**Title: The role of archaic human genes in the current epidemic of diabetes mellitus in Indigenous Australians.**

**Author List and Affiliations:**

**Manuscript Authors:**

1. **Dr. MM Brzozowska, MD, FRACP, PhD**
   - Endocrinology Department, Sutherland and St George Hospital, New South Wales, Australia,
   - St George & Sutherland Hospital Clinical School, University of New South Wales, Australia
Corresponding author: Dr MM Brzozowska,

Email:  m.brzozowska@garvan.org.au
Phone: 61 2 9540 7111

2. **Mr. RB Allen** BSc
   - School of Archeology, University of Oxford, United Kingdom

Email:  richard.allen@arch.ox.ac.uk

**Key words:** archaic human genes, obesity, diabetes mellitus, Indigenous Australians.

**Manuscript's word count: 2294**

**There are no supporting grants or conflicts of interest to declare**

**Abstract**

Indigenous Australians have been particularly affected by type 2 diabetes mellitus due to their genetic susceptibility and a range of environmental risk factors. Recent genetic studies link predisposition to some diseases, including diabetes, to archaic humans, such as Neanderthals and Denisovans, suggesting persistence of ancient alleles in the genomes of modern humans.  In this review we discuss the evolutionary role of the negative genetic selection associated with an adopted Western lifestyle as well as DNA variants influencing predisposition to obesity and diabetes in the Australian Indigenous population. We review the contribution of the ancient gene/pathways to the modern human phenotypes including the *Neanderthal* haplotype-tagging SNPs in *NTRK2* gene, which may continue to play a role in obesity in Indigenous Australians.

**Introduction**

The prevalence of diabetes mellitus (DM) has increased exponentially worldwide with some of the populations being particularly affected [1].  Type 2 diabetes mellitus in particular, has reached global epidemic proportions posing a significant challenge for the Australian public health system [1].  Aboriginal and Torres Strait Islander Australians experience disproportionately high levels of diabetes, which contributes to approximately 7% of the disease burden in the Aboriginal and Torres Strait Islander population [2].  Furthermore the diabetes mellitus rates among Aboriginal and Torres Strait Islander people are three to five times higher than for Australians of European descent [3].  Indigenous Australians are diagnosed with diabetes approximately 14 years earlier than Anglo-Celt patients [4].

1

Notably, Aboriginal and Torres Strait Islander children are eight times more likely to develop type 2 diabetes mellitus than their non-Indigenous peers[1].

Finally, Aboriginal and Torres Strait Islander people have poor clinical outcomes from diabetes treatment with six fold increased mortality rates when compared with non-Indigenous Australians[1].

The above differences in prevalence of diabetes mellitus, clinical complications and mortality rates between Indigenous and non-Indigenous Australians represent a complex interplay between genetic susceptibility and the environmental and lifestyle risk factors.  In particular obesity, which is closely associated with diabetes, as well as age (35 years and onwards) are two critical risk factors affecting the prevalence of type 2 diabetes mellitus in Aboriginal population [5].  Aboriginal Australians display body habitus which differ from non-Aboriginals in Australia with preferential abdominal fat deposition, exacerbating their insulin resistance [6] and higher risk for metabolic complications including diabetes mellitus with body mass index (BMI) scores above that of 22 kg/m$^2$ [7].

