Insights into early pig domestication provided by ancient DNA analysis

Amke Caliebe1,*, Almut Nebel2,*, Cheryl Makarewicz3, Michael Krawczak1 & Ben Krause-Kyora2,4

Pigs (Sus scrofa) were first domesticated between 8,500 and 8,000 cal BC in the Near East, from where they were subsequently brought into Europe by agriculturalists. Soon after the arrival of the first domestic pigs in northern Europe (~4500 BC), farmers are thought to have started to incorporate local wild boars into their swine herds. This husbandry strategy ultimately resulted in the domestication of European wild boars. Here, we set out to provide a more precise geographic and temporal framework of the early management of suid populations in northern Europe, drawing upon mitochondrial DNA haplotype data from 116 Neolithic Sus specimens. We developed a quantitative mathematical model tracing the haplotypes of the domestic pigs back to their most likely geographic origin. Our modelling results suggest that, between 5000 and 4000 BC, almost all matrilines in the north originated from domesticated animals from the south of central Europe. In the following period (4000–3000 BC), an estimated 78–100% of domesticates in the north were of northern matrilineal origin, largely from local wild boars. These findings point towards a dramatic change in suid management strategies taking place throughout south-central and northern Europe after 4000 BC.

Pigs were first domesticated in the Near East around 8500 BC and subsequently brought into Europe by agriculturalists1. Ancient mitochondrial DNA (mtDNA) studies further indicate that, by 4500 BC, domesticated pigs bearing Near Eastern haplotypes appeared in northern Europe2. Soon thereafter, agriculturalists are thought to have incorporated local wild boars into their domesticated swine herds1,2. On the basis of a relatively small and geographically diffuse set of mtDNA data, it has been hypothesized that the Near Eastern mtDNA lineages were rapidly replaced by European haplotypes, apparently within 500 years or less, after the first introduction of Near Eastern domesticates into the region1.

Here, we set out to derive a more precise geographic and temporal framework for the early management of indigenous wild boar populations in northern Europe, a process which ultimately led to domesticated European lineages. We examined ancient mtDNA haplotype data from a large sample of 87 domestic and 29 wild Neolithic Sus specimens (Fig. 1). The study targeted an 80-bp highly informative fragment of the mitochondrial control region that provides ample information about maternal ancestry1,3–5. The presence of Near Eastern haplotypes Y1 and Y2 in pigs located in Europe has been previously interpreted as a marker of their domesticated status1, demonstrating the role of humans in facilitating the translocation of suids from the Near East into Europe. Reflecting the difficulty in resolving the domesticated status of animals by lineage alone in regions inhabited by both domesticates and their wild progenitors, Sus with the European haplotype A or C could represent either wild or domestic animals5. In order to better resolve the source and timing of pig management processes, we developed a stochastic model to trace the mtDNA haplotypes from nascent domestic pig populations of northern Europe of 5000 BC to 3000 BC back to their likely origin. Our study is the first to use such a quantitative model to ascertain the origin of prehistoric mtDNA haplotypes of northern domestic pigs.

1Institute of Medical Informatics and Statistics, Kiel University, 24105 Kiel, Germany. 2Institute of Clinical Molecular Biology, Kiel University, 24105 Kiel, Germany. 3Institute of Prehistoric and Protohistoric Archaeology, Kiel University, 24098 Kiel, Germany. 4Max Planck Institute for the Science of Human History, 07745 Jena, Germany. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to A.C. (email: caliebe@medinfo.uni-kiel.de) or B.K.-K. (email: b.krause-kyora@ikmb.uni-kiel.de)
Results

Study design and aim. In our sample of 116 Neolithic Sus specimens, we observed the four known mtDNA haplotypes Y1, Y2, A and C1 (Tables 1 and 2). Interestingly, Near Eastern haplotype Y1 appeared as late as 4000 to 3000 BC in northern domestic pigs.

For the present study, pig remains recovered from sites located in the northern parts of Germany and the Netherlands were classified as belonging to the ‘northern group’ (n = 73), irrespective of haplotype (Fig. 1, Table S1). The northern region was home to late Mesolithic hunter-gatherers who eventually turned towards the use of domesticated plants and animals during the Neolithic (ca 4100 BC). Haplotype frequencies obtained from northern group pigs were compared to those among specimens recovered from southern Poland and central/southern Germany (‘southern group’, n = 43; Fig. 1), a region where Neolithic farmers had been established since ~5500 BC. The wild or domestic status of each specimen was determined based upon (i) its mtDNA haplotype (with Y1 or Y2 indicating domestic status) and (ii) standard metrical and non-metrical analyses (cp. Methods).

