

RESEARCH ARTICLE

Open Access



# The psoriasis-associated deletion of late cornified envelope genes *LCE3B* and *LCE3C* has been maintained under balancing selection since Human Denisovan divergence

Petar Pajic<sup>†</sup>, Yen-Lung Lin<sup>†</sup>, Duo Xu and Omer Gokcumen<sup>\*</sup> 

## Abstract

**Background:** A common, 32kb deletion of *LCE3B* and *LCE3C* genes is strongly associated with psoriasis. We recently found that this deletion is ancient, predating Human-Denisovan divergence. However, it was not clear why negative selection has not removed this deletion from the population.

**Results:** Here, we show that the haplotype block that harbors the deletion (i) retains high allele frequency among extant and ancient human populations; (ii) harbors unusually high nucleotide variation ( $\pi$ ,  $P < 4.1 \times 10^{-3}$ ); (iii) contains an excess of intermediate frequency variants (Tajima's  $D$ ,  $P < 3.9 \times 10^{-3}$ ); and (iv) has an unusually long time to coalescence to the most recent common ancestor (T<sub>Sel</sub>, 0.1 quantile).

**Conclusions:** Our results are most parsimonious with the scenario where the *LCE3BC* deletion has evolved under balancing selection in humans. More broadly, this is consistent with the hypothesis that a balance between autoimmunity and natural vaccination through increased exposure to pathogens maintains this deletion in humans.

**Keywords:** Copy number variation, Genomic structural variants, Atopic dermatitis, HLA, Defensins, Neanderthal, *LCE3A*, Human evolution

## Background

Genomic structural variants (SVs), which are deletions, duplications, inversions and translocations of genomic segments, account for the majority of variable base pairs in primate genomes [1, 2]. Because of their sheer size, SVs can have strong effects on gene function and regulation if they overlap with protein-coding (eg, [3]) or regulatory sequences (eg, [4]). Indeed, several studies have revealed important roles that SVs play in human evolution [2, 5] and adaptation [6, 7].

Disruption of a gene's function by deletion of its coding sequence likely reduces fitness and predisposes humans to several genetic disorders (eg, [8–10]).

Consistent with this notion, deletion variants among humans are distributed significantly away from coding sequences [11] and most exonic deletions are found in very low frequencies in human populations [12]. In a recent study, we searched for unusually old deletion variants that affect coding sequences. Specifically, we identified exonic deletions that evolved before Human-Neanderthal divergence (>500-1,000KYA) [13]. We surmised that it is unlikely for a loss-of-function gene deletion to be maintained for this long, especially under negative selection. Thus, we hypothesized that a number of these ancient deletion variants have been evolving under balancing selection.

Balancing selection has enjoyed a renewed interest in the evolutionary genomics community. In its most basic form, balancing selection can be thought of as the combination of adaptive forces that maintain variation longer

\* Correspondence: omergokc@buffalo.edu

<sup>†</sup>Equal contributors

Department of Biological Sciences, University at Buffalo, Cooke 639, Buffalo, NY 14260, USA



than expected under neutrality [14, 15]. Based on the analyses of recently available human and nonhuman primate genomes, several variants have been shown to be evolving under long-term balancing selection in the human-chimpanzee lineage [16, 17]. In addition, multiple instances of balancing selection within the human gene pool have been reported in the last decade [18–23].

In this paper, we investigate the evolution of an ancient gene deletion (*LCE3BC* deletion). This ~32kb deletion variant overlaps 2 genes, *LCE3B* and *LCE3C*, which are both involved in skin tissue repair. We recently showed that this deletion variant is derived in the *Homo* lineage and that the deletion is present in the Denisovan genome, but absent in the Neanderthal genome [13]. In the same study, we were able to rule out archaic introgression and concluded that incomplete lineage sorting best explains the observed allele sharing at this locus. The deletion is very common in humans, reaching up to 70% in some European populations. Moreover, this deletion has been strongly associated with psoriasis susceptibility, with odds ratios ranging from 1.3 (The Italian population) to 1.9 (The Chinese population) [24–27]. However, the adaptive reasons for why this deletion remains in the population are unknown.

