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

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

Published Version

Accessed
June 12, 2017 8:59:22 PM EDT

Citable Link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:11179848

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

(Article begins on next page)
Esperanto for histones: CENP-A, not CenH3, is the centromeric histone H3 variant


Received: 15 March 2013 / Revised: 17 March 2013 / Accepted: 17 March 2013 / Published online: 12 April 2013 © The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract The first centromeric protein identified in any species was CENP-A, a divergent member of the histone H3 family that was recognised by autoantibodies from patients with scleroderma-spectrum disease. It has recently been suggested to rename this protein CenH3. Here, we argue that the original name should be maintained both because it is the basis of a long established nomenclature for centromere proteins and because it avoids confusion due to the presence of canonical histone H3 at centromeres.
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 rationalise the nomenclature.

In 2012, an article entitled ‘A unified phylogeny-based nomenclature for histone variants’ appeared in...
the journal *Epigenetics and Chromatin* (Talbert et al. 2012). 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 et al. 1996), and a proposed standard nomenclature for the kinesin proteins (Lawrence et al. 2004). 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 human, we suggest, for the reasons detailed below, that the scientific community should maintain the original nomenclature.
(CENP-A) that was established for the centromeric histone H3 variant and avoid the usage of the misleading name (CenH3) proposed by Talbert et al. (2012).

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

It has now been suggested (Talbert et al. 2012) 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 et al. 1995), 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 et al. 2000). 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 on the basis of sequence similarity and was known from the outset to be the Drosophila variant of CENP-A.

The name proposed in Talbert et al., CenH3, adds an unnecessary layer of confusion that 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 et al. 2002; Foltz et al. 2006; Hori et al. 2008) to high-resolution light microscopy (Blower et al. 2002; Sullivan and Karpen 2004; Ribeiro et al. 2010) 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 et al. 2012) that make meaningful contacts with centromeric H3-containing chromatin are important for kinetochore assembly and function (Ohzeki et al. 2012). In fact, the interspersed H3 chromatin may represent a distinct chromosome domain, as it is post-translationally modified in a pattern that is distinct from both canonical heterochromatin and euchromatin (Sullivan and Karpen 2004). Recognising 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 chromatin-oriented authors of the Epigenetics and Chromatin article (Talbert et al. 2012) to rename CENP-A as CenH3 does not take into account the extensive preceding literature on centromeres or kinetochores, or the scientific confusion raised by such a change. It is notable 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 consulted concerning the Epigenetics and Chromatin nomenclature proposal. As systems and other forms of integrative biology become increasingly prevalent and communities that do not normally interact are brought into contact (and potential conflict), other nomenclature issues such as this will arise when the same protein means different things to diverse groups of scientists. Thus, the importance of cross-communication between communities and respect for precedence in naming (in this case, the precedence of the well-established CENP nomenclature) may actually increase over the next few years.

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 WCE and RCA are Principal Research Fellows of the Wellcome Trust [grant numbers 073915 and 095021, respectively]. The Wellcome Trust Centre for Cell Biology is supported by core grant numbers 077707 and 092076. BEB is funded by NIH R01-GM082989, a Career Award in the Biomedical Sciences from the Burroughs Wellcome Fund and a Rita Allen Foundation Scholar Award.
KB is funded by grant R37 GM32238. IMC is funded by NIH grant GM088313. DWC is an Investigator of the Ludwig Institute, whose centromere work is supported by NIH grant GM074150. GPC thanks the National Science Foundation (MCB-1121563) for support. JGD is supported by NIH R01GM088371. TF is supported by the Cabinet Office, Government of Japan, throug its ‘Funding Program for Next Generation World-Leading Researchers’ (LS122). RG is funded by an EMBO Installation Grant. DWG is funded by the European Community’s Seventh Framework Programme FP7/2007-2013 under grant agreements no. 241548 (MitoSys) and no. 258068 (Systems Microscopy) and from an ERC Starting Grant (no. 281198). DMG is supported by Cancer Research UK and the MRC. GIG is funded by NIH/NIGMS R01-50412, the Oklahoma Center for Advancement of Science and Technology and the McCasland Foundation. PH is funded by Seventh Framework Programme grant BioSynCen from the European Research Council. TH is funded by Research Grants from the Japanese Society for the Promotion of Science (JSPS) and the Ministry of Education, Culture, Sports, Science and Technology in Japan (MEXT). LETJ is funded by FCT grants BIA-BCM/100557/2008, BIAPRO/100537/2008 and an EMBO installation grant. GK is funded by NIH grant 5R01GM61169. GJPLK is supported by grants from the ERC (KINSIGN) and the Netherlands Organisation for Scientific Research (NWO-Vici-016.130.661). MAL is supported by NIH grant GM083988. AWdF is supported by grants from the Dutch Cancer Society (UU2009-4311, UU2011-5134) and the Netherlands Organisation for Scientific Research (NWO-Vici-91812610). AL is supported by the Spanish Ministry of Economy (grants CSD2007-015 and PTDC/SAU-ONC/112917/2009 from FCT (COMPETE/FEDER), the Human Frontier Science Program and the 7th framework program grant PRECISE from the European Research Council. ADM is funded by BBSRC grant (BB/I021353/1). RLM is funded by NIH grant R01GM088716. HM is supported in part by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, and the Kazusa DNA Research Institute Foundation. PSM is supported by a grant from the CCSRI (700824). BGM is funded by NSF award number 1024973. PM is funded by the Swiss National Foundation, University of Geneva and the Foundation Louis-Jeantet. RJO is supported by the NSF. EDS is funded by NIH grant R01GM074728. PTS is funded by NIH grant R01GM063045. BS is funded by NIH R01 GM098500 and March of Dimes 6-FY10-294. KFS is funded by Science Foundation Ireland grant 12/IA/1370. CES is funded by grants from the Fundação para Ciência e Tecnologia of Portugal, COMPETE and ON2. JRS is funded by BBSRC grant BB/H013024. CEW is funded by NIH grant GM059618. LW is funded by NIH grant GM69429. TY is funded by NIH grant GM083988, Core Grant CA06927, and an appropriation from the Commonwealth of Pennsylvania.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References


