

## A missense variant in *CREBRF* is associated with taller stature in Samoans.

**Short title:** *CREBRF* and height in Samoans

Jenna C. Carlson<sup>1,2\*</sup>, Samantha L. Rosenthal<sup>1\*</sup>, Emily M. Russell<sup>1</sup>, Nicola L. Hawley<sup>3</sup>, Guangyun Sun<sup>4</sup>, Hong Cheng<sup>4</sup>, Take Naseri<sup>5</sup>, Muagututi'a Sefuiva Reupena<sup>6</sup>, John Tuitele<sup>7</sup>, Ranjan Deka<sup>4</sup>, Stephen T. McGarvey<sup>8,9</sup>, Daniel E. Weeks<sup>1,2</sup>, Ryan L. Minster<sup>1</sup>.

1. Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA.
2. Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA.
3. Department of Chronic Disease Epidemiology, School of Public Health, Yale University, New Haven, CT, USA.
4. Department of Environmental Health, College of Medicine, University of Cincinnati, Cincinnati, OH, USA.
5. Ministry of Health, Government of Samoa, Apia, Samoa.
6. Lutia i Puava ae Mapu i Fagalele, Apia, Samoa.
7. Department of Public Health, Government of American Samoa, Pago Pago, American Samoa, USA.
8. International Health Institute and Department of Epidemiology, School of Public Health, Brown University, Providence, RI, USA.
9. Department of Anthropology, Brown University, Providence, RI, USA.

\*These authors contributed equally to this work

### Abstract

**Objectives:** Studies have demonstrated that rs373863828, a missense mutation in *CREBRF*, is associated with a number of anthropometric traits including body mass index (BMI), obesity, percent body fat, hip circumference, and abdominal circumference. Given the biological relationship between height and adiposity, we hypothesized that the effect of this variant on BMI might be due in part to a previously untested association of this variant with height.

**Methods:** We tested the hypothesis that minor allele of rs373863828 is associated with height in a Samoan population in two adult cohorts and in a separate cohort of children (age 5 - 18 years old) using linear mixed modeling.

**Results:** We found evidence of a strong relationship between rs373863828 and greater mean height in Samoan adults (0.77 cm greater average height for each copy of the minor allele) with the same direction of effect in Samoan children.

**Conclusions:** These results suggest that the missense variant rs373863828 in *CREBRF*, first identified through an association with larger BMI, may be related to an underlying biological mechanism affecting overall body size including stature.

**Keywords:** anthropometry, genetic association, diverse populations, height

## Introduction

Recently, we identified a novel association between a missense mutation (rs373863828, c.1370G>A, p.R457Q) in the gene for transcription factor *CREB3 regulatory factor (CREBRF)* and body mass index (BMI) in individuals of Polynesian ancestry (Krishnan et al., 2018; Minster et al., 2016). Despite fine-mapping and functional work suggesting the causal variant of this association is rs373863828 (Minster et al., 2016), the biological mechanism through which *CREBRF* operates to affect BMI is still unknown.

The rs373863828 variant is common in Samoans (minor allele frequency, MAF=0.259), but is very rarely seen in individuals without Pacific Island ancestry (gnomAD MAF for all populations= $3.53 \times 10^{-5}$ ; accessed 12 April 2019). However, this variant is polymorphic and associated with BMI in other Pacific Island populations including New Zealand Māori, Pukapukans, and other Polynesians (Krishnan et al., 2018).

This missense variant is associated with a constellation of anthropometric traits including obesity, percent body fat, hip circumference, and abdominal circumference (Minster et al., 2016). Of those traits, rs373863828 was most significantly associated with BMI, with an estimated 1.36–1.45 kg/m<sup>2</sup> ( $p_{\text{meta}} = 1.4 \times 10^{-20}$ ) higher mean BMI per copy of the alternate (A) allele (Minster et al., 2016). A meta-analysis of Polynesians living in New Zealand showed an estimated 1.34 kg/m<sup>2</sup> ( $p_{\text{meta}} = 1.60 \times 10^{-5}$ ) higher mean BMI per copy of the alternate allele (Krishnan et al., 2018).