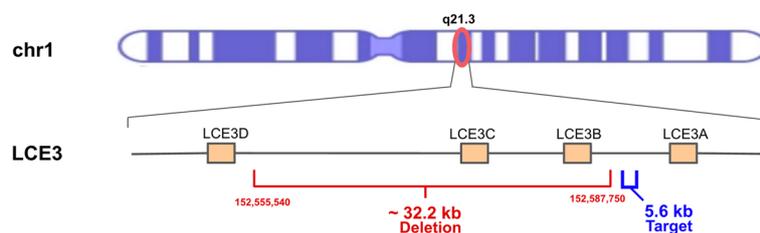
## Results

The *LCE3BC* deletion spans slightly more than 32kb in the human reference genome and overlaps with two conserved, protein coding genes, *LCE3B* and *LCE3C* (Fig. 1). We manually confirmed that this deletion has been shared with Denisovans, but not with Neanderthals (Additional file 1: Figure S1). We then determined single nucleotide variants that are in high linkage disequilibrium ( $R^2 > 0.9$ ) with the *LCE3BC* deletion among 2504 human genomes independent of population ancestry [28] (Additional file 2: Table S1). Using this dataset, we identified a haplotype block that comprises the *LCE3BC* deletion and its flanking sequences extending to 6.6kb downstream and 5.5kb upstream (Additional file 1: Figure S2).

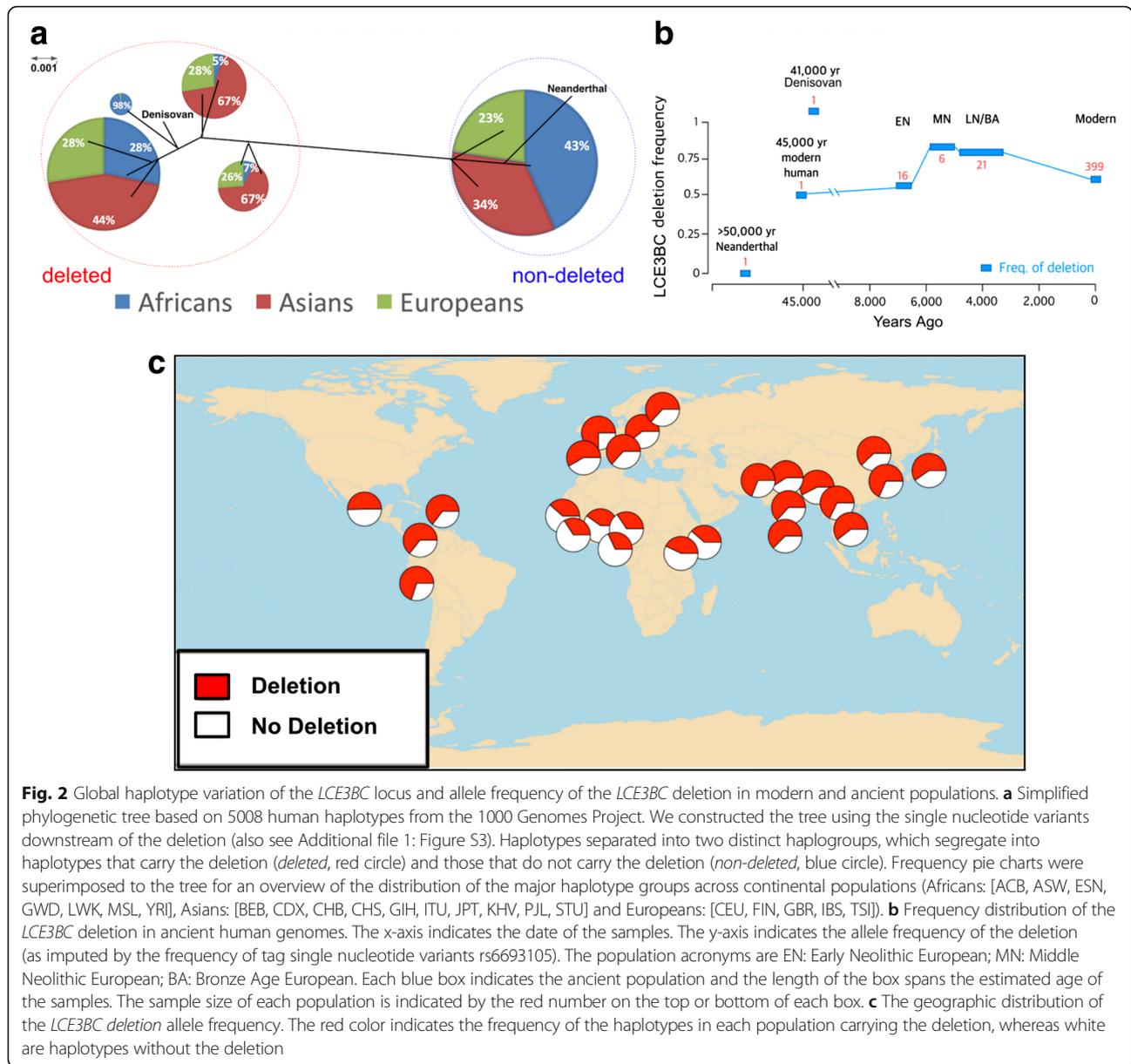
We constructed a maximum likelihood tree of the variation in the region downstream of the *LCE3BC* deletion (target region) among 5008 modern human haplotypes, as well as Denisovan, Altai Neanderthal, Chimpanzee and Rhesus Macaque haplotypes (Fig. 2a, Additional file 1: Figure S3). We found a clear separation of two haplogroups (named hereon *deleted* and *non-deleted*) with perfect segregation of haplotypes carrying the deletion and those that do not. Furthermore, the Denisovan haplotype, which carries the deletion, clusters with the *deleted* human haplotypes, while the Neanderthal haplotype, which does not carry the deletion, clusters with the *non-deleted* haplotypes (Fig. 2a). This observation provides further support that the *LCE3BC* deletion evolved before Human-Denisovan split and remains variable since then.