*Genetic adaptation to dietary and environmental changes*

Importantly, changes in food availability and diet composition might have significantly influenced the metabolic health of the Australian Indigenous population.  Before the recent arrival of Europeans, for thousands of years, Aboriginal people lived as hunter gatherers with a high reliance on animal-based foods coupled with the relatively low, by Western modern standards, carbohydrate content of wild plant foods.  Modern reconstructions of the worldwide hunter-gatherer diets have revealed their very high protein content from which Indigenous people derived between 19% and 50% of total energy supply [8].  On the contrary, the modern diet provides on average more than 70% of energy intake from refined sugars, refined vegetable oils, highly processed cereals, and dairy products [9].  The reported evolutionary mismatch between modern dietary constituents and the food available prior to the agricultural revolution has been considered a factor in the obesity epidemic [10].  A recent hypothesis proposed that regular consumption of a high carbohydrate diet, when combined with protein intake, may perpetuate obesity and abnormal gluconeogenesis in modern humans [11].  Furthermore, instead of focusing primarily on caloric imbalances, the researchers proposed an alternative obesitogenic pathway when processing of simple carbohydrates and starches as a main source of body energy leads to excessive conversion of protein into triglycerides and ultimately into the adipose tissue leading to obesity.  This hypothesis has been further supported by a recent study which identified the central role of a hepatic enzyme, alanine transaminase (ALT) in lipogenic pathway and body weight regulation [11].

Recently, the importance of genetic adaptation to the modern diet with progressive accumulation of genes coding for suboptimal energy metabolism has been attracting scientific attention.  The identification of genetic loci with a major role in dietary adaptations has been limited to genes coding for enzymes with highly specific functional roles in nutrient metabolism such as lactase (*LCT*) and amylase (*AMY1*).  The greater number of copies of the AMY1 gene coding for amylase (an enzyme important for carbohydrate digestion) has been found in populations that evolved under high-starch diets versus low-starch diets, consistent with an intense positive genetic selection imposed by diet on the amylase gene copy number during dietary changes [12].

The very recent introduction to the modern diet might have been deleterious to the metabolic health of Aboriginal populations in a situation of relaxed selection against genes affecting their energy balance and metabolism due to reduced mortality [13]. The concept of the limited opportunity for genetic adaptation towards the environmental factors in Australian Aboriginal populations has been further supported by the observation that the presence of European HLA haplotypes is protective against obesity [14].

It is this reduced genetic selection, especially prominent in the last century, which has been noted to play a role not only in obesity but also in its pathological sequelae including metabolic disorders. A recent study of 118 countries has found a negative correlation between natural selection as measured by a "Biological State Index " and the incidence of diabetes mellitus[15]. Interestingly the relationship between reduced natural selection, after controlling for income and urbanization, was stronger for the prevalence of type 1 diabetes mellitus ($R^2$=0.55) than for type 2 diabetes mellitus ($R^2$=0.13) [15]. Such a result could be partly explained by an improvement in life expectancy, through an effective insulin treatment, fostering the genetic susceptibility to type 1 and to a lesser extent, type 2 diabetes. In line with this theory, an analysis of 65 loci associated with susceptibility to type 2 diabetes in samples of African, European, and East Asian ancestry, has reported a weak prevalence of protective alleles being selected in these populations without evidence for the evolutionary elimination of harmful genetic variants [16].

**Overview of Genome –wide association studies**

The recent Genome-wide association studies (GWAS) of genetic variants influencing body mass index (BMI) and glucose homeostasis, which primarily included data of individuals with European ancestry, have mapped approximately 100 independent single nucleotide polymorphisms (SNPs) that modulate the risk of type 2 diabetes and glycaemic related traits [7] [17] [18]. Interestingly, despite the arrival of Aboriginal people to Australia > 50,000 years ago a recent analysis of genetic markers from 402 individuals from the Australian Aboriginal community including 89 patients with type 2 diabetes, has identified common DNA variants influencing predisposition towards obesity and type 2 DM [19]. The genetic profiling of the Aboriginal population has identified genes and pathways associated with BMI and type 2 diabetes across multiple ethnicities [19] including MC4R for BMI [20] and IGF2BP2 for DM [21], TCF7L2 [22] or KCNJ11 [23]. Interestingly, the effect size of these SNPs was noted to be smaller in comparison with other studies and none of the reported associations between identified SNPs and BMI or diabetes risk have achieved genome-wide significance.