Animals were grouped further according to the time period (t1: 5500–5000 BC, t2: 5000–4000 BC, t3: 4000–3000 BC) and region (north vs. south) they had been living in (Tables 1 and 2). The three time periods and the two geographical regions were characterized by distinct subsistence strategies, landscape interactions and cultural material used. The first time period (t1) was largely defined by the presence of Mesolithic hunter-gatherer-fishers in the north and the first farmers in the south (5500–5000 BC). The second period (t2) (5000–4000 cal BC) was described in the north by the transition to early farmers and in the south by the super-regional Linienbandkeramik culture. The third period (t3) (4000–3000 cal BC) was defined by the emergence of more regional agriculturalist Funnel Beaker and Globular Amphora culture. The mtDNA haplotypes of the t2 sample of wild boars from the north (n = 29) were used as a reference for the haplotype distribution of wild boars in periods t1 and t3 (Table 2). In our stochastic model, we investigated time periods t2 and t3 separately.
determined the probability that the mtDNA of a northern domestic pig derived from an animal of certain status (domestic or wild) living in a certain region (north or south) during the previous time period (i.e., t1 or t2). Based on the occurrence of European haplotypes A and C in southern domestic pigs, it is clear that wild boars had been continuously integrated into the domestic livestock in the south2. Therefore, the group of southern domestic pigs consists not only of domestic pigs originally descended from domestic pigs from the Near East, but also of previously domesticated southern wild boars.

Origin of mtDNA haplotypes of northern domestic pigs living 5000–4000 BC (t2). For time period t2 (5000–4000 BC), no domestic pigs have so far been identified in northern European contexts6,7. Therefore, the mtDNA of a northern t2 domestic pig can only originate from either a southern t1 domestic pig or a northern t1 wild boar (Fig. 2A). A single parameter is sufficient to describe this model, namely the probability pNG that the mtDNA of a northern t2 domestic pig descended from a northern t1 wild boar. Our data yielded a maximum likelihood estimate of 0.20 for pNG for two haplotypes (combining A and C on the one hand, and Y1 and Y2 on the other) and of 0 when all four haplotypes were considered. The 95% confidence interval for pNG was [0, 0.83] for two haplotypes and [0, 0.28] for four haplotypes (Table 3, Figures S1A and B). The two-haplotype confidence interval is considerably larger because the European haplotypes A and C were combined, which masks the fact that in southern t1 domestic pigs, haplotype A is more frequent than haplotype C (A: 64% vs. C: 36% among the European haplotypes) whereas in northern t1 wild boars haplotype C is almost five times more frequent than A (A: 17% vs. C: 83%). The four haplotype confidence interval takes the separate information about A and C into account and therefore yields a more precise estimate.

Origin of mtDNA haplotypes of northern domestic pigs living 4000–3000 BC (t3). Since domestic pigs were already present in northern Europe in time period t2, mtDNA haplotypes of northern t3 domestic pigs have three possible origins, namely southern t2 domestic pigs, northern t2 domestic pigs or northern t2 wild boars (Fig. 2B). In addition to parameter pNG, i.e. the probability that the haplotype was derived from a northern rather than a southern t2 animal, a second parameter pDP is required to distinguish between northern domestic and northern wild predecessors (pDP equals the probability that a northern predecessor was domestic rather than wild). For two haplotypes, maximum likelihood estimation did not yield a unique solution in our data. Instead, possible values of pNG and pDP were found to lie on a straight line (Fig. 3A). Consequently, pNG was estimated to lie.
Why did early agriculturalists in northern Europe change their suid management strategy after 4000 BC? It may be that the availability of domesticated Near Eastern pigs in the earliest Neolithic settlements located in northern Europe was initially limited, reflecting their location at the end of the Neolithic ‘supply chain’. The reliability of Near Eastern-derived livestock was perhaps also unstable, particularly if pig husbandry strategies applied to Near Eastern domesticates initially failed to adjust for northern European environments and dietary conditions. The overall number of domestic pigs carried to the north by the first Neolithic agriculturalists along with their livestock may have been relatively small. As a case in point, domesticated sheep and goats, core constituents of the Near Eastern Neolithic livestock package, were also exploited at low intensities only in the earliest Neolithic settlements of the southwestern Baltic region, with their remains generally comprising no more than 10% of the faunal assemblage. A small domestic source population would have been susceptible to isolation, disease and over-harvesting which would have made it difficult to maintain a stable and viable swine herd. In addition, the material cultural record indicates that the contact between northern and southern
European cultures was only sporadic at the time. If this was the case, then the strategy of interbreeding locally abundant female wild boars with domestic males would have facilitated an increase in swine herd size without having to rely upon a limited and uncertain supply of Near Eastern domesticated pigs from the south. Moreover, local stock are usually well adapted to their specific environmental conditions, for instance, by showing greater resistance to endemic pathogens. Out-crossing therefore may have produced more resilient, fertile and larger offspring. Another explanation for the dominance of local European mtDNA genomes in later pig specimens could be that the corresponding genomes conferred a selective advantage in the north, for example, in terms of energy metabolism that would eventually lead to their fixation in European domesticates.