BMI is often used as a measure of body adiposity independent of height, although BMI is not consistently a stature-independent measure of adiposity, especially in children and adolescents (Benn, 1971; Cole, 1986; J. C. Wells, 2014; J. Wells & Cole, 2002). Thus, we hypothesized that rs373863828, in addition to having a strong association with BMI, could also be associated with height. We sought to characterize this potential relationship in Samoan adults and children.

## Methods

### *Study Sample and Genotyping*

We used three cohorts in this study: an adult discovery cohort of 3,077 Samoans, an adult replication cohort of 2,103 Samoans and American Samoans, and a cohort of 409 Samoan children (Table 1). For all cohorts, height was recorded using a portable GPM anthropometer (Pfister Imports, New York, NY) and rs373863828 was genotyped using a TaqMan real-time PCR assay (Applied Biosystems), as previously described (Minster et al., 2016). These studies were approved by the Health Research Committee of the Samoan Ministry of Health, the American Samoan Department of Health Institutional Review Board (IRB; for Samoa and American Samoa-based studies respectively) and the Brown University IRB. All participants gave written informed consent.

### Discovery Cohort

The discovery cohort of 3,077 adults of Samoan ancestry, 23–70 years old, was drawn from a population-based sample recruited from 33 villages located across the two major islands of the Independent State of Samoa (hereinafter Samoa) (Table 1). The sample selection, data collection methods, phenotyping and extensive quality control using genome-wide genotyping data has been reported previously (Hawley et al., 2014; Minster et al., 2016).

### Replication and Children Cohorts

The replication cohort totaling 2,103 adults was taken from two studies from Samoa and the United States Territory of American Samoa (hereinafter American Samoa) (Table 1). The first sample totaling 1,020 participants was derived from a longitudinal study of adiposity and cardiovascular disease risk factors among adults from American Samoa and Samoa recruited between 1990 and 1995, using stature as recorded at baseline. Detailed descriptions of the sampling, recruitment, and phenotyping have been reported previously (Chin-Hong & McGarvey, 1996; Galanis, McGarvey, Sobal, Bausserman, & Levinson, 1995; McGarvey, Levinson, Bausser-Man, Galanis, & Hornick, 1993). Due to lack of genome-wide marker data on these samples, we were unable to infer relatedness, and so the participants were treated as unrelated in the analyses.

The second sample of 1,083 adults from American Samoa and Samoa, recruited in 2002–2003, was drawn from an extended family-based genetic linkage study of cardiometabolic traits (Åberg, Dai, Sun, et al., 2009; Åberg et al., 2008; Åberg, Dai, Viali, et al., 2009). Probands and relatives were unselected for obesity or related phenotypes, and all individuals self-reported Samoan ancestry. The recruitment process, criteria used for inclusion in this study, and phenotyping have been described in detail previously (Dai et al., 2007, 2008). The child cohort comprises 409 Samoan children, 5–18 years old, who are members of the recruited families. Relatedness was derived from known family structures which have been verified to be consistent with relatedness estimates derived using genome-wide microsatellite markers (Dai et al., 2007, 2008).

### *Statistical Analysis*

The association between rs373863828 and height was assessed separately for adults and children to account for differences in height during maturation and adulthood. In adults, the association was tested for separately in the discovery cohort of 3,077 adults and in the replication cohort of 2,103 adults.

In the discovery cohort, height residuals were calculated adjusting for the fixed effects age, age<sup>2</sup>, sex, age × sex interaction, and age<sup>2</sup> × sex interaction. An empirical kinship matrix was generated with KING (Manichaikul et al., 2010) and PC-Relate (Conomos et al., 2019) to distinguish recent and distant relatedness in the sample using markers from the Genome-Wide Human SNP 6.0 array (Affymetrix). The percent of total variation in height and BMI was estimated using the coefficient of determination ( $R^2$ ) from a simple linear regression of each phenotype as predicted by an additive effect of rs373863828.

