Deep Learning Derived Adipocyte Size Reveals Adipocyte Hypertrophy is under Genetic Control

Emil Jørsboe[1,2,3,4], Phil Kubitz[5], Julius Honecker[5], Andrea Flaccus[6], Dagmar Mvondo[6], Matthias Raggi[7], Torben Hansen[3], Hans Hauner[5,8], Matthias Blüher[9,10], Philip D. Charles[1,11,12,+], Cecilia M. Lindgren[1,2,13,+], Christoffer Nellåker[1,14,+], Melina Claussnitzer[13,15,16,+]

1. Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, United Kingdom

2. Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

3. Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

4. Center for Liver Research, Department of Gastroenterology and Hepatology, Odense University Hospital, Odense, Denmark

5. Technical University of Munich, Else Kröner-Fresenius-Center for Nutritional Medicine, Chair of Nutritional Medicine, School of Life Sciences, Gregor-Mendel-Straße 2, 85354, Freising-Weihenstephan, Germany

6. Institute of Nutritional Sciences, University of Hohenheim, Stuttgart, Germany

7. Department of General and Visceral Surgery, Karl-Olga-Krankenhaus, Stuttgart, Germany

8. Institute for Nutritional Medicine, School of Medicine, Technical University of Munich, Georg-Brauchle-Ring 62, 80992, Munich, Germany

9. Helmholtz Institute for Metabolic, Obesity and Vascular Research (HI-MAG) of the Helmholtz Zentrum München at the University of Leipzig and University Hospital Leipzig, Leipzig, Germany

10. Medical Department III - Endocrinology, Nephrology, Rheumatology, University of Leipzig Medical Center, Leipzig, Germany

11. Nuffield Department of Medicine, University of Oxford; Oxford, United Kingdom

12. Target Discovery Institute, Centre for Medicines Discovery, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

13. Broad Institute of MIT and Harvard, Medical and Population Genetics Program & Type 2 Diabetes Systems Genomics Initiative, Cambridge, MA, USA

14. Nuffield Department of Women's & Reproductive Health, University of Oxford, Oxford, United Kingdom

15. Novo Nordisk Foundation Center for Genomic Mechanisms of Disease, Broad Institute of MIT and Harvard, Cambridge, MA, USA

16. Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Corresponding authors:

Christoffer Nellåker: christoffer.nellaker@wrh.ox.ac.uk, Melina Claussnitzer: melina@broadinstitute.org

+ = Shared senior authors

Abstract
Fat distribution and macro structure of white adipose tissue are important factors in predicting obesity-associated diseases, but cellular microstructure of white adipose tissue has been less explored. To investigate the relationship between adipocyte size and obesity-related traits, and their underlying disease-driving genetic associations, we performed the largest study of automatic adipocyte phenotyping linking histological measurements and genetics to date. We introduce deep learning based methods for scalable and accurate semantic segmentation of subcutaneous and visceral adipose tissue histology samples (N=2,667) across 5 independent cohorts, including data from 9,000 whole slide images, with over 27 million adipocytes. Estimates of mean size of adipocytes were validated against Glastonbury et al. 2020. We show that adipocyte hypertrophy correlates with an adverse metabolic profile with increased levels of leptin, fasting plasma glucose, glycated hemoglobin and triglycerides, and decreased levels of adiponectin and HDL cholesterol. We performed the largest GWAS ($N_{Subcutaneous}$ = 2066, $N_{Visceral}$ = 1878) and subsequent meta-analysis of mean adipocyte area, and find two genome-wide significant loci (rs73184721, rs200047724) associated with increased 95%-quantile adipocyte size in respectively visceral and subcutaneous adipose tissue. Stratifying by sex, in females we find two genome-wide significant loci, with one variant (rs140503338) associated with increased mean adipocyte size in subcutaneous adipose tissue, and the other (rs11656704) is associated with decreased 95%-quantile adipocyte size in visceral adipose tissue.

Introduction

Obesity is a rapidly growing global health care problem. Individuals with obesity are characterised by an increased risk of developing a range of diseases, including type 2 diabetes, coronary artery disease, non-alcoholic fatty liver disease (1,2). Recent large-scale genome-wide association studies (GWAS) of body fat distribution, fasting insulin levels and type 2 diabetes highlight white adipose tissue as a key tissue in which disease-associated variants manifest their effect (3).