It has been suggested before that interbreeding between pigs and wild boars was mainly unintentional and resulted from chance encounters due to escaped feral domesticates, loose swine management systems and mobile swine herding. Loose pig husbandry is thought to have been characteristic of European agricultural systems for millennia. However, this scenario would not explain the introgression of European wild sow mtDNA haplotypes (i.e. of female animals) into expanding Neolithic swineherds; mating between domesticated female sows and European wild boars would not have changed the mtDNA pool of domestic herds. The latter instead would have required the incorporation of additional wild females, which could, for instance, be achieved by capture of wild female piglets as part of an active herd-building strategy. Therefore, a combination of loose management and intentional integration of female wild boars seems to have occurred.

The results of our quantitative analyses strongly suggest that pig management was an ever-evolving process which depended heavily upon interbreeding domesticated animals with local wild boar, in particular wild sows. This insight demonstrates how a change in animal handling ~6000 years ago may have influenced livestock composition up to the present day. What is more, our mathematical model presented here can also be adapted to other proxies (i.e. animals) and haploid DNA markers as well as scenarios and periods.

**Methods**

**Sample specification and determination of the wild or domestic status of samples.** The domestication status of the sampled *Sus* remains was determined in a two-stage process. Since Near Eastern mtDNA haplotypes Y1 and Y2 classify a pig as domestic with very high probability, samples with one of these two haplotypes were assigned domestic status (n = 25). Since pigs with European haplotype A or C can be either domestic or wild, we had to rely upon standard metrical and non-metrical morphological analyses for the classification of *Sus* with these haplotypes (n = 91).

Metrical data retrieved from the appendicular skeleton and from teeth, in particular the third mandibular molar, are commonly used to evaluate change in animal body size associated with intensive husbandry. Large bodied animals are generally associated with the wild condition. Size diminution over time is thought to reflect the introduction of anthropogenic selective pressures and the onset of domesticated phenotypes with small-bodied animals reflecting their domesticated status. However, especially in the early stages of the domestication process when progenitor species may have been only loosely managed, the relationship between body size and animal ‘domestication’ status is less clear-cut. Status determination taking into account body size alone can therefore be error-prone. For this study, pig bone specimens were classified as belonging to ‘domesticated, or ‘wild’ according to criteria established independently by numerous zooarchaeologists over the course of several decades. For the majority of pig bone specimens obtained from the northern region, domesticated status was established using pig bone and tooth specimens from the medieval site of Hedehy located in Schleswig-Holstein, northern Germany. For each skeletal element, all Mesolithic or Neolithic bone specimens smaller than Hedehy specimens were classified as ‘domesticated’ (Schmölcke, personal communication). Bone and tooth specimens within the range of modern European wild boar were classified as ‘wild’, while bone and tooth specimens larger than those from Hedehy but smaller than modern wild boar were classified as ‘indeterminate’. The latter were not considered in the present study. Bone specimens recovered from sites located in regions outside of northern Europe were identified as belonging to wild or domesticated animals according to criteria established by individual analysts (Table S1). In addition, this study included mtDNA haplotype information to assist in classification (e.g. specimens yielding Near Eastern Y1 or Y2 lineages were classified as domesticated). In addition, the genetic analysis of a relatively large number of ancient specimens largely allows for zooarchaeological misclassification errors. The type of skeletal element used for aDNA analysis and the archaeological context from which each specimen is derived are presented in Table S1.

**aDNA quality control procedures and mtDNA sequence analysis.** All analyses were performed in the aDNA laboratory established at Kiel University, following the strict guidelines and procedures for work with minute amounts of degraded DNA. Using previously described methods, mtDNA haplotypes were determined for an 80 bp diagnostic fragment from the control region of the *Sus* mitochondrial genome. All sequencing results are shown in Table S1.