The analysis has provided stronger evidence for an association between Indigenous BMI and SNPs that lie in the intergenic region between SLC28A3 and NTRK2 on chromosome 9q21.33 as well as with genes PIK3C2G and CNTNAP2 although to a lesser degree respectively. NTRK2 has been previously established to regulate the BMI as well as it was linked with mood disorders [24]. In addition, NTRK2 has been shown to regulate the mammalian eating behaviour and energy balance downstream of the melanocortin-4 receptor (MC4R) through the brain-derived neurotrophic factor (BDNF) [25] [26].

Interestingly, the recent report identified Neanderthal haplotype-tagging SNPs in *NTRK2* gene which although downregulated in modern humans, may be associated with obesity in Aboriginal patients [27]. Further studies are required to determine the

regulatory mechanisms of *NTRK2* expression in Aboriginal population and whether these BMI associated loci are in linkage disequilibrium with the Neanderthal haplotype-tagging SNPs in *NTRK2* gene.

Furthermore, the results from the Aboriginal genome analysis have identified an association between a novel gene BCL9 and an increased diabetes risk. The BCL9 gene along with TCF7L2 modulates the function of WNT-signalling pathway [28] in the transcription of key incretins including a glucagon-like peptide 1 (GLP1) hormone which stimulates insulin secretion [29]. Additionally this first genome-wide analysis in an Australian Aboriginal population has implicated other diabetes susceptible genes including *KCNJ6*, *KCNA1*, and *GABRR1* which are known to influence pancreatic function[19].

The results from the First Genome-Wide Association Study have not pointed to susceptibility genes which could explain disproportionally higher (sixfold) risk for type 2 diabetes in ancient Aboriginal populations. Notably, recent genomic analysis reported that Aboriginal Australians first diverged from ancient Eurasian populations approximately 50 000 years ago and subsequently admixed with separate archaic populations including Denisovans and Neanderthals [30]. Therefore the future identification of the risk variants for abnormal glucogenesis may need to include genetic admixture from archaic populations [31] which has been implicated in the pathogenesis of diabetes in Mexican and Latino Americans populations [32].

## Impact of archaic genome on diseases of modern humans

Advances in DNA technology have only recently allowed us to study the intersections between ancient DNA, large-scale genomic data, and modern epidemiology. Recent genetic studies link predisposition to some diseases, including diabetes, to ancient communities suggesting historical interbreeding of modern humans with Neanderthals or Denisovans (an archaic hominin group from Siberia) [33]. Denisovans share a common ancestral population with Neanderthals, with up to 17% of the Denisova genome representing Neanderthal DNA [34]. Importantly, due to shared ancestral alleles it may be difficult to distinguish between Neanderthal and Denisovan admixture in genomic regions where these populations are not highly differentiated.

Genomes of modern Eurasians contain a small fraction (~1.5–4%) of DNA originating from Neanderthals [35]. Notably, a small portion of Melanesian, Papuan and Australian people derive 3–6% of their genes from Denisovans with an increase in allele sharing between the Denisovans and the Aboriginal Australians genome compared to other Eurasians and Africans [36] [37] [38]. Smaller amounts of Denisovan ancestry (0.2%) are also found in East Asia [39]. The approximately 4% contribution to the Indigenous Australian nuclear genome comes from an ancient hominin lineage yet to be identified [30].

The recent study, which analysed the contribution of common Neanderthal variants to over 1,000 electronic health record (EHR)-derived phenotypes in ~28,000 adults of European ancestry, reported that archaic admixture continues to influence disease risk in modern humans [33]. The study analyses confirmed the impact of Neanderthal DNA on neurological, psychiatric, immunological, and dermatological modern human phenotypes. In particular, the presence of Neanderthal alleles explained a significant percent of the risk for medical conditions including actinic keratosis ($P = 0.0059$), mood disorders ($P = 0.018$), depression ($P = 0.020$), obesity ($P = 0.030$) and seborrheic keratosis ($P = 0.045$).The Neanderthal haplotype

has been shown to influence hypercoagulable state, thiamine metabolism, urinary tract disorders and nicotine addiction.