During development of obesity, white adipose tissue expands due to a combination of increasing adipocyte size (hypertrophy) and number of adipocytes (hyperplasia). While white adipose tissue expansion through hyperplasia is not associated with obesity-associated metabolic complications, hypertrophy is associated with fat accumulation in liver, skeletal muscle, and heart tissue, local inflammation (4), cardiometabolic risk (4) and impaired glucose metabolism (5). Expansion of white adipose tissue in adults, is primarily characterised by hypertrophy (6). In individuals with obesity and adipocyte hyperplasia, the storage capacity of triglycerides is limited, and further caloric overload leads to fat accumulation in other tissues (such as the liver, skeletal muscle, and heart) and in visceral adipose depots, an event commonly defined as “lipotoxicity” (4). Enlargement of adipocytes triggers low-grade chronic inflammation, insufficient angiogenesis, and excessive collagen deposition, which further lead to dysfunctional adipokine release (7). Adipocyte hypertrophy in visceral adipose tissue correlates stronger with insulin resistance or diabetes than in subcutaneous adipose tissue (8). Furthermore, solely in visceral adipose tissue, adipocyte size is associated with glucose intolerance and insulin resistance independently of body mass index (BMI) (8).

Obesity and fat distribution are highly heritable traits (9,10), and recent large-scale genetic association studies have identified well over 1,000 independent genetic risk loci for fat mass and distribution related traits. However, the disease-driving cellular processes are mostly still unclear. We hypothesized that adipocyte hypertrophy is one such disease-driving phenotype. To date, there is a lack of large-scale, well-powered GWAS mapping of the genetic determinants of adipocyte morphology (11,12), and as such the underlying genes and regulatory pathways involved in the size of adipocytes are yet to be determined (13). The limiting factor is the lack of tools that allow for scalable and affordable mapping of adipocyte size.

White adipose tissue histology samples can be used for studying adipocyte morphology, however manual annotation is cumbersome and extremely slow, and software for image-based measurements of adipocyte measurement software has been too inefficient to apply to large datasets. We previously developed a deep learning approach, Adipocyte U-Net, (14) capable of efficiently estimating the size of adipocytes across thousands of white adipose tissue histology samples. In that study, adipocyte size estimates were validated against two established (non deep-learning) methods: Adiposoft (15) and CellProfiler (16). Ultimately, that study did not identify any significant genetic associations, however this may have been due to limited statistical power (sample size around 1,000 individuals), and heterogeneity of the cohorts.

The strategy used for the model in Glastonbury et al., 2020 was one of splitting whole slides images (WSIs) into subimages (referred to as tiles), and only using tiles containing adipocytes for segmentation of adipocytes, thereby discarding a lot of tiles and not segmenting all adipocytes. Furthermore, adipocytes on the edge of a tile were discarded from the segmentation results. The authors note that size estimation of their method is on average smaller, compared to the area estimates obtained by Adiposoft and CellProfiler, which they attributed to a difference in filtering criteria. We will explore in this paper if the approach used in Glastonbury et al., 2020 leads to a bias in estimation of size of adipocytes, as we expect that larger adipocytes have a higher probability of touching an edge of a tile.

Methods

Training the deep learning model

We implemented a U-net (17) based model in PyTorch (18), for semantic segmentation of adipocytes, in WSIs. The model is based on the design in (14), with added dilations for the lowest resolution layer. The model takes 1024x1024 pixel RGB images as input. A reference pixel size of 0.2500 µm was selected for this model.

The model was trained with an ADAM optimizer (19) from PyTorch, using an initial learning rate of 0.0001, decaying by a factor of 0.5 per 30 epochs, and a batch size of 2. 200 tiles were shown to the model per epoch, and the model was trained for 200 epochs. The loss function was a dice binary cross-entropy loss (half dice coefficient and half cross entropy), and a dice coefficient metric was used to assess performance. We applied early stopping, storing the weights with the lowest validation loss.

$$Dice\ loss\ = 1 - \frac{2\sum_{n=1}^{N} p_n r_n + \epsilon}{\sum_{n=1}^{N} p_n + \sum_{n=1}^{N} r_n + \epsilon}$$

$$Binary\ cross\ entropy\ loss\ = \ -\frac{1}{N}\sum_{n=1}^{N} r_n log(p_n) + (1 - r_n) log(1 - p_n)$$

$$loss\ = \ 0.5(Dice\ loss) + 0.5(Binary\ cross\ entropy\ loss)$$

Where p_n is the predicted probabilities, r_n is the ground truth labels of the *n*th pixel *N* is the total number of pixels and ϵ is a term to prevent division by 0.

During training each input was normalised to Z-scores using the mean and standard deviation from the training data, and augmentations from the albumentations (20) python library were applied, selecting from horizontal flip, random rotate 90 degrees, random brightness contrast, blurring and gaussian noise. For each training tile four random crops were sampled from the training tiles. The crops were selected from the top left, top right, bottom left and bottom right of four randomly selected tiles, the crops were sampled so when merged together the resulting tile would still be 1024x1024 pixels. Augmentation was applied to the merged tile.

Augmentations and merging of random crops were not performed during validation. The training, validation and hold-out test split was 70%,15% and 15% respectively.