We then traced back the allele frequency of rs6693105, which tags the *LCE3BC* deletion in all modern humans ( $R^2 = 0.95$ ), in recently available ancient genomes (Additional file 3: Table S2). We found that the genome of the 45,000 year old individual from Central Asia (Altai) [29] is heterozygous for the deletion allele (Fig. 2b). In fact, we were able to directly confirm the loss of read-depth due to heterozygous deletion in this sample (Additional file 1: Figure S1). We also found that the deletion haplotype has always been found in high frequency in ancient European populations (Fig. 2b). The *LCE3BC* deletion is very common in extant human populations as well, reaching major allele status in all non-African populations (Fig. 2c).

To explain the evolutionary forces that shape this ancient (older than Human-Denisovan divergence) and very common allele (>50% frequency in most populations), we consider three scenarios. First, it is plausible that the *LCE3BC* deletion evolved under neutrality (eg, [30]). Indeed, in a recent study, we provided evidence that a neighboring gene, *FLG*, harbors neutrally evolving common loss-of-function variants [31]. In this scenario, we would expect the haplotype block carrying the deletion variant to show no significant deviation in tests of



**Fig. 1** Genomic location of the *LCE3BC* deletion. *Top*: Location of the *LCE3BC* deletion on chromosome 1q21.3 (red circle). *Below*: The zoomed-in look at the region harboring the *LCE3BC* deletion. The red bar shows where the deletion occurs (Hg19, chr1:152,555,540–152,587,750). The blue bar represents the 5.6 kb “target” (chr1: 152,587,904–152,593,549) sequence that we used to conduct the majority of the population genetics analyses in this study



neutrality, when compared to neutral regions of the genome. Second, we considered positive selection, which would increase the frequency of this variant in human populations. Such cases were reported for other loss of function variants [32]. In this scenario, we expect increased population divergence ( $F_{ST}$ ) [33] and deviation from expected homozygosity (delta integrated homozygosity score ( $\Delta iHH$ )) [34], but reduced nucleotide diversity ( $\pi$ ). Third, we considered balancing selection, where we expect high  $\pi$  and a deviation from the expected frequency site spectrum (eg, positive values of Tajima's D).

To test the above hypotheses, we compared different population genetics statistics for the *LCE3BC* haplotype

block to those calculated for neutrally evolving regions as defined by Arbiza et al. [35] (Additional file 4: Table S3). We found no significant increase of  $\Delta iHH$  or  $F_{ST}$  in the *LCE3BC* haplotype block as compared to neutral regions (Additional file 1: Figure S4). In contrast, we observed that nucleotide diversity ( $\pi$ ) in the *LCE3BC* haplotype block is at least 2 fold higher than neutral regions (Wilcoxon Test, CEU -  $P < 4.1 \times 10^{-3}$ ; CHB -  $P < 7.7 \times 10^{-4}$ ; YRI -  $P < 2.2 \times 10^{-3}$ ) (Table 1, Fig. 3a) Similarly, Tajima's D statistics in the *LCE3BC* haplotype block was significantly higher than neutral regions in *all* populations tested (Wilcoxon Test, CEU -  $P < 1.5 \times 10^{-3}$ ; CHB -  $P < 1.5 \times 10^{-3}$ ; YRI -  $P < 3.9 \times 10^{-3}$ ) (Table 1,

**Table 1** Summarized mean and standard deviation values for Tajima’s D and Pi for different population and for the target and neutral regions, as well as regions adjacent to comparable nonexonic ancient deletions

