Correspondence

Estimating human age from T-cell DNA rearrangements

D. Zubakov[1], F. Liu[1], M.C. van Zelm[2], J. Vermeulen[2], B.A. Oostra[3], C.M. van Duijn[4], G.J. Driessen[5], J.J.M. van Dongen[2], M. Kayser[1,§,*], and A.W. Langerak[2,§]

Predicting human phenotypes from genotypes is a newly emerging field with relevance for personalized medicine [1] and forensics [2]. However, only a few phenotypic traits can currently be identified from DNA information with accuracies sufficient for practical applications [1], most notably human eye (iris) color [3]. It could be expected that individual age is too biologically complex to allow a simple and accurate molecular estimation from biological materials. Indeed, previously proposed genetic methods for human age estimation, based on the accumulation of mitochondrial DNA deletions or on telomere shortening, show low accuracies and various technical problems, and are therefore not suitable for practical applications [4]. Proposed biochemical methods, such as those based on the accumulation of D-aspartic acid, involve the destructive analysis of specific body parts (such as bones, teeth and ligaments), and suffer from technical issues and bio-degradation [4]. In the present study, we demonstrate that human individual age can be estimated accurately and reliably from blood using T-cell DNA rearrangements, and we provide a robust and sensitive real-time quantitative PCR protocol for application in various areas of bioscience.

T lymphocytes employ specific receptors, called T-cell receptors (TCRs), to recognize foreign antigens. In order to create a broad repertoire of TCR molecules, each immature T lymphocyte undergoes unique somatic rearrangements in its TCR loci during intra-thymic development. During this rearrangement process, the intervening DNA sequences in the TCR loci are deleted and circularized into episomal DNA molecules, also called signal joint TCR excision circles (sjTRECs) [5]. One particular sjTREC is the δRec-ψJα sjTREC that arises through an intermediate rearrangement in the TCRD/TCRA locus in developing TCRαβ$^+$ T lymphocytes (see Supplemental Figure S1 in the online Supplemental Information section). This sjTREC is seen in approximately 70% of all newly-formed mature TCRαβ$^+$ T lymphocytes [6] (see Supplemental Information for more details). As shown previously, there is a log-linear decline in sjTREC number with increasing human age, reflecting a life-long process of thymus involution that starts shortly after birth by replacement with adipose tissue and consequent loss of thymic function [7]. Here we show that this biological phenomenon can be used for estimating human individual age accurately and reliably.

We employed a TaqMan qPCR approach to quantify sjTREC levels, normalized to the single-copy albumin gene to account for the amount of input DNA, in whole-blood samples of 195 healthy Dutch individuals ranging in age from a few weeks to 80 years (for sample and methodological details see Supplemental Information). We then used the measured sjTREC abundance as the single predictor in a linear regression model, which explained a large and highly statistically significant portion of the total age variance (R^2 = 0.835, P = 8.16×10^{-215}, standard error of the estimate ± 8.9 years; Figure 1). The R^2 estimates were robust against 10,000 cross-validations with an 80% training and 20% testing split (mean R^2 = 0.836, StD = 0.036): for further details see Supplemental Information. We subsequently employed a model-based prediction approach using a multinomial logistic regression model [3] for estimating human age across four categories each spanning 20 years that approximately correlate to different generations. The prediction accuracy was evaluated by means of the area under the ROC curves (AUC), ranging from 0.5 (random) to 1 (perfect) prediction, as well as by using additional prediction estimates (see Supplemental Information). We obtained very high AUC values (from 0.89 up to 0.97) for all four age categories (Figure 1), which were robust against the cross-validations (see Supplemental Information). Notably, the achieved AUC values include the highest prediction