The dice score of our validation with the lowest loss (0.36) was 0.83 (Supplementary Fig. 1), the dice score of our held-out test set was 0.87.

Deep learning derived phenotypes

For prediction we adapted the code base for HAPPY (21) implemented in Python and using PyTorch for semantic segmentation. Tiles of 1024x1024 pixels are used, starting from the top left corner of the WSI and moving left to right in rows (Fig. 1). Tiles include a 256 pixel overlap with both vertically and horizontally adjacent neighbours. Tiles are inferred to be empty, and excluded from further processing, if all pixels are close to white or black, or if the ratio of the top decile mean pixel value to the bottom decile mean pixel value is above 0.95.

The weights for the model are from the training described in the “Training deep learning model” section.

The segmentation mask (using a 0.8 segment network confidence cutoff for saving predictions) of each tile is converted to polygons using the “measure.find_contours()” package from the scikit-image (a.k.a. skimage) python library. The polygons are constructed using the shapely python library.

Segmentation is performed twice on each WSI with respectively a pixel size of 0.2500 µm and 0.5034 µm.

After the whole WSI has been segmented each polygon is searched for polygons in its vicinity using an STRtree (22) from the shapely library, and then merged with intersecting polygons (Fig. 1).

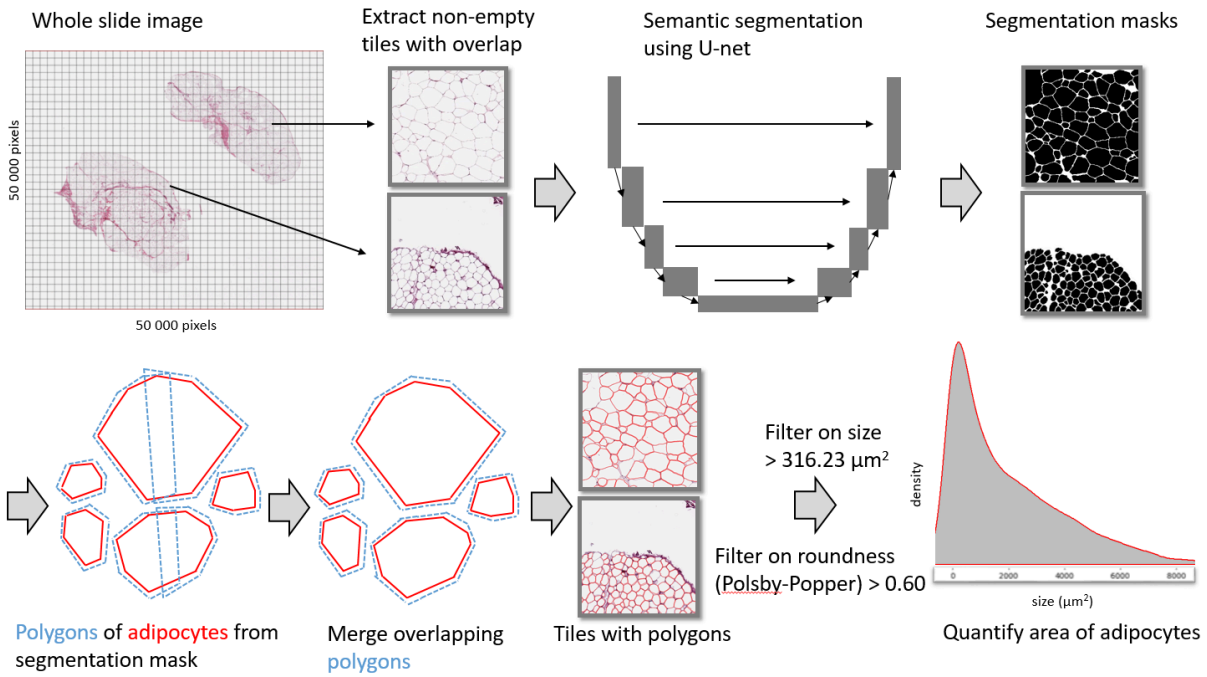


Figure 1. Schematic overview of the pipeline presented in this paper.

Post-Processing

The merged polygon list was filtered for area greater than 316.23 µm^2 and a Polsby-Popper (PP) roundness value (23) greater than 0.6 (manually selected thresholds; prefiltered distributions and selected thresholds are shown in Supplementary Fig. 2 & 3) (Fig. 1). To remove WSIs without any or very little adipose tissue, WSIs with fewer than 1500 polygons left after filtering and more than 65% of the polygons with an area below 750 µm^2 were discarded (manually selected thresholds, see Supplementary Fig. 4 & 5). Additionally, we removed faulty and low quality WSIs, upon quality control from laboratory technical staff based at Leipzig, Munich and Hohenheim.