Furthermore, the genetic analysis which examined 9.2 million single nucleotide polymorphisms (SNPs) in Mexican and Latin American population in 3,848 people with type 2 DM and 4,366 non-diabetic controls has identified SLC16A11 as a novel candidate gene for type 2 DM [32]. The effect size of the SLC16A11 haplotyp*e* was more prominent in younger (by 2.1years), leaner (BMI lower by 0.9 kg/m$^2$) people with type 2 DM and it was replicated in independent samples ($P$=1.1×10−4; OR=1.20). Interestingly this particular risk haplotype was not found in the previous genome-wide association as being rare or absent in samples from Europe and Africa, however being present in approximately half of Native Americans samples. The subsequent analysis of an archaic unpublished genome of a human from Denisova cave reported a presence of four missense SNPs at SLC16A11 gene indicating that the observed SLC16A11 haplotyp*e* has entered into modern humans from Neanderthals. Furthermore this Neanderthal genome sequence was nearly identical to 1000 genomes from the modern Native Americans who were homozygous for the 5 SNP haplotype. The presence of 5 SNP haplotype significantly increases the risk for type 2 DM contributing to approximately 20% (9.2%-29%) higher prevalence of type 2 diabetes mellitus in individuals with Native American ancestry. Notably the people affected by the 5SNP of SLC16A11 haplotype develop the 2 DM at the younger age with a lower BMI than non-carriers of SLC16A11. The researchers have proposed that SLC16A11 may influence the predisposition to diabetes through its effects on hepatic lipid metabolism affecting intracellular triacylglycerol levels.

Conversely, a recent study proposed that Neanderthal alleles resulted in selective advantage for the modern Europeans through their impact on the lipid catabolism [40]. The adaptive introgression of ancient allelles from Neanderthals and Denisovans to the modern human gene pool might have also conferred a substantial immune advantage the modern humans. The recent report described a cluster of three Toll-like receptors (*TLR6-TLR1-TLR10*), present in modern Asians and Europeans, which acquired their advantageous alleles by an admixture with archaic humans[41]. Furthermore, these introgressed alleles of immunity genes continue to have functional effects in modern humans as archaic-like alleles underlie differences in the expression of the Toll-like receptors (TLR) genes being essential for eliciting inflammatory and anti-microbial responses and for activating an adaptive immune response. They also play an important predisposing role to increased microbial resistance and to increased hypersensitivity to non-pathogenic allergens, resulting in allergic diseases in present-day people [42].

Moreover, the recent report has further confirmed that Neanderthal-inherited sequences continue to have measurable impacts on gene expression that contribute to variation in modern human phenotypes [27]. Although this analysis has emphasised the downregulation of several Neanderthal alleles in modern humans, it is important to point out that the NTRK2 gene, which contained a pair of adjacent Neanderthal tag SNPs, has recently been shown to be associated with obesity of Indigenous Australian patients

**Conclusions**

In summary, the Australian Indigenous people are disproportionately affected by type 2 diabetes mellitus compared to non-Indigenous Australians likely due to the complex interplay

between acquired and inherited risk factors. As current medical interventions have a modest effect in alleviating the burden of this potentially preventable condition more DNA based studies are needed to allow for more personalised therapies in an Aboriginal population. Furthermore, considering the recent evidence of the archaic genetic admixture, the possibility of the archaic legacy leading to an increased diabetes risk in the Indigenous population remains to be defined. Further genetic studies are required to identify the pathways from archaic human genomes which may advance our understanding of phenotypic differences between populations and provide novel targets for medical therapies.