In the replication cohort, height residuals were calculated adjusting for the fixed effects of study (1994-95 or 2002-03 recruitment), polity (American Samoan or Samoa), age, age<sup>2</sup>, sex, age × sex, and age<sup>2</sup> × sex. Kinship was estimated from pedigree information with the kinship2 package in R (Therneau & Sinnwell, 2014).

In children, the association was measured in the 409 children from the replication cohort. First, age- and sex-specific z-scores of height were calculated for children based on World Health Organization (WHO) References 2007 (de Onis et al., 2007) (Table 1). The resulting height z-scores were adjusted for the effects of age, age<sup>2</sup>, sex, age × sex, and age<sup>2</sup> × sex. Kinship was estimated from pedigree information with the kinship2 package (Therneau & Sinnwell, 2014).

Within each cohort, age was centered to the overall mean for that cohort to avoid collinearity among the effects. Additionally, within each cohort, height (or height z-score) residuals were examined, and no evidence of non-normality was found (results not shown). These residuals were then used to assess the relationship between rs373863828 and height via linear mixed models adjusting for kinship as a random effect using the lme4 function in R (Therneau, 2018). The resulting effects from the two analyses in adults were combined using an inverse-variance effect-based fixed-effect meta-analysis as implemented in the rmeta package (Lumley, 2018).

Sex-specific effects of rs373863828 on height were also examined through a test of statistical interaction between sex and rs373863828 genotype in each cohort.

## Results

Association between rs373863828 and height was examined using linear mixed models in a discovery cohort of 3,077 adults and in a replication cohort of 2,103 adults, adjusting in the replication cohort analyses for the average effect of polity and study to account for the observed differences. Each cohort demonstrated significant association between the A allele of rs373863828 and taller average height ( $\beta_{\text{discovery}} = 0.60$ ; 95% CI [0.30 - 0.91];  $p_{\text{discovery}} = 1.00 \times 10^{-4}$ ;  $\beta_{\text{replication}} = 1.06$ ; 95% CI [0.65 - 1.46];  $p_{\text{replication}} = 2.80 \times 10^{-7}$ ; Figure 1). Fixed-effects meta-analysis yielded a weighted effect estimate of 0.77 cm taller average height per copy of the A minor allele (95% CI [0.53 - 1.01];  $p_{\text{meta}} = 5.56 \times 10^{-10}$ ). The Cochran test for heterogeneity of effects was not significant ( $p = 0.08$ ).

Notably, 0.18% of the total variation in unadjusted height was explained by an additive effect of rs373863828, compared to 1.95% of the total variation in unadjusted BMI.

The association between rs373863828 and taller stature was also observed in 409 children, with an effect estimate of 0.38 taller average height z-score per copy of the A allele (95% CI [0.23 - 0.53];  $p_{\text{child}} = 1.39 \times 10^{-6}$ ; Figure 1).

We also examined whether the effect of rs373863828 on average height differed for males and females in each cohort by testing for a statistical interaction between sex and genotype. We did not observe a significant interaction effect ( $\beta_{\text{discovery}} = 0.59$ ,  $p_{\text{discovery}} = 0.04$ ;  $\beta_{\text{replication}} = -0.05$ ,  $p_{\text{replication}} = 0.09$ ,  $\beta_{\text{meta}} = 0.38$ ,  $p_{\text{meta}} = 0.11$ ).

## Discussion

We have established an association of the minor allele (A) of rs373863828 with greater mean height in the Samoan population. This association confers an estimated taller mean height of 0.77 cm per copy of the minor allele in adults and an estimated 0.38 standard deviations taller height z-score in children, irrespective of sex. This observed effect was seen across two cohorts of adults recruited at three different time periods between 1990 and 2010.

A similar association between rs373863828 and height was observed in a cohort of Māori and Pacific adults from New Zealand ( $\beta = 1.25$  cm;  $p = 3.90 \times 10^{-6}$ ) (T. R. Merriman, personal communication, April 15, 2019). Additionally, association between rs373863828 and early childhood growth between ages 2 and 4 years old has been previously reported in a separate cohort including self-identified Māori and Pacific from New Zealand (Berry et al., 2018b, 2018a; Major et al., 2018).