From the filtered polygons from each individual we calculated mean adipocyte size and upper 95% quantile of adipocyte size, and the ratio between mean adipocyte size in visceral and subcutaneous visceral adipose tissue.

Validation of U-net based estimates of adipocyte size against Adiposoft

Using WSIs from the Leipzig cohort we ran Adiposoft on 10 tiles (1024x1024 pixels) from 17 WSIs. The WSIs were chosen randomly, the tiles were chosen so that there was no empty space, and only adipocytes in the tile. We used the same lower size cutoff of 316.23 µm^2 as for the U-net based approach. We chose to discard adipocytes touching the edge of the tile.

For comparison with U-net we extracted the polygons within the coordinates of the tiles used with Adiposoft, we also removed polygons touching the edge of the tile.

Cohorts
White adipose tissue samples from subcutaneous and visceral depots were obtained from five different study cohorts (Table 1). Each study participant gave written informed consent and the protocols were approved by the corresponding ethics committee. Munich and Hohenheim: 5716/13, 1946/07, 409/16s. Leipzig: 159-12-21052012, 017-12-23012012. ENDOX: REC: 09/H0604/58, IRAS: 8282, GTEX: NIH (project id phs000424.v7.p2).

Cohort	Individuals with WSI (SC / VC)	Phenotypes	BMI (mean, SD)	Age (mean, SD)	Sex (M / F)
Munich	211 / 232	Anthropometric and metabolic	45.13, 12.58	46.13, 12.71	105 / 204
Leipzig	1151 / 1145	Anthropometric and metabolic	46.62, 11.49	49.77, 13.01	403 / 766
Hohenheim	47 / 50	Anthropometric and metabolic	46.55, 6.23	42.31, 13.09	37 / 114
Endox (Oxford)	425 / 55	Anthropometric	26.60, 5.80	32.78, 7.59	0 / 352
GTEX	918 / 759	Anthropometric and metabolic	27.32, 4.13	52.76, 12.91	653 / 327

Table 1. An overview of the five cohorts used in this study and the number of WSIs in each cohort. BMI, age and sex information is the number of individuals with phenotypes, these numbers might therefore differ from the number of individuals with WSIs.

Munich, Leipzig and Hohenheim cohort
Male and female patients undergoing abdominal laparoscopic surgery were included. Subcutaneous adipose tissue biopsy derived from beneath the skin at the abdominal surgical incision site and visceral adipose tissue was obtained at the proximity of the angle of His. After sampling, approx. 5 mm^3 pieces of the tissue, biopsies were sectioned and fixed in 4% paraformaldehyde for histology.

ENDOX cohort
Female premenopausal participants were recruited at the Oxford Endometriosis CaRe Centre between 2012 and 2018 for endometriosis screening via laparoscopy. Biopsies of the endometrium, peritoneal fluid and adipose tissue from subcutaneous and visceral depots were obtained (14).

GTEX cohort
Whole slide images from HE stained subcutaneous and visceral adipose tissue samples were obtained from the Genotype Tissue and Expression (GTEx) project. Samples were obtained from individuals post-mortem. Subcutaneous adipose tissue biopsies were sampled at the leg, approx. 2 cm below the patella, and visceral adipose tissue was obtained from the omentum (24).

Histology and imaging
Dehydration and clearing of fixated histology samples was performed automatically (TP1020, Leica, Germany). Afterwards, samples were embedded in paraffin and 5 µm thick tissue sections were obtained using a rotary microtome (RM2255, Leica, Germany). Subsequently, the tissue sections were transferred to a glass microscope slide, hematoxylin and eosin (H&E) staining was applied using a fully automated multistainer (ST5020, Leica, Germany) and the samples were immediately coverslipped. Digital whole slide images were obtained using a slide scanner (Aperio AT2, Leica, Germany).

Genotyping
The Leipzig cohort was genotyped in two different batches. We checked, doing PCA with pruning and only with autosomal markers, and with a MAF>0.05 and a missingness<0.01 filter, for no separation between the two batches looking at the top 10 genetic PCs (Supplementary Fig. 6). The two genotype files were merged into one genotype file which was used for all subsequent analyses.

Cohort	Individuals genotyped	Sites	Genotype Platform
Leipzig	1,251	730,059	Global Screening Array-24 v3.0 BeadChip
ENDOX (part I)	56	655,448	Affymetrix Axiom
ENDOX (part II)	127	700,078	Illumina Infinium Global Screening Array
Hohenheim & Munich (part I)	175	730,059	Global Screening Array-24 v3.0 BeadChip
Munich (part II)	192	700,078	Infinium Global Screening Array
GTEx	866	69,763,935	Whole genome sequenced on: Illumina HiSeq 2000 machine and Illumina HiSeq X machine

Table 2. Overview of the genotyping of each cohort. GTEx donors were whole genome sequenced to a median depth of 32X. 79 GTEx donors were sequenced on an Illumina HiSeq 2000 machine, 801 donors were sequenced on an Illumina HiSeq X machine. For more information on how the GTEx cohort was sequenced see (25).

