



OPINION ARTICLE

REVISED Changing the name of the NBPF/DUF1220 domain to the Olduvai domain [version 2; referees: 3 approved]

Previously titled: A proposal to change the name of the NBPF/DUF1220 domain to the Olduvai domain

James M. Sikela ¹, Frans van Roy^{2,3}¹Department of Biochemistry and Molecular Genetics, Human Medical Genetics and Neuroscience Programs, University of Colorado School of Medicine, Aurora, CO, 80045, USA²Department of Biomedical Molecular Biology, Ghent University, Ghent, 9052, Belgium³VIB-UGent Center for Inflammation Research, Ghent, 9052, Belgium**v2** First published: 28 Dec 2017, 6:2185 (doi: [10.12688/f1000research.13586.1](https://doi.org/10.12688/f1000research.13586.1))
Latest published: 17 Jul 2018, 6:2185 (doi: [10.12688/f1000research.13586.2](https://doi.org/10.12688/f1000research.13586.2))**Abstract**

We are jointly proposing a new name for a protein domain of approximately 65 amino acids that has been previously termed NBPF or DUF1220. Our two labs independently reported the initial studies of this domain, which is encoded almost entirely within a single gene family. The name Neuroblastoma Breakpoint Family (*NBPF*) was applied to this gene family when the first identified member of the family was found to be interrupted in an individual with neuroblastoma.

Prior to this discovery, the Pfam database had termed the domain DUF1220, denoting it as one of many protein domains of unknown function. It has been Pfam's intention to use "DUF" nomenclature to serve only as a temporary placeholder until more appropriate names are proposed based on research findings.








We believe that additional studies of this domain, primarily from our laboratories over the past 10 years, have resulted in furthering our understanding of these sequences to the point where proposing a new name for this domain is warranted. Because of considerable data linking the domain to human-specific evolution, brain expansion and cognition, we believe a name reflecting these findings would be appropriate. With this in mind, we have chosen to name the domain (and the repeat that encodes it) Olduvai. The gene family will remain as *NBPF* for now. The primary domain subtypes will retain their previously assigned names (e.g. CON1-3; HLS1-3), and the three-domain block that expanded dramatically in the human lineage will be termed the Olduvai triplet.



The new name refers to Olduvai Gorge, which is a site in East Africa that has been the source of major anthropological discoveries in the early-mid 1900's. We also chose the name as a tribute to the scientists who made important contributions to the early studies of human origins and our African genesis.

Keywords

DUF1220, NBPF, protein domain, human brain evolution, copy number, gene duplication, genome evolution, Olduvai Gorge

Open Peer ReviewReferee Status: 

	Invited Referees		
	1	2	3
REVISED			
version 2	report	report	
published 17 Jul 2018			
version 1			
published 28 Dec 2017	report	report	report

- 1 **Alex Bateman** , European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), UK
- 2 **Christian J.A. Sigrist**, Swiss Institute of Bioinformatics (SIB), Switzerland
- 3 **Elsbeth A. Bruford** , European Bioinformatics Institute (EMBL-EBI), UK

Discuss this article

Comments (0)

Corresponding authors: James M. Sikela (james.sikela@ucdenver.edu), Frans van Roy (frans.vanroy@ugent.be)

Author roles: **Sikela JM:** Conceptualization, Writing – Original Draft Preparation, Writing – Review & Editing; **van Roy F:** Conceptualization, Writing – Review & Editing

Competing interests: JMS is founder of GATC Science, LLC, a biotech company involved in genomic research.

Grant information: JMS is funded by NIH R01 grant MH108684. FVR is supported by the Foundation Against Cancer – Belgium, the Research Foundation – Flanders (FWO-Vlaanderen), and the Queen Elisabeth Medical Foundation (G.S.K.E.), Belgium.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Sikela JM and van Roy F. This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Data associated with the article are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

How to cite this article: Sikela JM and van Roy F. **Changing the name of the NBPF/DUF1220 domain to the Olduvai domain [version 2; referees: 3 approved]** *F1000Research* 2018, **6**:2185 (doi: [10.12688/f1000research.13586.2](https://doi.org/10.12688/f1000research.13586.2))

First published: 28 Dec 2017, **6**:2185 (doi: [10.12688/f1000research.13586.1](https://doi.org/10.12688/f1000research.13586.1))

REVISED Amendments from Version 1

We appreciate the helpful comments of reviewers and have modified our original paper to address specific reviewer suggestions. Specifically we have:

1. Altered the title to now read "Changing the name of the NBPF/DUF1220 domain to the Olduvai domain".
2. After "propose that, for now, the gene family name should remain *NBPF*," we added the following sentence: "It is also worth noting that there is an Olduvai domain that is separate from those encoded by *NBPF* genes: a single, likely ancestral, copy of the domain is found within the *PDE4DIP/myomegalin* gene."
3. Changed PFAM to Pfam
4. (after "our African genesis".) we added the following sentence. "While we are aware that naming a protein domain after a geographic location is unusual, we believe the name Olduvai is more than a place. Rather it has also come to be thought of as a symbol of the efforts of our species to understand itself."
5. Added the name Mark Johnston to the Acknowledgments.