This association, of rs373863828 with greater average height, is consistent in both adults ages 23–70 and children ages 5–18 years old, suggesting that the effect of rs373863828 on height may occur early in life. However, it is important to note that these data are cross-sectional and unadjusted for pubertal status, so future longitudinal studies will be required to determine if the effect of this variant on height can be attributed to overall growth trajectory or increased growth during a particular developmentally crucial period.

In addition to being one of the largest effect sizes reported for a common variant ( $\beta = 0.77$ , MAF = 0.259) on height, rs373863828 is a Pacific Island-specific polymorphism. Combined with previous association studies connecting rs373863828 with greater average adiposity and lower risk of diabetes, these results suggest that rs373863828 may operate broadly as a body size variant, affecting overall growth as opposed to a singular growth dimension. It is worth noting that this variant explains more of the variation in BMI than height (1.95% versus 0.18% respectively). However, the function of *CREBRF* is minimally characterized, so the mechanisms by which rs373863828 affects these phenotypes are currently unclear. Individual studies have indicated that this transcription factor impacts glucocorticoid signaling, autophagy, unfolded protein response, and starvation resistance, as well as murine decidualization, embryo implantation, and maternal behavior (Audas, Li, Liang, & Lu, 2008; Li et al., 2016; Martyn et al., 2012; Minster et al., 2016; Xue et al., 2016; Yang et al., 2013). Future studies should aim to refine the effect of this variation on the related phenotypes of height, weight, and BMI in other Pacific Island populations, as well as other anthropometric traits. Functional studies will be crucial to determining the function and targets of the *CREBRF* gene.

## **Acknowledgements**

The authors would like to thank the Samoan participants of the study, local village authorities, and the many Samoan and other field workers over the years. We acknowledge the Samoan Ministry of Health, the Samoa Bureau of Statistics, and the American Samoan Department of Health for their support of this research. We give particular thanks to two research assistants, Melania Selu and Vaimoana Lupematasila, who contributed to the 2010 recruitment and continue to assist us in our work in Samoa. This work was funded by the National Institute of Health grants R01-HL093093 (STM), R01-HL133040 (RLM), R01-AG09375 (STM), R01-HL52611 (MI Kamboh), R01-DK59642 (STM), and R01-DK55406 (RD). Genotyping was performed in the Core Genotyping Laboratory at the University of Cincinnati, funded by National Institutes of Health grant P30-ES006096 (SM Ho).

## **Author Contributions**

JCC and SLR performed the statistical analysis, created the figures and tables, and wrote the manuscript, with guidance from RLM, STM, NLH, and DEW. EMR contributed to the statistical analysis under the guidance of RLM. NLH led the field work data collection and phenotype analyses with guidance from STM. GS led and directed genotyping experiments (using the Affymetrix 6.0 chip) and assay development for validation and replication (using the TaqMan platform) with guidance from RD. MSR facilitated fieldwork in Samoa. TN contributed to the discussion of the public health implications of the findings. All authors contributed to this work, discussed the results, and critically reviewed and revised the manuscript.

## **Conflict of Interest Statement**

DEW, NLH, RD, STM, and RLM are co-inventors on the United States Patent Application 20180245155, "Compositions and Methods for Identifying Genetic Predisposition to Obesity and for Enhancing Adipogenesis".