For an overview of the genetic data see Table 2, for how we applied quality control see the “Genetic Quality Control” section in Supplementary Methods.

Imputation
Each SNP-chip was phased and imputed independently, after applying a quality control filter (see "Genetic Quality Control" section in Supplementary Methods), and splitting by chromosome, and preparing the input files according to the TOPMed guidelines (https://topmedimpute.readthedocs.io/en/latest/). The TOPMed Imputation server was used (26). Eagle v2.4 (27) was used for the phasing, Minimac4 (26) was used for the imputation both with the multiethnic TOPMed r2 reference panel (approx. 97,000 individuals). The "vs. TOPMed panel" was selected as the population.

Association analysis
The GWAS was done on bimbam mean genotype files, also known as genetic dosages, including genetic variants with an imputation Rsq > 0.3. A linear mixed model was run, implemented in the software GEMMA (28), on each cohort. The model was adjusted for sex, age, age^2, sex · age, and the top ten genetic PCs. The ENDOX cohort was analysed jointly and the linear mixed model was additionally adjusted for which SNP-chip. The Hohenheim cohort and part of the Munich cohort genotyped on the same SNP-chip were also analysed together. The phenotypes were analysed raw, and inverse normal quantile transformed. The outputs from each cohort were meta-analysed together and the joint estimates of effect size and P-value were derived using the inverse variance based strategy from METAL (29).

Furthermore, for our lead variants we did a Phenome-wide association study (PheWAS). We utilised a meta-analysis based approach like for the GWAS. We focused on traits that were linked to hypertrophy in the literature: BMI, waist-hip ratio (WHR), plasma glycated haemoglobin, fasting plasma glucose, fasting plasma insulin, plasma 2 hour blood glucose, Interleukin 6, Adiponectin, type II diabetes status and C-reactive protein.

Colocalisation
To identify colocalisations between our adipocyte derived traits and other metabolic traits, colocalization was assessed using a Bayesian statistical test with the coloc R-package (version 5.2.2) (30). Genomic regions were defined by windows of 1 Mb around suggestive and significant GWAS signals ($P < 10^{-5}$). Overlapping regions were merged. The minimal data required by coloc were used as input: SNP IDs, positions, effect sizes, standard errors, data type (quantitative or case-control), minor allele frequencies and sample sizes. To identify potential causal variants, 95% credible sets of variants based on their posterior inclusion probabilities were created. A posterior probability PPH4 (H4: associations with both traits due to a single causal variant) > 80% was considered sufficient for identification of colocalisation.

We analysed all the traits listed in Fig. 2 & 3.

Adipocyte hypertrophy in the Leipzig cohort
In a linear model we explore which phenotypes associate with mean adipocyte size. This was done to maximise statistical power instead of using a variable with hypertrophy or hyperplasia. Adipocyte size was divided by 1000, to make the effect sizes more meaningful. The linear model was adjusted for age, sex, BMI and type 2 diabetes state:

$$y_i = \alpha_i + \beta_{i1} \cdot meanAdipocyte_i + \beta_{i2} \cdot age_i + \beta_{i3} \cdot sex_i + \beta_{i4} \cdot BMI_i + \beta_{i5} \cdot ifDiabetes_i + e_i$$

For more on estimating the effect of subcutaneous adipose tissue sampling location and adipocyte size epidemiology using linear mixed models see Supplementary Methods.

Results

We have benchmarked our U-net based estimated mean size of adipocytes against Adiposoft (15), the linear correlation is $R^2 = 0.76$ (Supplementary Fig. 7).

Our U-net based segmentation approach detects adipocytes larger than in (14) (Fig. 4). With the same filtering applied to both methods, we still see the same pattern (Supplementary Fig. 8-9). Furthermore, in Supplementary Fig. 10-11 we show the linear correlation between our estimates and the estimates from (14) ($R^2_{subcutaenous} = 0.99$, $R^2_{visceral} = 0.99$), the current method consistently gives larger estimates than the previous implementation in (14) ($\beta_{subcutaenous} = 1.44$, $\beta_{visceral} = 1.4$).

We replicate the finding from (14) of larger mean adipocyte size in subcutaneous than visceral adipose tissue (except for the Hohenheim cohort) (Supplementary Fig. 12). Furthermore we show a strong correlation between mean adipocyte size and BMI with R2 values range from (excluding Hohenheim) 0.24 to 0.67 in subcutaneous, and 0.32 to 0.52 in visceral adipose tissue (Supplementary Fig. 13).

