A at centromeres or to the CENP-A itself, we recommend that the proposed name CenH3 be abandoned and that this important marker for centromeric chromatin should be referred to by the name originally given to it in 1985 – CENP-A.

ACKNOWLEDGMENTS

W.C.E. is a Principal Research Fellow of The Wellcome Trust [grant number 073915]. The Wellcome Trust Centre for Cell Biology is supported by core grant numbers 077707 and 092076.

REFERENCES

1. **Talbert, P.B. et al. A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin* 5, 7 (2012).**
2. **Alnemri, E.S. et al. ICE/CED3 protease nomenclature. *Cell* 87, 171 (1996).**
3. **Lawrence, C.J. et al. A standardized kinesin nomenclature. *J. Cell Biol.* 167, 19-22 (2004).**
4. **Earnshaw, W.C. & Rothfield, N. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma (Berl.)* 91, 313-321 (1985).**
5. **Palmer, D.K., O'Day, K., Le Trong, H., Charbonneau, H. & Margolis, R.L. Purification of the centromeric protein CENP-A and demonstration that it is a centromere specific histone. *Proc. Nat. Acad. Sci. (USA)* 88, 3734-3738 (1991).**
6. **Sullivan, K.F., Hechenberger, M. & Masri, K. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J. Cell Biol.* 127, 581-192 (1994).**
7. **Stoler, S., Keith, K.C., Curnick, K.E. & Fitzgerald-Hayes, M. A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes dev.* 9, 573-586 (1995).**
8. **Henikoff, S., Ahmad, K., Platero, J.S. & van Steensel, B. Heterochromatic deposition of centromeric histone H3-like proteins. *Proc. Natl. Acad. Sci. U.S.A.* 97, 716-721 (2000).**
9. **Foltz, D.R. et al. The human CENP-A centromeric nucleosome-associated complex. *Nat. Cell Biol.* 8, 458-469 (2006).**
10. **Blower, M.D., Sullivan, B.A. & Karpen, G.H. Conserved organization of centromeric chromatin in flies and humans. *Dev. Cell* 2, 319-330 (2002).**
11. **Sullivan, B.A. & Karpen, G.H. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* 11, 1076-1083 (2004).**
12. **Ribeiro, S.A. et al. A super-resolution map of the vertebrate kinetochore. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10484-10489 (2010).**