**References:**

1. Zimmet, P.Z., et al., *Diabetes: a 21st century challenge.* Lancet Diabetes Endocrinol, 2014. **2**(1): p. 56-64.
2. Vos, T., et al., *Burden of disease and injury in Aboriginal and Torres Strait Islander Peoples: the Indigenous health gap.* Int J Epidemiol, 2009. **38**(2): p. 470-7.
3. Hoy, W.E., et al., *Quantifying the excess risk for proteinuria, hypertension and diabetes in Australian Aborigines: comparison of profiles in three remote communities in the Northern Territory with those in the AusDiab study.* Aust N Z J Public Health, 2007. **31**(2): p. 177-83.
4. Davis, T.M., et al., *Characteristics and outcome of type 2 diabetes in urban Aboriginal people: the Fremantle Diabetes Study.* Intern Med J, 2007. **37**(1): p. 59-63.
5. O'Dea, K., K.G. Rowley, and A. Brown, *Diabetes in Indigenous Australians: possible ways forward.* Med J Aust, 2007. **186**(10): p. 494-5.
6. Adegbija, O., W.E. Hoy, and Z. Wang, *Corresponding waist circumference and body mass index values based on 10-year absolute type 2 diabetes risk in an Australian Aboriginal community.* BMJ Open Diabetes Res Care, 2015. **3**(1): p. e000127.
7. Daniel, M., et al., *Diabetes and impaired glucose tolerance in Aboriginal Australians: prevalence and risk.* Diabetes Res Clin Pract, 2002. **57**(1): p. 23-33.
8. Cordain, L., et al., *Plant-animal subsistence ratios and macronutrient energy estimations in worldwide hunter-gatherer diets.* Am J Clin Nutr, 2000. **71**(3): p. 682-92.
9. Ilich, J.Z., et al., *Low-grade chronic inflammation perpetuated by modern diet as a promoter of obesity and osteoporosis.* Arh Hig Rada Toksikol, 2014. **65**(2): p. 139-48.
10. Eaton, S.B., M. Konner, and M. Shostak, *Stone agers in the fast lane: chronic degenerative diseases in evolutionary perspective.* Am J Med, 1988. **84**(4): p. 739-49.
11. Grantham, J.P., et al., *Modern diet and metabolic variance--a recipe for disaster?* Nutr J, 2014. **13**: p. 15.
12. Santos, J.L., et al., *Copy number polymorphism of the salivary amylase gene: implications in human nutrition research.* J Nutrigenet Nutrigenomics, 2012. **5**(3): p. 117-31.
13. Budnik, A. and M. Henneberg, *Worldwide Increase of Obesity Is Related to the Reduced Opportunity for Natural Selection.* PLoS One, 2017. **12**(1): p. e0170098.
14. Williams, D.R., et al., *Diabetes and glucose tolerance in New South Wales coastal Aborigines: possible effects of non-Aboriginal genetic admixture.* Diabetologia, 1987. **30**(2): p. 72-7.

15.     Ruhli, F., K. van Schaik, and M. Henneberg, *Evolutionary Medicine: The Ongoing Evolution of Human Physiology and Metabolism.* Physiology (Bethesda), 2016. **31**(6): p. 392-397.

16.     Ayub, Q., et al., *Revisiting the thrifty gene hypothesis via 65 loci associated with susceptibility to type 2 diabetes.* Am J Hum Genet, 2014. **94**(2): p. 176-85.

17.     Lyssenko, V., et al., *Clinical risk factors, DNA variants, and the development of type 2 diabetes.* N Engl J Med, 2008. **359**(21): p. 2220-32.

18.     McCarthy, M.I., *Genomics, type 2 diabetes, and obesity.* N Engl J Med, 2010. **363**(24): p. 2339-50.

19.     Anderson, D., et al., *First genome-wide association study in an Australian aboriginal population provides insights into genetic risk factors for body mass index and type 2 diabetes.* PLoS One, 2015. **10**(3): p. e0119333.

20.     Larsen, L.H., et al., *Prevalence of mutations and functional analyses of melanocortin 4 receptor variants identified among 750 men with juvenile-onset obesity.* J Clin Endocrinol Metab, 2005. **90**(1): p. 219-24.