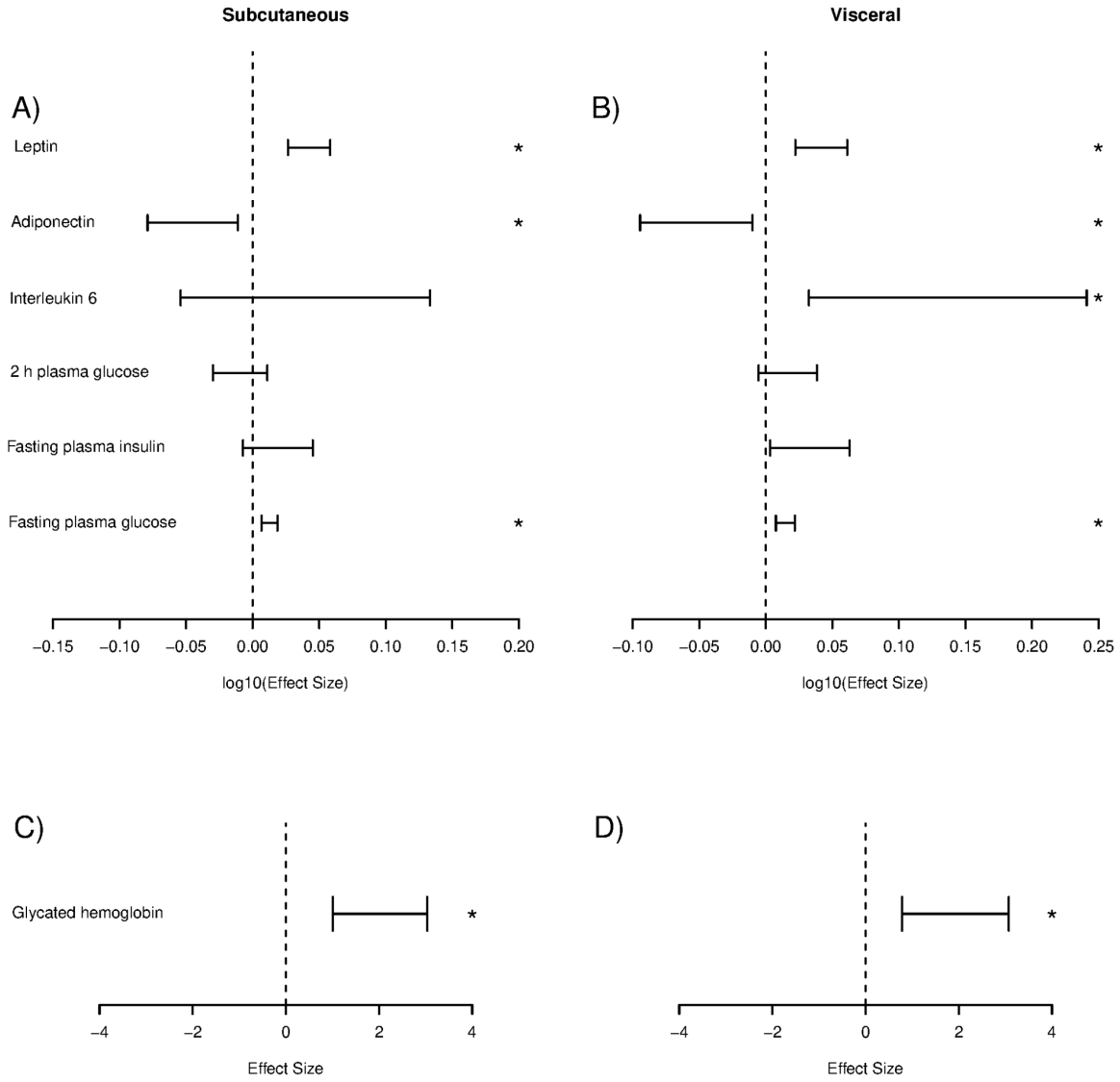


Figure 2. Effect size of mean size of adipocytes on the listed phenotypes, in subcutaneous adipose tissue A) and C) and visceral adipose tissue B) and D) with the 95% confidence intervals. * denotes significant after false discovery rate (FDR) correction.