Population	Region	Tajima’s D	Pi
CEU	Target	2.65+/-0.13	45.48 +/-5.62
	Neutral	0.41+/-0.87	29.65 +/-12.56
	Non-Exon	0.49 +/-1.05	31.32 +/-35.12
CHB	Target	2.97+/-0.13	48.00 +/-8.67
	Neutral	0.49 +/-0.96	27.73 +/-12.53
	Non-Exon	0.46 +/-1.19	29.21 +/-35.38
YRI	Target	0.66+/-0.16	53.58 +/-8.14
	Neutral	-0.27+/-0.54	39.30 +/-13.21
	Non-Exon	-0.17 +/-0.70	40.29 +/-31.30

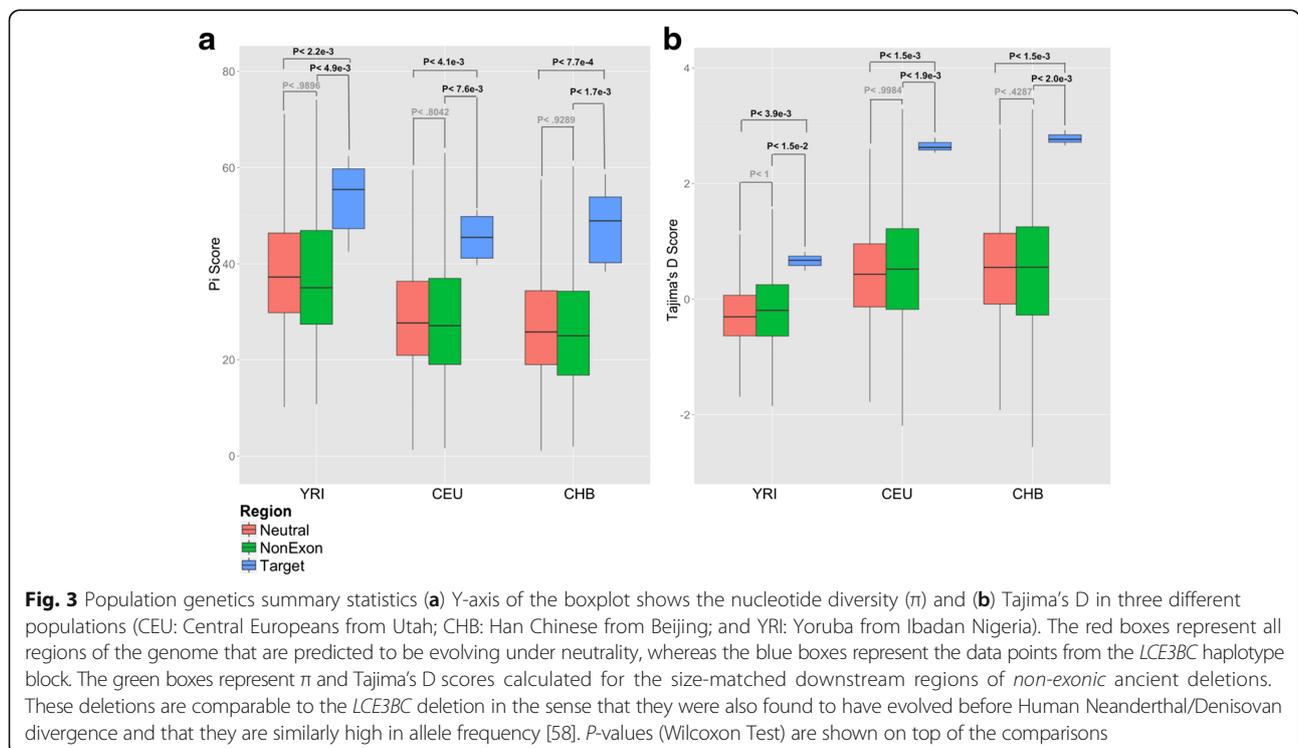
Fig. 3b). Tajima’s D compares the pairwise differences between haplotypes and the number of segregating sites [36]. In practice, positive values of Tajima’s D indicate an excess of intermediate frequency variants, which is a hallmark of balancing selection. Last but not least, we found that the region 10kb downstream of the *LCE3BC* deletion is within the 10<sup>th</sup> percentile of genome-wide distribution of pairwise time to most recent common ancestor as measured by Tsel method [37].