See referee reports

Protein domains are portable units within proteins that can serve important biological functions. They have been implicated in a broad range of key biological phenomena, including development, disease and evolution^{1,2}. Here we jointly propose a new name for a protein domain of approximately 65 amino acids that has been previously termed NBPF³ or DUF1220⁴. Our two labs independently reported the initial studies of this domain, which is encoded almost entirely within a single gene family. The name Neuroblastoma Breakpoint Family (*NBPF*) was applied to this gene family when the first identified member of the family was found to be interrupted in an individual with neuroblastoma^{3,5}. Prior to this discovery, the Pfam database had termed the domain DUF1220, denoting it as one of many protein domains of unknown function⁶. It has been Pfam's intention to use "DUF" nomenclature to serve only as a temporary placeholder until more appropriate names are proposed based on research findings. We believe that additional studies of this domain, primarily from our laboratories over the past 10 years, have resulted in furthering our understanding of these sequences to the point where proposing a new name for this domain is warranted.

Key findings relevant to assigning a name to the domain are as follows:

1. **The domain has been repeatedly linked with human-specific evolution.** The haploid human genome is estimated to encode approximately 300 copies of the NBPF/DUF1220 domain, while the copy number for other species is substantially lower: great apes 90–120, monkeys 30–40, and all other mammalian species 1–9^{7,8}. The increase in humans (at least 165 additional human copies) represents the largest human lineage-specific copy number increase of any coding region in the genome. These findings, involving the copy number of a protein coding domain, provide strong support for an involvement in human-specific evolution.

2. **The domain has been linked with human brain evolution and cognitive function.** Over the past 10 years we have published several papers on NBPF/DUF1220 protein domains and the *NBPF* gene family. These have implicated the copy number of the domain in human brain evolution^{7,9–11}, brain size-related phenotypes^{7,9–11}, brain disorders (autism/schizophrenia/micro- and macrocephaly)^{9,12–14}, and measures of cognitive function^{13,15}. Also, our finding of a robust linear association between NBPF/DUF1220 copy number and brain size across primate species was confirmed by an independent study¹⁶.

Given the above research findings, a new name for this protein domain that is related to human-specific evolution would be appropriate. We believe a name that would do this is "Olduvai" (ohl'-du-vi) (when necessary it can be abbreviated as "ODV"). This name refers to Olduvai Gorge, which is located in the rift valley of Eastern Africa. Olduvai has been the site of key paleoanthropological discoveries related to human origins and has been called "the Cradle of Mankind", and "the Grand Canyon of Human Evolution"¹⁷. Deposits at the gorge are estimated to cover a time span from 2.1 million to 15,000 years ago, and the fossil remains that have been identified there are thought to represent more than 60 hominins (members of the human lineage). These findings are believed to constitute the most continuous known record of human evolution over the past 2 million years, and the longest known archaeological record of the development of stone-tool industries. Olduvai Gorge was designated part of a UNESCO World Heritage site in 1979¹⁸. Just as the protein domain appears to be important to human-specific evolution, so too, Olduvai Gorge has provided key insights into human's evolutionary origin. We believe both are central to the story of what made us human.

Finally, we have also chosen this name because it reflects an appreciation for the important contributions of the scientists who made major anthropological discoveries in Africa in the early/mid-20th century that stimulated further research into human origins and our African genesis. While we are aware that naming a protein domain after a geographic location is unusual, we believe the name Olduvai is more than a place. Rather it has also come to be thought of as a symbol of the efforts of our species to understand itself.

While we believe that the domain and repeat should now be called "Olduvai", we also propose that, for now, the gene family name should remain *NBPF*. It is also worth noting that there is an Olduvai domain that is separate from those encoded by *NBPF* genes: a single, likely ancestral, copy of the domain is found within the *PDE4DIP/myomegalin* gene. In summary, the NBPF domain and DUF1220 domain will be termed the Olduvai domain, and the NBPF repeat and DUF1220 repeat will be termed the Olduvai repeat. The primary gene family that encodes these sequences will continue to be called *NBPF*, and the primary domain subtypes will retain established nomenclature (CON1-3, HLS1-3)⁸. The three-domain block, composed of HLS1, HLS2 and HLS3 subtypes, that is tandemly repeated

within several human *NBPF* genes and responsible for the great majority of additional human copies of the domain, will be called the Olduvai triplet.