## References

- Åberg, K., Dai, F., Sun, G., Keighley, E. D., Indugula, S. R., Roberts, S. T., ... McGarvey, S. T. (2009). Susceptibility Loci for Adiposity Phenotypes on 8p, 9p, and 16q in American Samoa and Samoa. *Obesity*, *17*(3), 518–524. <https://doi.org/10.1038/oby.2008.558>
- Åberg, K., Dai, F., Sun, G., Keighley, E., Indugula, S. R., Bausserman, L., ... McGarvey, S. T. (2008). A genome-wide linkage scan identifies multiple chromosomal regions influencing serum lipid levels in the population on the Samoan islands. *Journal of Lipid Research*, *49*(10), 2169–2178. <https://doi.org/10.1194/jlr.M800194-JLR200>
- Åberg, K., Dai, F., Viali, S., Tuitele, J., Sun, G., Indugula, S. R., ... McGarvey, S. T. (2009). Suggestive linkage detected for blood pressure related traits on 2q and 22q in the population on the Samoan islands. *BMC Medical Genetics*, *10*(1), 107. <https://doi.org/10.1186/1471-2350-10-107>
- Audas, T. E., Li, Y., Liang, G., & Lu, R. (2008). A novel protein, Luman/CREB3 recruitment factor, inhibits Luman activation of the unfolded protein response. *Molecular and Cellular Biology*, *28*(12), 3952–3966. <https://doi.org/10.1128/MCB.01439-07>
- Benn, R. T. (1971). Some mathematical properties of weight-for-height indices used as measures of adiposity. *British Journal of Preventive & Social Medicine*, *25*(1), 42–50.
- Berry, S. D., Walker, C. G., Ly, K., Snell, R. G., Atatoa Carr, P. E., Bandara, D., ... Grant, C. C. (2018a). Re: “Widespread prevalence of a CREBRF variant amongst Māori and Pacific children is associated with weight and height in early childhood.” *International Journal of Obesity*, *42*(7), 1392–1393. <https://doi.org/10.1038/s41366-018-0026-0>
- Berry, S. D., Walker, C. G., Ly, K., Snell, R. G., Atatoa Carr, P. E., Bandara, D., ... Grant, C. C. (2018b). Widespread prevalence of a CREBRF variant amongst Māori and Pacific children is associated with weight and height in early childhood. *International Journal of Obesity*, *42*(4), 603–607. <https://doi.org/10.1038/ijo.2017.230>
- Chin-Hong, P. V., & McGarvey, S. T. (1996). Lifestyle incongruity and adult blood pressure in Western Samoa. *Psychosomatic Medicine*, *58*(2), 131–137. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8849629>
- Cole, T. J. (1986). Weight/height<sup>p</sup> compared to weight/height<sup>2</sup> for assessing adiposity in childhood: influence of age and bone age on p during puberty. *Annals of Human Biology*, *13*(5), 433–451. <https://doi.org/10.1080/03014468600008621>
- Conomos, M. P., Gogarten, S. M., Brown, L., Chen, H., Rice, K., Sofer, T., ... Yu, C. (2019). GENESIS: GENetic ESTimation and Inference in Structured samples (GENESIS): Statistical methods for analyzing genetic data from samples with population structure and/or relatedness. Retrieved from <https://github.com/UW-GAC/GENESIS>
- Dai, F., Keighley, E. D., Sun, G., Indugula, S. R., Roberts, S. T., Åberg, K., ... McGarvey, S. T. (2007). Genome-wide scan for adiposity-related phenotypes in adults from American Samoa. *International Journal of Obesity*, *31*(12), 1832–1842. <https://doi.org/10.1038/sj.ijo.0803675>
- Dai, F., Sun, G., Åberg, K., Keighley, E. D., Indugula, S. R., Roberts, S. T., ... McGarvey, S. T. (2008). A Whole Genome Linkage Scan Identifies Multiple Chromosomal Regions Influencing Adiposity-Related Traits among Samoans. *Annals of Human Genetics*, *72*(6), 780–792. <https://doi.org/10.1111/j.1469-1809.2008.00462.x>