These standard tests of neutrality can be affected by the mutation rate, as well demographic history [38]. One specific worry was whether preselecting a known

“ancient” deletion that is shared with an archaic hominin species may bias our results. For example, it is possible that increased  $\pi$  may be due to the fact that the variation in this locus is older than many other parts of the genome. To address this, we compared Tajima’s D and  $\pi$  observed for the *LCE3BC* haplotype block with those calculated for size matched regions downstream of 340 other “ancient”, non-exonic deletions. These deletions were previously found to be shared with Neanderthals and/or Denisovans due to incomplete lineage sorting and also present in high allele frequency in contemporary human populations. As such, they are comparable to the *LCE3BC* deletion both for their age and allele frequency. This analysis confirmed our previous observations that  $\pi$  (Wilcoxon Test, CEU -  $P < 7.6 \times 10^{-3}$ ; CHB -  $P < 1.7 \times 10^{-3}$ ; YRI -  $P < 4.9 \times 10^{-3}$ ) and Tajima’s D (Wilcoxon Test, CEU -  $P < 1.9 \times 10^{-3}$ ; CHB -  $P < 2.0 \times 10^{-3}$ ; YRI -  $P < 1.5 \times 10^{-2}$ ) calculated for the *LCE3BC* haplotype block is significantly higher than those calculated for other ancient deletion haplotype blocks (Fig. 3a and b). We also considered the potential impact of repetitive sequences to variant calling quality. As such, we confirmed all our main results, omitting variants coinciding with repetitive sequences (Additional file 1: Figures S5–S7).

**Discussion**

In this study, we analyzed the haplotype block specifically because it harbors an ancient, exonic and disease-



associated deletion variant. Our analyses of the haplotypic variation harboring this deletion best fits a model where the *LCE3BC* deletion has been maintained under balancing selection in human populations. What remains an open question is the adaptive pressure(s) on the *LCE3B* and *LCE3C* gene functions at the organismal and ecological levels.

The phenotypic impact of the *LCE3BC* deletion has been discussed extensively, especially within the context of psoriasis biology (reviewed in [39]). Briefly, *LCE3B* and *LCE3C* coding sequences, which remain highly conserved across mammals (Additional file 1: Figure S8), are both active in skin barrier repair. As such, their expression is mostly confined to injured skin [40] and deletion of these genes likely leads to inefficient skin barrier repair. In addition, a likely regulatory region is also eliminated by the deletion [41] (Additional file 1: Figure S8). It also appears that the deletion haplotype leads to a significant increase in *LCE3A* expression in sun-exposed skin (Additional file 1: Figure S9), which may be a partial compensatory response to loss of *LCE3B* and *LCE3C* activity.

It is unknown how the lack of *LCE3B* and *LCE3C* activity due to their deletion leads to psoriasis susceptibility. Variants that are related to skin structure (eg, *LCE3BC* deletion [24]) and variants related to immune function [42–44] are independently associated with psoriasis. Moreover, epistatic interactions between *HLA-C\*06* and *LCE3BC* loci were reported within the context of psoriasis [24, 25]. Briefly, it appears that the *LCE3BC* deletion leads to slower repair of the epidermal barrier, which in turn leads to increased exposure to environmental antigens and pathogens. The higher level of exposure consequently leads to higher activity of immune elements, and occasionally pathological autoimmune response.

Based on the above-described mechanism, Bergboer et al. [39] hypothesized that the *LCE3BC* deletion would be favored to increase the effectiveness of the acquired immunity system (ie, natural vaccination), with the drawback of increasing the susceptibility to autoimmune disorders. In fact, it is plausible that the effect of the *LCE3BC* deletion on the immune system may be more than skin-deep, as the deletion was also associated with more systemic autoimmune disorders, such as psoriatic arthritis [45] and lupus [46]. Our findings presented here are concordant with this hypothesis.

## Conclusion

The recent availability of high quality whole genome data at the population level provides novel opportunities to investigate complex evolutionary forces that shape disease susceptibility loci without ascertainment bias. Using such an approach, we provide multiple lines of

evidence that a common, 32kb deletion strongly associated with psoriasis has evolved under balancing selection in the human lineage [47]. Our study presents empirical evidence that balancing selection on the *LCE3BC* deletion contributes to this very interesting dynamic. Our results will also contribute to the renewed discussion in the community on balancing selection maintaining advantageous diversity in human populations [17, 18, 22].

## Methods

This 5.6kb region was selected based on the linkage disequilibrium between the single nucleotide variants and the deletion ( $R^2 > 0.8$ ) (Additional file 2: Table S1). Therefore, we surmised that this region is the haplotype block that harbors the deletion and we expect that the characteristics of the genetic variation within this region would inform us with regards to the evolutionary forces that shape the deletion as well. To investigate genetic variation within the *LCE3BC* haplotype block, we utilized the 1000 Genomes Project dataset, which includes 2504 human genomes across 26 populations [28], multiple ancient genomes [29, 48–50], chimpanzee and rhesus macaque reference haplotypes [51, 52] (Additional file 2: Table S1 and Additional file 3: Table S2). To do this, we used a custom pipeline, which is available at [github](https://github.com/duoduoo/VCFtoTree) (<https://github.com/duoduoo/VCFtoTree>).