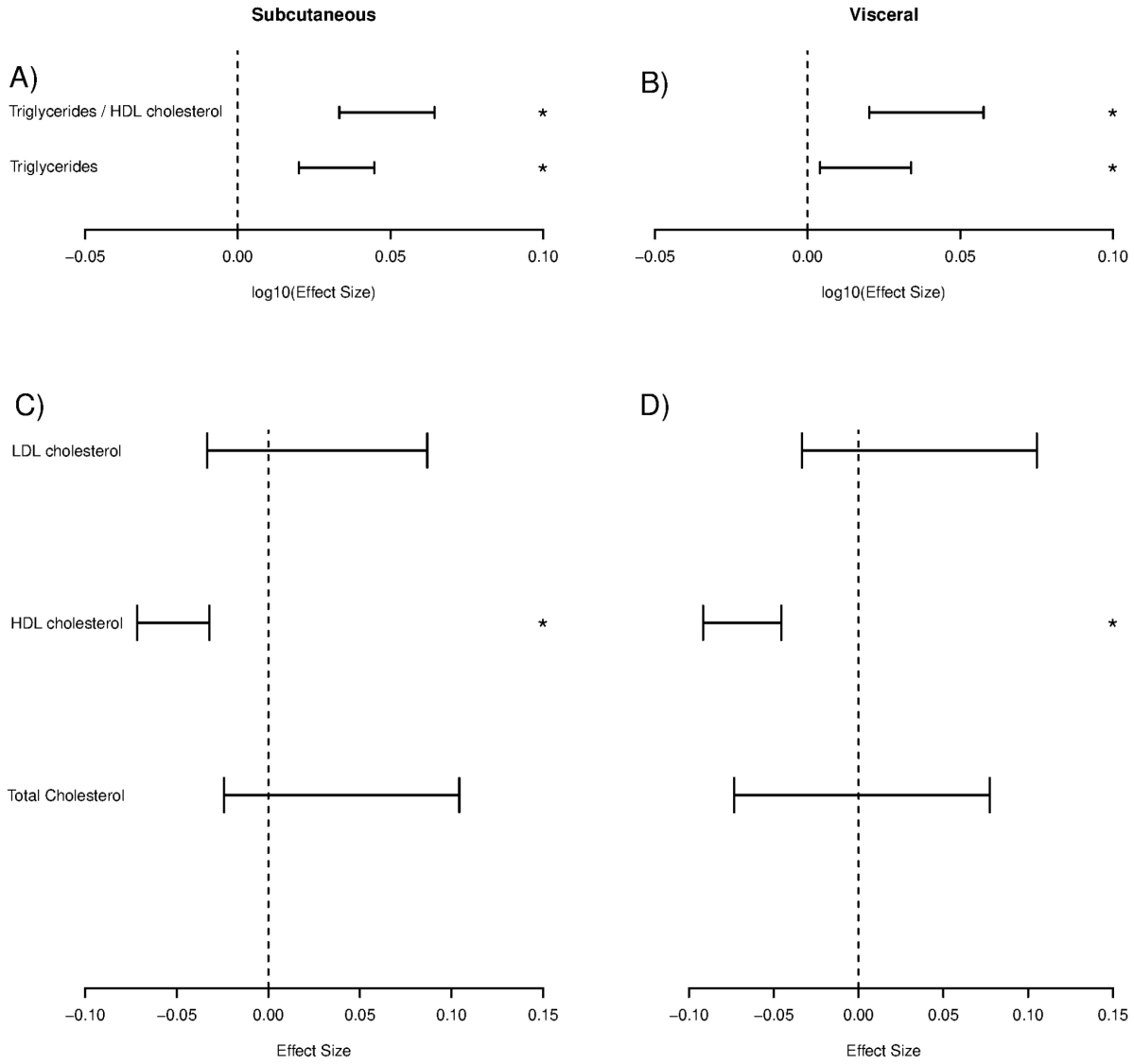


Figure 3. Effect size of mean size of adipocytes on the listed phenotypes, in subcutaneous adipose tissue A) and C) and visceral adipose tissue B) and D) with the 95% confidence intervals. * denotes significant after FDR correction.