- de Onis, M., Onyango, A. W., Borghi, E., Siyam, A., Nishida, C., & Siekmann, J. (2007). Development of a WHO growth reference for school-aged children and adolescents. *Bulletin of the World Health Organization*, 85(9), 660–667. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18026621>
- Galanis, D. J., McGarvey, S. T., Sobal, J., Bausserman, L., & Levinson, P. D. (1995). Relations of body fat and fat distribution to the serum lipid, apolipoprotein and insulin concentrations of Samoan men and women. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, 19(10), 731–738. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8589767>
- Hawley, N. L., Minster, R. L., Weeks, D. E., Viali, S., Reupena, M. S., Sun, G., ... McGarvey, S. T. (2014). Prevalence of adiposity and associated cardiometabolic risk factors in the samoan genome-wide association study. *American Journal of Human Biology*, 26(4), 491–501. <https://doi.org/10.1002/ajhb.22553>
- Krishnan, M., Major, T. J., Topless, R. K., Dewes, O., Yu, L., Thompson, J. M. D., ... Merriman, T. R. (2018). Discordant association of the CREBRF rs373863828 A allele with increased BMI and protection from type 2 diabetes in Māori and Pacific (Polynesian) people living in Aotearoa/New Zealand. *Diabetologia*, 61(7), 1603–1613. <https://doi.org/10.1007/s00125-018-4623-1>
- Li, X., Lin, P., Chen, F., Wang, N., Zhao, F., Wang, A., & Jin, Y. (2016). Luman recruiting factor is involved in stromal cell proliferation during decidualization in mice. *Cell and Tissue Research*, 365(2), 437–447. <https://doi.org/10.1007/s00441-016-2392-z>
- Lumley, T. (2018). rmeta: Meta-Analysis. R package version 3.0. Retrieved from <https://cran.r-project.org/package=rmeta>
- Major, T. J., Krishnan, M., Topless, R. K., Dewes, O., Thompson, J., Zoysa, J. de, ... Merriman, T. R. (2018). Re: “Widespread prevalence of a CREBRF variant among Māori and Pacific children is associated with weight and height in early childhood.” *International Journal of Obesity*, 42(7), 1389–1391. <https://doi.org/10.1038/s41366-018-0025-1>
- Manichaikul, A., Mychaleckyj, J. C., Rich, S. S., Daly, K., Sale, M., & Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. *Bioinformatics*, 26(22), 2867–2873. <https://doi.org/10.1093/bioinformatics/btq559>
- Martyn, A. C., Choleris, E., Gillis, D. J., Armstrong, J. N., Amor, T. R., McCluggage, A. R. R., ... Lu, R. (2012). Luman/CREB3 recruitment factor regulates glucocorticoid receptor activity and is essential for prolactin-mediated maternal instinct. *Molecular and Cellular Biology*, 32(24), 5140–5150. <https://doi.org/10.1128/MCB.01142-12>
- McGarvey, S. T., Levinson, P. D., Bausser-Man, L., Galanis, D. J., & Hornick, C. A. (1993). Population change in adult obesity and blood lipids in American Samoa from 1976–1978 to 1990. *American Journal of Human Biology*, 5(1), 17–30. <https://doi.org/10.1002/ajhb.1310050106>
- Minster, R. L., Hawley, N. L., Su, C.-T., Sun, G., Kershaw, E. E., Cheng, H., ... McGarvey, S. T. (2016). A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nature Genetics*, 48(9), 1049–1054. <https://doi.org/10.1038/ng.3620>
- Therneau, T. M. (2018). coxme: Mixed Effects Cox Models. R package version 2.2-10. Retrieved from <https://cran.r-project.org/package=coxme>
- Therneau, T. M., & Sinnwell, J. (2014). kinship2: Pedigree functions. R package version 1.6.4.



Retrieved from <https://cran.r-project.org/package=kinship2>

- Wells, J. C. (2014). Commentary: The paradox of body mass index in obesity assessment: not a good index of adiposity, but not a bad index of cardio-metabolic risk. *International Journal of Epidemiology*, 43(3), 672–674. <https://doi.org/10.1093/ije/dyu060>
- Wells, J., & Cole, T. (2002). Adjustment of fat-free mass and fat mass for height in children aged 8 y. *International Journal of Obesity*, 26(7), 947–952. <https://doi.org/10.1038/sj.ijo.0802027>
- Xue, H., Zhang, J., Guo, X., Wang, J., Li, J., Gao, X., ... Li, G. (2016). CREBRF is a potent tumor suppressor of glioblastoma by blocking hypoxia-induced autophagy via the CREB3/ATG5 pathway. *International Journal of Oncology*, 49(2), 519–528. <https://doi.org/10.3892/ijo.2016.3576>
- Yang, Y., Jin, Y., Martyn, A. C., Lin, P., Song, Y., Chen, F., ... Wang, A. (2013). Expression pattern implicates a potential role for luman recruitment factor in the process of implantation in uteri and development of preimplantation embryos in mice. *The Journal of Reproduction and Development*, 59(3), 245–251. <https://doi.org/10.1262/JRD.2012-137>