It has been shown that single nucleotide variation calling in next generation resequencing data may be affected in regions with duplicated or repeat-rich regions of the genome. We previously verified that all the single nucleotide variants that we used in our analyses have passed the 1000 Genomes quality filters (Quality score = 100) [11]. However, to further ensure that our analyses are not biased with repeated segments of the genome, we checked the target region where we conducted our analyses for presence of segmental duplications and other repetitive sequences. In addition, we used BLAT to check sequence uniqueness of this region. These analyses confirmed that there are no segmental duplications within our target region, but detected an L1 element (Fig. 1). Based on this search, we found an L1 element within the target region (Additional file 1: Figure S5, chr1:152,587,904–152,590,051). Since, it is plausible that this element can cause spurious single nucleotide variant calls, creating increased heterozygosity, we re-conducted our main analyses on the newly constructed target region (chr1:152,590,052–152,593,549), omitting the LINE element. The results confirm that our original conclusions remain unchanged, based on identical differentiation of haplogroups on the maximum likelihood tree (Additional file 1: Figure S6) and unchanged results for Tajima's  $D$  and  $\pi$  (Additional file 1: Figure S7).

We constructed a maximum likelihood tree using RAxML [53] with single nucleotide variants downstream of the deletion [Hg19-chr1: 152,587,904–152,593,549]. The alignments can be found as a supplementary file as well as on our website (gokcumenlab.org/data-and-codes/). We used Dendroscope [54] program for visualization. We constructed a rooted maximum likelihood tree using PhyML (HKY85 model) and we bootstrapped with 1000 replicates for branch support. We used values calculated by the 1000 Genomes Selection Browser [55] for comparing multiple population genetics parameters. To visually inspect the deletion variation in archaic humans, we used Integrated Genome Viewer [56]. We used ENCODE and Gtex databases to search for functional variants within the LCE3BC haplotype block. For compiling allele frequencies in ancient genomes, we used PLINK [57]. We used Python and R for custom bioinformatic and statistical analyses of the genetic variation. The R codes and processed datasets to replicate our core analyses are available as supplemental files (Additional File 5: File S1, Additional File 6: File S2) as well as on our website (<http://gokcumenlab.org/data-and-codes/>).

## Additional files

- Additional file 1:** Supplementary Figures. (PDF 1236 kb)
- Additional file 2: Table S1.** Linkage disequilibrium between LCE3 BC deletion and adjacent SNPs. (XLSX 64 kb)
- Additional file 3: Table S2.** rs6693105 (ie, linked with L CE3BC deletion) allele frequency in ancient human samples. (XLSX 39 kb)
- Additional file 4: Table S3.** Tajima's D, Pi, Fst, ΔiHH for Neutral, ancient and Target regions. (XLSX 7546 kb)
- Additional file 5: File S1.** Codes. (TXT 2 kb)
- Additional file 6: File S2.** Input file for generating the World Map. (TXT 1 kb)

## Abbreviations

ACB: African Caribbeans in Barbados; ASW: Americans of African Ancestry in SW USA; BA: Bronze Age European; BEB: Bengali from Bangladesh; *bp/kb*: Basepairs/Kilobases; CDX: Chinese Dai in Xishuangbanna, China; CEU: Central Europeans in Utah; CHB: Han Chinese in Beijing, China; *chr*: Chromosome; CHS: Southern Han Chinese; EN: Early Neolithic European; ESN: Esan in Nigeria; FIN: Finnish in Finland; FLG: Filaggrin; GBR: British in England and Scotland; GIH: Gujarati Indian from Houston, Texas; GWD: Gambian in Western Divisions in the Gambia; HLA: Human leukocyte antigen; IBS: Iberian Population in Spain; IGV: Integrative genome viewer; ITU: Indian Telugu from the UK; JPT: Japanese in Tokyo, Japan; KHV: Kinh in Ho Chi Minh City, Vietnam; LCE3A/LCE3B/LCE3C/LCE3BC: Late cornified envelope 3 A/B/C/BC; LD: Linkage disequilibrium; LWK: Luhya in Webuye, Kenya; MN: Middle Neolithic European; MSL: Mende in Sierra Leone; PhyML: Phylogenetic estimation using maximum likelihood; PJJ: Punjabi from Lahore, Pakistan; SNPs: Single nucleotide polymorphisms; STU: Sri Lankan Tamil from the UK; SVs: Structural variants; TSI: Toscani in Italia; YRI: Yoruba in Ibadan, Nigeria

## Acknowledgements

We would like to thank Derek Taylor and Victor Albert for careful reading of the earlier versions of this manuscript, as well as to Jessica Poulin for edits on later versions. We would like to thank Pavlos Pavlidis, Nina Jablonski and

Ran Blehman for discussions with regards to function of epidermal differentiation complex and population genetic analyses. We would like to recognize the help from Hayley Hunter-Zinck for providing raw Tsel scores across the genome and discussions regarding their interpretation.