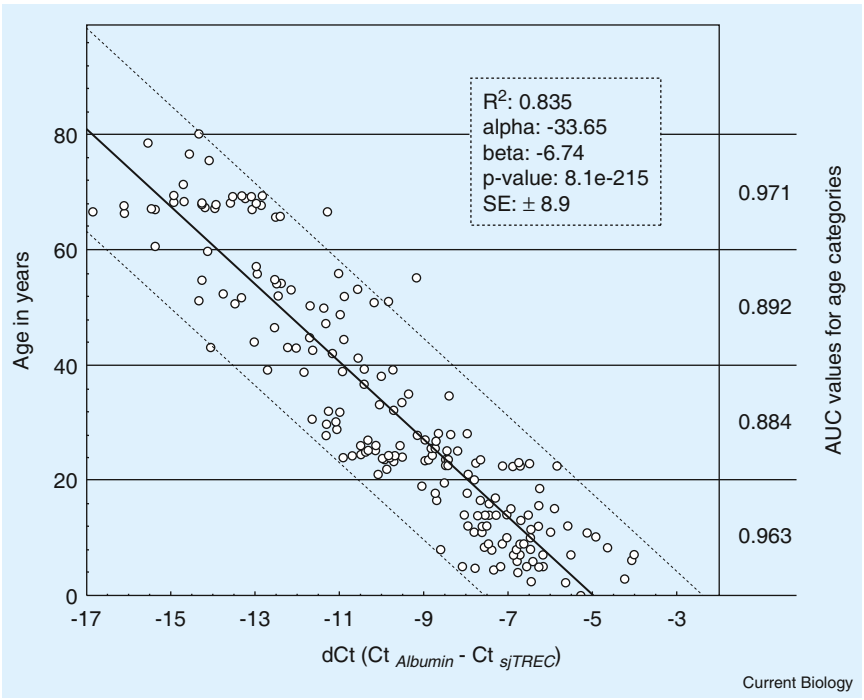

Figure 1. Age prediction from blood-derived DNA samples.
Linear regression of the relationships between human individual age and normalized sjTREC abundance in peripheral blood of 195 individuals (dotted lines correspond to 95% prediction interval) and the mean AUC values from 1,000 cross-validations in predicting age categories each spanning 20 years.

accuracies obtained so far for any human trait, including those recently described for blue and brown eye color [3] for which practical application has been suggested [8]. Furthermore, we tested for the influence of sample storage time by analyzing fresh and 1.5-year-old blood samples of the same individuals and found no statistically significant difference between the sjTREC quantifications (see Supplemental Information for details). This provides a first indication that our approach, employing small amplicon sizes of 140 bp for sjTREC and 118 bp for albumin (used for normalization), is remarkably robust against the negative effects of DNA degradation and can be successfully applied to aged blood stains. Furthermore, the test we propose here requires small amounts of DNA; although all 195 samples were analyzed from 50 ng of DNA to ensure accurate results, our preliminary sensitivity testing revealed that the assay works reliably from 5 ng, at least in young individuals (see Supplemental Information for details).

Unlike odontological or skeletal approaches for age estimation [9] or some biochemical methods [4], our DNA-based approach using blood does not require the availability of invasive samples such as bones or teeth, and thus expands the availability of biological age estimations for practical applications. The sensitivity, high prediction accuracy and apparent time-wise stability are important prerequisites for practical applications of the proposed DNA age test. Within the forensic context, our approach is expected to provide investigative leads in criminal cases by allowing an accurate estimation of the generation age (clearly reflected in a person's appearance) of unknown individuals from minute blood stains. Our approach is also expected to be applied to disaster victim identification where body parts (containing blood) are available and where age information can be crucial for final identification. Furthermore, our method is relevant for several other practical applications, such as: in the determination of 'immunological' age (as compared to chronological age) under certain conditions of aging (immunosenescence) or disease; in anthropological studies (including ancient DNA studies) where the age information of an individual is important but not available; in clinical applications where age records were lost; and potentially for age estimation in wild animals for ecological research and wildlife management. The application of our method is restricted to blood samples and body parts containing blood and is not possible for other body parts or fluids, such as semen or saliva, that do not contain T cells in quantities required for sjTREC detection.