**Stochastic models for the origin of mtDNA haplotypes in northern domestic pigs.** In the following, pig subspecies (ps) are denoted by DP (domestic pig) and WB (wild boar), and regional affiliation (r) are denoted by NG (northern group) and SG (southern group). Then, $P(h|t_i, r, ps)$ is the conditional probability that an animal from subspecies ps living in region r during time period $t_i$ carried haplotype h. For the sake of brevity, we define $P_{Dom}(t_i)$ as the probability that a northern $t_i$ domestic pig descended from a northern $t_i-1$ pig (either DP or WB), and $P_{North}(t_i)$ is the probability that a northern $t_i$ domestic pig descended from a domestic $t_i-1$ pig, given that this predecessor came from the northern group.

In modelling the origin of mtDNA haplotypes of domestic pigs from the northern group, we will assume that no domestic pigs were present in that group during time period $t_i$. A northern $t_i$ domestic pig therefore
descended either from a northern \( t_1 \) wild boar or from a southern \( t_1 \) domestic pig. This results in the following model equation:

\[
P(h|t_2; NG; DP) = P(h|t_2; NG; WB)P_{NG}(t_2) + P(h|t_2; SG; DP)(1 - P_{NG}(t_2)).
\]  

(1)

Northern \( t_2 \) domestic pigs are assumed to have descended from (i) northern \( t_2 \) domestic pigs, (ii) northern \( t_2 \) wild boars or (iii) from southern \( t_2 \) domestic pigs. The resulting model equation thus reads

\[
P(h|t_2; NG; DP) = P(h|t_2; NG; WB)P_{NG}(t_2) + P(h|t_2; SG; DP)(1 - P_{NG}(t_2)).
\]  

(2)

The mathematical models underlying the statistical analyses are relatively simple in several regards. Only one or two parameters are considered and the time and geography are treated as discrete (i.e. three time periods, two geographic regions). These simplifications take into account the limitation of (i) the sample size, which prohibits employment of multi-parameter models, and (ii) the level of accuracy achievable by way of archaeological dating. Therefore, the models appear adequate for the aims and scope of the present study acknowledging that more refined temporal or geographic structures could not be resolved.

Parameter estimation. The two model parameters \( p_{NG}(t_i) \), the probability that a northern \( t_i \) domestic pig descended from a northern \( t_i \) pig, and \( p_{DP}(t_i) \), the probability that a northern \( t_i \) predecessor was domestic rather than wild, were estimated by maximum likelihood using absolute frequencies \( n_h(t_i) \), for short \( n_h \), of haplotypes A, C, Y1 and Y2 in the sampled domestic pigs from the northern group in time period \( t_i \). Assuming that the sampled pigs were stochastically independent, these data follow a multinomial distribution:

\[
\text{Likelihood}(p_{NG}(t_i), p_{DP}(t_i)| n_A(t_i), n_C(t_i), n_{Y1}(t_i), n_{Y2}(t_i)) = 
\sum_{n} \left( \frac{n!}{n_A!n_C!n_{Y1}!n_{Y2}!} \right) \left( n_A \cdot P(A|t_i; NG; DP) \right)^{n_A} \left( n_C \cdot P(C|t_i; NG; DP) \right)^{n_C} \left( n_{Y1} \cdot P(Y1|t_i; NG; DP) \right)^{n_{Y1}} \left( n_{Y2} \cdot P(Y2|t_i; NG; DP) \right)^{n_{Y2}}
\]  

(3)

Parameters \( p_{NG}(t_i) \) and \( p_{DP}(t_i) \) enter into the likelihood model via Equations (1) or (2), depending upon the respective time period. Conditional probabilities \( P(h|t_2; NG; DP) \), \( P(h|t_2; NG; WB) \), \( P(h|t_2; SG; DP) \) and \( P(h|t_2; SG; DP) \) can be estimated from the available mtDNA data. For \( P(h|t_2; NG; WB) \), however, no such data were available so that we had to assume \( P(h|t_2; NG; WB) \).

Calculations were performed with statistics software R\(^{18} \). Numerical maximization of the likelihood was carried out with function optim allowing for box constraints on probabilities\(^{19} \). To deal with multiple maxima, we chose starting values from a grid of 12 equally spaced values between 0 and 1 and inspected the likelihood surface by means of a 2D or 3D plot. For the estimation of \( p_{NG} \) for time period \( t_2 \), confidence intervals were calculated by bootstrapping (\( n = 100,000 \) bootstrap samples). For the estimation of \( p_{NG} \) and \( p_{DP} \) for time period \( t_3 \), bootstrapping and therefore the calculation of a confidence interval was not feasible because the three-dimensional likelihood surface was too complex.