## Funding

This study is supported by OG's start-up funds from University at Buffalo Research Foundation.

## Availability of data and materials

All data used in this study is publically available. Please refer to the Methods section for sources. The codes and alignments presented in the paper is available through the corresponding author's website - <https://gokcumenlab.org/data-and-codes/>.

## Authors' contributions

PP and YL conducted the analyses and wrote the paper. DX designed and conducted phylogenetic analyses. OG designed the study and wrote the paper. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Not applicable.

Received: 18 June 2016 Accepted: 23 November 2016

Published online: 05 December 2016

## References

- Conrad DF, et al. Origins and functional impact of copy number variation in the human genome. *Nature*. 2009;464:704–12.
- Gokcumen O, et al. Primate genome architecture influences structural variation mechanisms and functional consequences. *Proc Natl Acad Sci U S A*. 2013;110:15764–9.
- Perry GH, et al. Diet and the evolution of human amylase gene copy number variation. *Nat Genet*. 2007;39:1256–60.
- Iskow RC, et al. Regulatory element copy number differences shape primate expression profiles. *Proc Natl Acad Sci U S A*. 2012;109:12656–61.
- Dennis MY, et al. Evolution of human-specific neural SRGAP2 genes by incomplete segmental duplication. *Cell*. 2012;149:912–22.
- Iskow RC, Gokcumen O, Lee C. Exploring the role of copy number variants in human adaptation. *Trends Genet*. 2012;28:245–57.
- Sudmant PH, et al. Global diversity, population stratification, and selection of human copy-number variation. *Science*. 2015;349.
- Stefansson H, et al. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008;455:232–6.
- Turner TN, et al. Genome sequencing of autism-affected families reveals disruption of putative noncoding regulatory DNA. *Am J Hum Genet*. 2016;98:58–74.
- Brandler WM, et al. Frequency and complexity of De Novo structural mutation in Autism. *Am J Hum Genet*. doi:10.1016/j.ajhg.2016.02.018.
- Sudmant PH, et al. An integrated map of structural variation in 2,504 human genomes. *Nature*. 2015;526:75–81.
- Eaaswarkhanth M, Pavlidis P, Gokcumen O. Geographic distribution and adaptive significance of genomic structural variants: an anthropological genetics perspective. *Hum Biol*. 2014;86:260–75.
- Lin Y-L, Pavlidis P, Karakoc E, Ajay J, Gokcumen O. The evolution and functional impact of human deletion variants shared with archaic hominin genomes. *Mol Biol Evol*. 2015. doi:10.1093/molbev/msu405.
- Schierup MH, Vekemans X, Charlesworth D. The effect of subdivision on variation at multi-allelic loci under balancing selection. *Genet Res*. 2000;76:51–62.
- Charlesworth D. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genet*. 2006;2:e64.
- Teixeira, et al. Long-term balancing selection in LAD1 maintains a missense trans-species polymorphism in humans, chimpanzees and bonobos. *bioRxiv*. 2014. doi:10.1101/006684.