The Olduvai domain hyper-amplification in the human lineage was one of the most extreme and rapid copy number expansions in the human genome, and we look forward to additional studies that may provide further insights into the role this protein domain family plays in human disease and evolution.

Competing interests

JMS is founder of GATC Science, LLC, a biotech company involved in genomic research.

Grant information

JMS is funded by NIH R01 grant MH108684. FVR is supported by the Foundation Against Cancer – Belgium, the Research Foundation – Flanders (FWO-Vlaanderen), and the Queen Elisabeth Medical Foundation (G.S.K.E.), Belgium.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

The authors thank Jonathon Davis, Veronica Searles Quick, Ilea Heft, Laura Dumas, Mark Johnston, Vanessa Andries and Karl Vandepoele for constructive discussions.

References

- Mark M, Rijli FM, Chambon P: **Homeobox genes in embryogenesis and pathogenesis.** *Pediatr Res.* 1997; **42**(4): 421–429.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Li WH: **Molecular Evolution.** Sinauer Associates, Sunderland, Massachusetts; 1997.
- Vandepoele K, Van Roy N, Staes K, *et al.*: **A novel gene family NBPF: Intricate structure generated by gene duplications during primate evolution.** *Mol Biol Evol.* 2005; **22**(11): 2265–2274.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Popesco MC, Maclaren EJ, Hopkins J, *et al.*: **Human lineage-specific amplification, selection, and neuronal expression of DUF1220 domains.** *Science.* 2006; **313**(5791): 1304–1307.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Vandepoele K, Andries V, Van Roy N, *et al.*: **A constitutional translocation t(1;17)(p36.2;q11.2) in a neuroblastoma patient disrupts the human NBPF1 and ACCN1 genes.** *Bielinsky A-K, ed. PLoS One.* 2008; **3**(5): e2207.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Bateman A, Coin L, Durbin R, *et al.*: **The Pfam protein families database.** *Nucleic Acids Res.* 2004; **32**(Database issue): D138–41.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dumas L, Sikela JM: **DUF1220 domains, cognitive disease, and human brain evolution.** *Cold Spring Harb Symp Quant Biol.* 2009; **74**: 375–382.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- O'Bleness MS, Dickens CM, Dumas LJ, *et al.*: **Evolutionary history and genome organization of DUF1220 protein domains.** *G3 (Bethesda, Md).* *Genes/Genomes/Genetics.* 2012; **2**(9): 977–986.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dumas LJ, O'Bleness MS, Davis JM, *et al.*: **DUF1220-domain copy number implicated in human brain-size pathology and evolution.** *Am J Hum Genet.* 2012; **91**(3): 444–454.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Keeney JG, Dumas L, Sikela JM: **The case for DUF1220 domain dosage as a primary contributor to anthropoid brain expansion.** *Front Hum Neurosci.* 2014; **8**: 427.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Keeney JG, Davis JM, Siegenthaler J, *et al.*: **DUF1220 protein domains drive proliferation in human neural stem cells and are associated with increased cortical volume in anthropoid primates.** *Brain Struct Funct.* 2015; **220**(5): 3053–60.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Davis JM, Searles VB, Anderson N, *et al.*: **DUF1220 dosage is linearly associated with increasing severity of the three primary symptoms of autism.** *PLoS Genet.* 2014; **10**(3): e1004241.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Davis JM, Searles Quick VB, Sikela JM: **Replicated linear association between DUF1220 copy number and severity of social impairment in autism.** *Hum Genet.* 2015; **134**(6): 569–575.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Searles Quick V, Davis JM, Olincy A, *et al.*: **DUF1220 copy number is associated with schizophrenia risk and severity: Implications for understanding autism and schizophrenia as related diseases.** *Transl Psychiatry.* 2015; **5**(12): e697.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Davis JM, Searles VB, Anderson N, *et al.*: **DUF1220 copy number is linearly associated with increased cognitive function as measured by total IQ and mathematical aptitude scores.** *Hum Genet.* 2015; **134**(1): 67–75.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Zimmer F, Montgomery SH: **Phylogenetic analysis supports a link between DUF1220 domain number and primate brain expansion.** *Genome Biol Evol.* 2015; **7**(8): 2083–2088.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ardrey R: **African Genesis.** Collins, London, UK; 1961.
[Reference Source](#)
- Olduvai Gorge. In: *Encyclopedia Britannica.*
[Reference Source](#)

Open Peer Review

Current Referee Status:



Version 2

Referee Report 25 July 2018

doi:[10.5256/f1000research.16945.r36115](https://doi.org/10.5256/f1000research.16945.r36115)