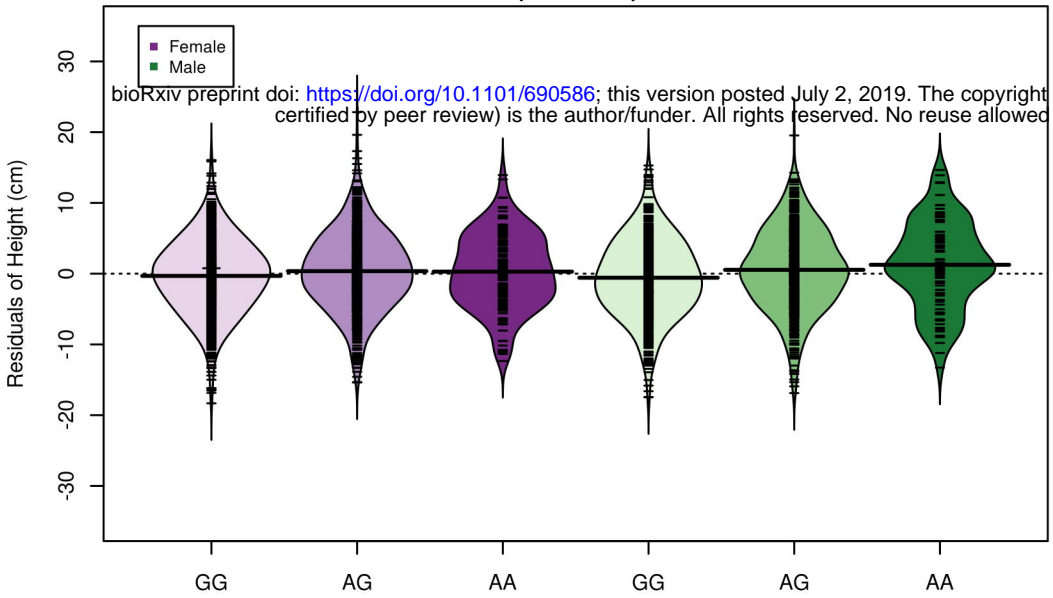
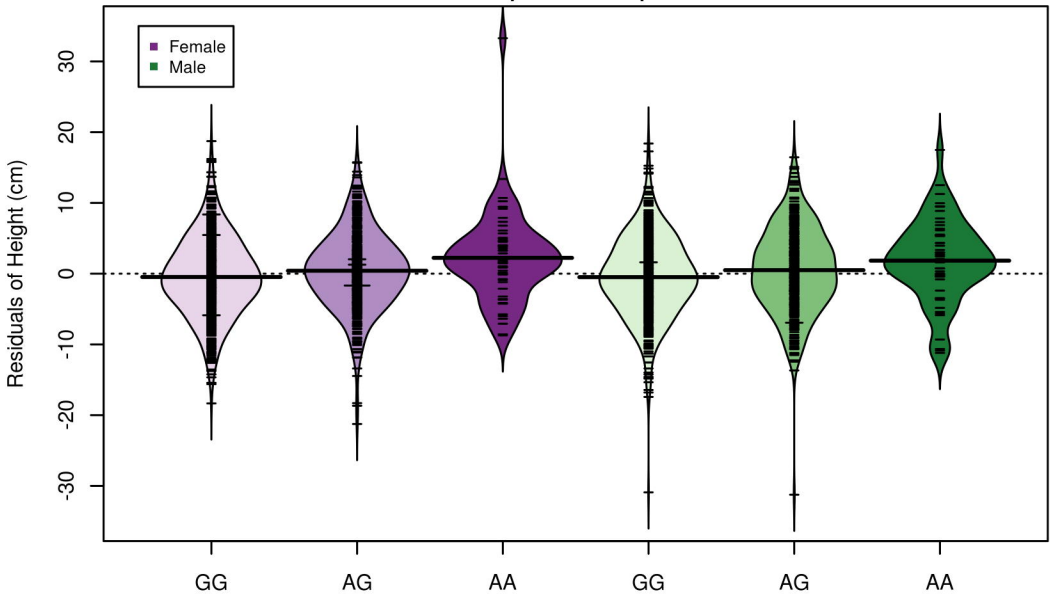
## Tables