17. Leffler EM, et al. Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science*. 2013;339:1578–82.
18. DeGiorgio M, Lohmueller KE, Nielsen R. A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genet*. 2014;10:e1004561.
19. Gokcumen O, et al. Balancing selection on a regulatory region exhibiting ancient variation that predates human–neandertal divergence. *PLoS Genet*. 2013;9:e1003404.
20. Andrés AM, et al. Targets of balancing selection in the human genome. *Mol Biol Evol*. 2009;26:2755–64.
21. Andrés AM, et al. Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. *PLoS Genet*. 2010;6:e1001157.
22. Key FM, Teixeira JC, de Filippo C, Andrés AM. Advantageous diversity maintained by balancing selection in humans. *Curr Opin Genet Dev*. 2014;29:45–51.
23. Barreiro LB, et al. The heritage of pathogen pressures and ancient demography in the human innate-immunity CD209/CD209L region. *Am J Hum Genet*. 2005;77:869–86.
24. de Cid R, et al. Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat Genet*. 2009;41:211–5.
25. Riveira-Munoz E, et al. Meta-analysis confirms the LCE3C\_LCE3B deletion as a risk factor for psoriasis in several ethnic groups and finds interaction with HLA-Cw6. *J Invest Dermatol*. 2011;131:1105–9.
26. De Cid R, Riveira-Munoz E, Zeeuwen P, et al. Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Genetics*. 2009;41(2):211–5.
27. Li M, et al. Deletion of the late cornified envelope genes LCE3C and LCE3B is associated with psoriasis in a Chinese population. *J Invest Dermatol*. 2011;131:1639–43.
28. 1000 Genomes Project Consortium, et al. A global reference for human genetic variation. *Nature*. 2015;526:68–74.
29. Fu Q, et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature*. 2014;514:445–9.
30. Somel M, et al. A scan for human-specific relaxation of negative selection reveals unexpected polymorphism in proteasome genes. *Mol Biol Evol*. 2013;30:1808–15.
31. Eaaswarkhanth M, Xu D, Flanagan C, Rzhetskaya M, Hayes MG, Blekman R, Jablonski NG, Gokcumen O. Atopic Dermatitis Susceptibility Variants in Filaggrin Hornerin Selective Sweep. *Genome Biology and Evolution*. 2016;8(10):3240–3255.
32. MacArthur DG, et al. Loss of ACTN3 gene function alters mouse muscle metabolism and shows evidence of positive selection in humans. *Nat Genet*. 2007;39:1261–5.
33. Hudson RR, Boos DD, Kaplan NL. A statistical test for detecting geographic subdivision. *Mol Biol Evol*. 1992;9:138–51.
34. Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. *PLoS Biol*. 2006;4:e72.
35. Arbiza L, Zhong E, Keinan A. NRE: a tool for exploring neutral loci in the human genome. *BMC Bioinformatics*. 2012;13:301.
36. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989;123:585–95.
37. Hunter-Zinck H, Clark AG. Aberrant Time to Most Recent Common Ancestor as a Signature of Natural Selection. *Mol Biol Evol*. 2015;32:2784–97.
38. Hahn MW, Rauscher MD, Cunningham CW. Distinguishing between selection and population expansion in an experimental lineage of bacteriophage T7. *Genetics*. 2002;161:11–20.
39. Bergboer JGM, Zeeuwen PLJM, Schalkwijk J. Genetics of psoriasis: evidence for epistatic interaction between skin barrier abnormalities and immune deviation. *J Invest Dermatol*. 2012;132:3230–31.
40. Bergboer JGM, et al. Psoriasis risk genes of the late cornified envelope-3 group are distinctly expressed compared with genes of other LCE groups. *Am J Pathol*. 2011;178:1470–7.
41. de Guzman Strong C, et al. A milieu of regulatory elements in the epidermal differentiation complex syntenic block: implications for atopic dermatitis and psoriasis. *Hum Mol Genet*. 2010;19:1453–60.
42. Hollox EJ, et al. Psoriasis is associated with increased beta-defensin genomic copy number. *Nat Genet*. 2008;40:23–5.
43. Nair RP, et al. Sequence and haplotype analysis supports HLA-C as the psoriasis susceptibility 1 gene. *Am J Hum Genet*. 2006;78:827–51.
44. Orrù S, Giuressi E, Carcassi C, Casula M, Contu L. Mapping of the Major Psoriasis-Susceptibility Locus (PSORS1) in a 70-Kb Interval around the Corneodesmosin Gene (CDSN). *Am J Hum Genet*. 2005;76:164–71.
45. Bowes J, et al. Variants in linkage disequilibrium with the late cornified envelope gene cluster deletion are associated with susceptibility to psoriatic arthritis. *Ann Rheum Dis*. 2010;69:2199–203.
46. Lu X, et al. Deletion of LCE3C\_LCE3B is associated with rheumatoid arthritis and systemic lupus erythematosus in the Chinese Han population. *Ann Rheum Dis*. 2011;70:1648–51.
47. Baurecht H, et al. Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms. *Am J Hum Genet*. 2015;96:104–20.
48. Haak W, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*. 2015;522:207–11.
49. Prüfer K, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature*. 2014;505:43–9.
50. Meyer M, et al. A high-coverage genome sequence from an archaic Denisovan individual. *Science*. 2012;338:222–6.
51. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*. 2005;437:69–87.
52. Rhesus Macaque Genome Sequencing and Analysis Consortium, et al. Evolutionary and biomedical insights from the rhesus macaque genome. *Science*. 2007;316:222–34.
53. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30:1312–3.
54. Huson DH, et al. Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics*. 2007;8:460.
55. Pybus M, et al. 1000 Genomes Selection Browser 1.0: a genome browser dedicated to signatures of natural selection in modern humans. *Nucleic Acids Res*. 2014;42:D903–9.
56. Robinson JT, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–6.
57. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559–75.
58. Lin Y-L, Pavlidis P, Karakoc E, Ajay J, Gokcumen O. The evolution and functional impact of human deletion variants shared with archaic hominin genomes. *Mol Biol Evol*. 2015;32:1008–19.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