Previous studies detected a slight gender difference in age-dependent sjTREC numbers [10], and we also observed a small but statistically significant gender effect on sjTREC quantification (see Supplemental Information). However, including gender as an additional factor into regression and categorical prediction modeling did not alter the age estimates to any significant level (see Supplemental Information). A potential drawback of the proposed age estimation approach could be in the effect of the immune system health status on the sjTREC copy numbers [7]. It remains to be addressed in future studies whether severe pathological conditions involving the immune system, such as HIV/AIDS or leukemia, substantially influence the accuracy of this method. Another issue to be evaluated is the potential influence of a person's bio-geographic ancestry on our approach for age estimation. Future research into other aspects of age-dependent biological changes may allow further improvement of accuracy in individual age estimation from biological materials.

Acknowledgments

We are grateful to all volunteers for providing samples and to Kaye Ballantyne for valuable comments on the manuscript. This study was financially supported in parts by the Centre for Medical Systems Biology (CMSB), an EUR-Fellowship to M.C.v.Z., the Netherlands Forensic Institute to M.K., and by the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) within the framework of the Forensic Genomics Consortium Netherlands (FGCN) to M.K.

Supplemental Information

Supplemental Information including methods and supplemental results can be found with this article online at doi:10.1016/j.cub.2010.10.022.

References

1. Janssens, A.C., and van Duijn, C.M. (2008). Genome-based prediction of common diseases: advances and prospects. Hum. Mol. Genet. *17*, R166–R173.
2. Kayser, M., and Schneider, P.M. (2009). DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. Forensic Sci. Int. Genet. *3*, 154–161.
3. Liu, F., van Duijn, K., Vingerling, J.R., Hofman, A., Uitterlinden, A.G., Janssens, A.C., and Kayser, M. (2009). Eye color and the prediction of complex phenotypes from genotypes. Curr. Biol. *19*, R192–R193.
4. Meissner, C., and Ritz-Timme, S. (2010). Molecular pathology and age estimation. Forensic Sci. Int. doi:10.1016/j.forsciint.2010.07.010
5. Breit, T.M., Verschuren, M.C., Wolvers-Tettero, I.L., Van Gastel-Mol, E.J., Hahlen, K., and van Dongen, J.J. (1997). Human T cell leukemias with continuous V(D)J recombinase activity for TCR-delta gene deletion. J. Immunol. *159*, 4341–4349.
6. Verschuren, M.C., Wolvers-Tettero, I.L., Breit, T.M., Noordzij, J., van Wering, E.R., and van Dongen, J.J. (1997). Preferential rearrangements of the T cell receptor-delta-deleting elements in human T cells. J. Immunol. *158*, 1208–1216.
7. Douek, D.C., McFarland, R.D., Keiser, P.H., Gage, E.A., Massey, J.M., Haynes, B.F., Polis, M.A., Haase, A.T., Feinberg, M.B., Sullivan, J.L., et al. (1998). Changes in thymic function with age and during the treatment of HIV infection. Nature *396*, 690–695.
8. Walsh, S., Liu, F., Ballantyne, K.N., van Oven, M., Lao, O., and Kayser, M. (2010). IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci. Int. Genet. doi:10.1016/j.fsigen.2010.02.004
9. Schmeling, A., Geserick, G., Reisinger, W., and Olze, A. (2007). Age estimation. Forensic Sci. Int. *165*, 178–181.
10. Lorenzi, A.R., Patterson, A.M., Pratt, A., Jefferson, M., Chapman, C.E., Ponchel, F., and Isaacs, J.D. (2008). Determination of thymic function directly from peripheral blood: a validated modification to an established method. J. Immunol. Methods *339*, 185–194.

Departments of [1]Forensic Molecular Biology, [2]Immunology, [3]Clinical Genetics, [4]Epidemiology, [5]Pediatrics, Erasmus MC University Medical Center Rotterdam, PO Box 2040, 3000 CA Rotterdam, The Netherlands.
[§]These authors contributed equally to this work and are listed in alphabetical order.
*E-mail: m.kayser@erasmusmc.nl