Christian J.A. Sigrist

Swiss Institute of Bioinformatics (SIB), Geneva, Switzerland

The authors have taken into account the reviewers remarks. I hope that this paper will encourage the NBPF and myomegalin communities to join efforts to unravel the function of the Olduvai domain and its role in the emergence of mankind.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 17 July 2018

doi:[10.5256/f1000research.16945.r36117](https://doi.org/10.5256/f1000research.16945.r36117)



Alex Bateman 

European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK

I approve the new version of the manuscript.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Referee Report 18 January 2018

doi:[10.5256/f1000research.14757.r29429](https://doi.org/10.5256/f1000research.14757.r29429)



Elsbeth A. Bruford 

HUGO Gene Nomenclature Committee (HGNC), European Bioinformatics Institute (EMBL-EBI), Hinxton, UK

This brief opinion article describes the authors' proposal to rename a domain found in a discrete group of proteins, which was originally named DUF1220 (domain of unknown function 1220) and subsequently reassigned as NBPF. The NBPF nomenclature referenced the fact that this domain is found in proteins encoded by a family of genes initially reported as the "Neuroblastoma BreakPoint Family"; however, as human gene nomenclature tries to move away from directly referencing specific phenotypes and conditions, the reference to neuroblastoma has been removed from the names for the NBPF genes, and they are now simply the "NBPF family".

Interestingly, as well as this family the only other protein this domain is found in is the large "myomegalin" protein in mammals, encoded by the PDE4DIP gene. I think the fact that this is not mentioned anywhere in the article is perhaps an oversight on the part of the authors which could be remedied.

It should be noted that the authors have already contacted PFAM to request this change and HGNC were also consulted and agreed this update would not create problems from a gene naming perspective. As we said then, we think it would be more informative for the new domain name to reflect functional information as opposed to something more esoteric, but the authors are not advocating any parallel change in the NBPF gene family nomenclature which is quite widely used for this family. A parallel change to the gene names would perhaps be more difficult to justify. Therefore we told the authors that we see no issue with their proposal.

Importantly, this article represents the collaborative effort of two independent groups, who easily could be viewed as "rivals". Therefore this article effectively shows that such collaborations are possible and beneficial to the scientific community at large; this kind of activity should be encouraged wherever possible and I applaud the authors for making the effort to work productively together.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 17 January 2018

doi:[10.5256/f1000research.14757.r29374](https://doi.org/10.5256/f1000research.14757.r29374)



Christian J.A. Sigrist

Swiss Institute of Bioinformatics (SIB), Geneva, Switzerland

This paper proposes to change the name of a protein domain from NBPF (Neuroblastoma Breakpoint Family) to Olduvai, arguing that this new name illustrates the fact that the domain copy number has been linked with human evolution. Although, it is in principle not recommended to multiply the names for a domain, the fact that the change is submitted by two major players in the field makes the proposal acceptable as it should rapidly spread in future works on this promising domain. In addition, the limited distribution of the domain should facilitate its diffusion. However, it should be noted that although having a very limited distribution found in roughly two protein families, i.e. NBPF and myomegalin (*PDE4DIP*), neither of the current names, NBPF or DUF1220, is commonly used by the myomegalin community. In this respect, I think that a paragraph presenting the myomegalin family and discussing the presence of a single Olduvai “repeat” amongst its members should be added prior to indexing.

The previous name, NBPF, was derived from a breakpoint in the first member identified and was also applied to the protein family. In the future, the use of the new name Olduvai only for the repeat will avoid the confusion between NBPF designating both the protein family and the repeat.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Referee Report 02 January 2018

doi:10.5256/f1000research.14757.r29376



Alex Bateman 

European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK

This manuscript describes a straightforward proposal to rename the NBPF/DUF1220 domain to the Olduvai domain. This is welcome because the domain has been referred to by two different names in the past. The authors are the major players in the field, who have used the divergent nomenclature and thus will be able to ensure that the proposed naming does become the standard. The authors had already

contacted me to ask about having the domain renamed within the Pfam database, which has now been done (although not visible on the public website yet). Perhaps calling this paper a proposal is a bit weak. The authors should consider renaming the article to “Changing the name of the NBPF/DUF1220 domain to the Olduvai domain”.

There are no agreed upon standard naming schemes for protein domains or families. To date the most popular schemes are based on naming after one or more proteins that contain the domain, sometimes based upon the initial letters of these members. A few domains are named after scientists. I don't recall any domains or families named after places. However, the renaming makes sense within the context of the evolution of this domain and human evolution.

I think it would be worth mentioning the reasons that the NBPF nomenclature is retained for the gene family. I understand this is to fit with the HGNC guidelines.

Minor points

Please change PFAM to Pfam throughout the text.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research