References

1. Larson, G. et al. Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. Proc Natl Acad Sci USA 104, 15276–15281 (2007).
2. Krause-Kyora, B. et al. Use of domesticated pigs by Mesolithic hunter-gatherers in northwestern Europe. Nat Common 4, 2348 (2013).
3. Vai, S. et al. The Biezo case in northern Italy: is the temporal dynamic of swine mitochondrial DNA lineages in Europe related to domestication? Sci Rep 5, 16514 (2015).
4. Larson, G. et al. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. Science 307, 1618–1621 (2005).
5. Ottone, C. et al. Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics. Mol Biol Evol 30, 824–832 (2013).
6. Rowley-Conwy, P. & Zeder, M. Mesolithic domestic pigs at Rosenhof – or wild boar? A critical re-appraisal of ancient DNA and geometric morphometrics. World Archaeology 46, 813–824 (2014).
7. Rowley-Conwy, P. & Zeder, M. Wild Boar or Domestic Pigs? Response to Evin et al. World Archaeology 46, 835–840 (2014).
8. Rowley-Conwy, P. Westward Ho! The spread of agriculture from Central Europe to the Atlantic. Carr Anthropol 52, S431–S451 (2011).
9. Frantz, L. A. et al. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. Nature genetics 47, 1141–1148 (2015).
10. Hübner, K.-D., Saur, R. & Reichstein, H. Palynologische und säugetierkundliche Untersuchungen zum Siedlungsplatz Hüde I am Dümmer Landkreis Diepholz. In Dümmer Landkreis Diepholz. In: Jacob-Friesen, G., Schütrumpf, R., Hübner, K.-D., Saur, R., Reichstein, H. (eds) Wachholtz 1988).
11. Heinrich, D. Die Tierknochen des frühneolithischen Wohnplatzes Wangels LA 505. Ein Vorbericht. Curr Anthropol 54, 43–48 (1998).
12. Marshall, F. R., Dobney, K., Denham, T. & Capriles, J. M. Evaluating the roles of directed breeding and gene flow in animal domestication. Proc Natl Acad Sci USA 111, 6153–6158 (2014).
13. Scandura, M., Iacolina, L. & Apollonio, M. Genetic diversity in the European wild boar Sus scrofa: phylogeography, population structure and wild x domestic hybridization. Mammal Rev 41, 125–137 (2011).
14. White, S. From globalized pig breeds to capitalist pigs: a study in animal cultures and evolutionary history. Environ Hist-Us 16, 94–120 (2011).
15. Zeder, M. A., Emshwiller, E., Smith, B. D. & Bradley, D. G. Documenting domestication: the intersection of genetics and archaeology. Trends in Genetics 22, 139–155 (2006).
16. Lee, E. et al. Ancient DNA insights from the Middle Neolithic in Germany. Archaeol Anthropol Sci 6, 199–204 (2014).
17. Lee, E. J. et al. Collective burials among agro-pastoral societies in later Neolithic Germany: perspectives from ancient DNA. *J Archaeol Sci* 51, 174–180 (2014).

18. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/ (2013).

19. Byrd, R. H., Lu, P. H., Nocedal, J. & Zhu C. Y. A limited memory algorithm for bound constrained optimization. *Siam J Sci Comput* 16, 1190–1208 (1995).

**Acknowledgements**

Special thanks for support go to Ulrich Schmölcke, Centre for Baltic and Scandinavian Archaeology, Schleswig. We thank the Archaeological State Museum Schleswig-Holstein, the Archaeological State Offices of Brandenburg, Lower Saxony and Saxony, the Institute of Archaeology of the Nicolaus Copernicus University Bydgoska, the Groningen Institute of Archaeology of the University of Groningen, the IPNA of the University of Basel and the following individuals who provided sample material: Jörg Ewersen, Michael Geschwinde, Silke Grefen-Peters, Susanne Hanik, Daniel Makowiecki, Wietske Prummel, Uwe Reuter, Marguerita Schäfer, Reinhold Schoon, Ulrich Schmölcke and Regina Smolnik. We are grateful to Arne Jochens for fruitful discussions and comments on the subject of the manuscript. This work was funded by the Graduate School ‘Human Development in Landscapes’ at Kiel University (CAU). Radiocarbon dating was carried out at the Leibniz Laboratory, CAU.

**Author Contributions**

B.K.-K., A.N., M.K. and A.C. conceived and designed the research. B.K.-K. generated and A.C. analyzed the ancient DNA data. All authors interpreted the findings. A.N., A.C. and M.K. wrote the manuscript with input from B.K.-K. and C.M.

**Additional Information**

Supplementary information accompanies this paper at http://www.nature.com/srep

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** Caliebe, A. et al. Insights into early pig domestication provided by ancient DNA analysis. *Sci. Rep.* 7, 44550; doi: 10.1038/srep44550 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

© The Author(s) 2017