21.     Hedbacker, K., et al., *Antidiabetic effects of IGFBP2, a leptin-regulated gene.* Cell Metab, 2010. **11**(1): p. 11-22.

22.     Lyssenko, V., et al., *Mechanisms by which common variants in the TCF7L2 gene increase risk of type 2 diabetes.* J Clin Invest, 2007. **117**(8): p. 2155-63.

23.     Qin, L.J., Y. Lv, and Q.Y. Huang, *Meta-analysis of association of common variants in the KCNJ11-ABCC8 region with type 2 diabetes.* Genet Mol Res, 2013. **12**(3): p. 2990-3002.

24.     Yilmaz, Z., et al., *The role of leptin, melanocortin, and neurotrophin system genes on body weight in anorexia nervosa and bulimia nervosa.* J Psychiatr Res, 2014. **55**: p. 77-86.

25.     Kernie, S.G., D.J. Liebl, and L.F. Parada, *BDNF regulates eating behavior and locomotor activity in mice.* EMBO J, 2000. **19**(6): p. 1290-300.

26.     Xu, B., et al., *Brain-derived neurotrophic factor regulates energy balance downstream of melanocortin-4 receptor.* Nat Neurosci, 2003. **6**(7): p. 736-42.

27.     McCoy, R.C., J. Wakefield, and J.M. Akey, *Impacts of Neanderthal-Introgressed Sequences on the Landscape of Human Gene Expression.* Cell, 2017. **168**(5): p. 916-927 e12.

28.     Kramps, T., et al., *Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex.* Cell, 2002. **109**(1): p. 47-60.

29.     Wagner, R., et al., *Untangling the interplay of genetic and metabolic influences on beta-cell function: Examples of potential therapeutic implications involving TCF7L2 and FFAR1.* Mol Metab, 2014. **3**(3): p. 261-7.

30.     Malaspinas, A.S., et al., *A genomic history of Aboriginal Australia.* Nature, 2016. **538**(7624): p. 207-214.

31.     Tobler, R., et al., *Aboriginal mitogenomes reveal 50,000 years of regionalism in Australia.* Nature, 2017. **544**(7649): p. 180-184.

32.     Consortium, S.T.D., et al., *Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico.* Nature, 2014. **506**(7486): p. 97-101.

33.     Simonti, C.N., et al., *The phenotypic legacy of admixture between modern humans and Neandertals.* Science, 2016. **351**(6274): p. 737-41.

34.     Pennisi, E., *Human evolution. More genomes from Denisova Cave show mixing of early human groups.* Science, 2013. **340**(6134): p. 799.

35.     Green, R.E., et al., *A draft sequence of the Neandertal genome.* Science, 2010. **328**(5979): p. 710-722.

36.     Reich, D., et al., *Genetic history of an archaic hominin group from Denisova Cave in Siberia.* Nature, 2010. **468**(7327): p. 1053-60.
37.     Reich, D., et al., *Denisova admixture and the first modern human dispersals into Southeast Asia and Oceania.* Am J Hum Genet, 2011. **89**(4): p. 516-28.
38.     Meyer, M., et al., *A high-coverage genome sequence from an archaic Denisovan individual.* Science, 2012. **338**(6104): p. 222-6.
39.     Prufer, K., et al., *The complete genome sequence of a Neanderthal from the Altai Mountains.* Nature, 2014. **505**(7481): p. 43-9.
40.     Khrameeva, E.E., et al., *Neanderthal ancestry drives evolution of lipid catabolism in contemporary Europeans.* Nat Commun, 2014. **5**: p. 3584.
41.     Dannemann, M., A.M. Andres, and J. Kelso, *Introgression of Neandertal- and Denisovan-like Haplotypes Contributes to Adaptive Variation in Human Toll-like Receptors.* Am J Hum Genet, 2016. **98**(1): p. 22-33.
42.     Akira, S., S. Uematsu, and O. Takeuchi, *Pathogen recognition and innate immunity.* Cell, 2006. **124**(4): p. 783-801.