**Table 1. Sample characteristics for males and females in each cohort.** Mean (SD) is given for age, height, and BMI for adults and children. Additionally in children, mean (SD) is given for age- and sex-specific height and BMI z-scores based on World Health Organization criteria.

	<i>n</i>		Age (years)		Height (cm)		BMI (kg/m <sup>2</sup> )	
	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>
<b>Discovery (<i>n</i> = 3,077)</b>								
<i>2010 Samoa</i>	1,839	1,238	44.63 (11.12)	45.40 (11.49)	160.94 (5.51)	171.75 (5.91)	34.86 (6.84)	31.31 (5.93)
<b>Replication (<i>n</i> = 2,103)</b>								
<i>1990-1995 Samoa</i>	1,125	978	41.36 (13.76)	40.69 (13.53)	161.24 (5.98)	171.97 (6.23)	34.12 (7.56)	31.01 (6.43)
<i>1990-1995 American Samoa</i>	316	291	39.08 (9.07)	38.63 (9.13)	159.82 (5.45)	170.81 (6.18)	30.94 (5.34)	28.93 (4.85)
<i>2002-2003 Samoa</i>	225	188	39.21 (10.52)	40.41 (9.94)	161.63 (5.87)	172.94 (5.94)	35.97 (6.97)	33.91 (5.88)
<i>2002-2003 American Samoa</i>	248	245	44.02 (16.95)	40.91 (16.31)	160.64 (5.99)	170.96 (6.51)	33.25 (7.68)	28.80 (5.38)
	336	254	42.97 (16.00)	43.04 (16.48)	162.75 (6.18)	173.57 (5.79)	36.53 (8.43)	33.36 (7.60)
					Height (cm) Height z-score		BMI (kg/m <sup>2</sup> ) BMI z-score	
	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>	<i>female</i>	<i>male</i>
<b>Children (<i>n</i> = 409)</b>								
<i>2002-2003 Samoa</i>	220	189	11.64 (3.50)	11.28 (3.50)	143.96 (16.65)	143.51 (18.22)	20.09 (4.19)	19.11 (3.50)
					-0.45 (1.18)	0.07 (1.22)	0.79 (1.03)	0.46 (0.93)

## Figure Legends

**Figure 1. Beanplot displaying height residuals by rs373863828 genotype (GG, GA, AA) across sex and cohort.** (A) Height residuals (cm) for the discovery cohort were adjusted for age, age<sup>2</sup>, sex, age × sex, and age<sup>2</sup> × sex. (B) Height residuals (cm) for the replication cohort were adjusted for recruitment year, polity, age, age<sup>2</sup>, sex, age × sex, and age<sup>2</sup> × sex. (C) Age- and sex-specific height z-scores based on World Health Organization criteria were adjusted additionally for age, age<sup>2</sup>, sex, age × sex, and age<sup>2</sup> × sex. In each plot, individual observations are indicated with tick mark and the mean height residuals within each stratum is given by the black horizontal bar. Additionally, the minor allele frequency (MAF) of rs373863828, as well as the effect estimates ( $\beta$ ) and  $p$  values from a linear mixed model testing the additive association between rs373863828 and height residuals are given within each cohort.

**A****Discovery Cohort**MAF = 0.276,  $\beta = 0.6$ ,  $p = 1.00 \times 10^{-4}$ **B****Replication Cohort**MAF = 0.262,  $\beta = 1.06$ ,  $p = 2.80 \times 10^{-7}$ **C****Children**MAF = 0.242,  $\beta = 0.38$ ,  $p = 1.39 \times 10^{-6}$ 