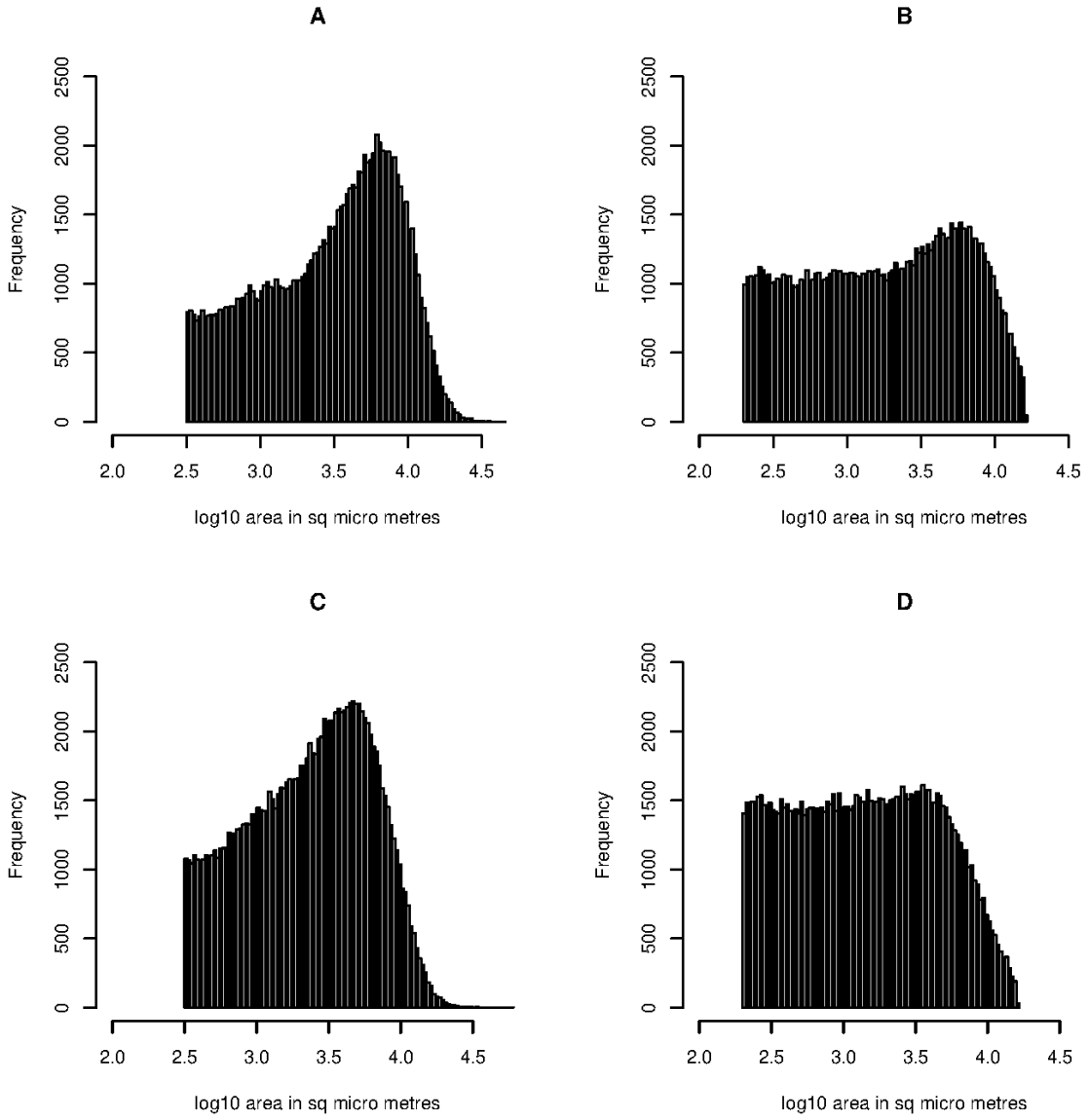


Figure 4. 500 randomly sampled segmented adipocytes from 206 WSIs from subcutaneous adipose tissue and 248 WSIs from visceral adipose tissue both from the Munich cohort A) using our filters B) using the filters from C. Glastonbury et al. 2020 - both subcutaneous adipose tissue. C) and D) same as A and B but with visceral adipose tissue. The histograms are shown with the log10 of size of adipocytes.

Adipocyte hypertrophy in the Leipzig cohort

Mean adipocyte size correlated positively with levels of plasma glycated haemoglobin (mmol/mol) ($P_{SC} = 1.06 \cdot 10^{-4}$, $\beta_{SC} = 2.02$ and $P_{VC} = 1.04 \cdot 10^{-3}$, $\beta_{VC} = 1.93$), and with levels of fasting plasma glucose in subcutaneous adipose tissue ($P_{SC\_LOG10} = 3.94 \cdot 10^{-5}$, $\beta_{SC\_LOG10} = 0.013$) and visceral adipose tissue ($P_{VC\_LOG10} = 5.06 \cdot 10^{-5}$, $\beta_{VC\_LOG10} = 0.015$). We observed a negative correlation with adiponectin ($P_{SC\_LOG10} = 0.011$, $\beta_{SC\_LOG10} = -0.045$, $P_{VC\_LOG10} = 0.017$, $\beta_{VC\_LOG10} = -0.052$) but a positive correlation with levels of leptin ($P_{SC\_LOG10} = 2.03 \cdot 10^{-7}$, $\beta_{SC\_LOG10} = 0.042$, $P_{VC\_LOG10} = 3.19 \cdot 10^{-5}$, $\beta_{VC\_LOG10} = 0.042$). Lastly, the data revealed a positive association with interleukin 6 levels in visceral adipose tissue ($P_{VC\_LOG10} = 0.012$, $\beta_{VC\_LOG10} = 0.14$) (Fig. 2). These results are significant after controlling for FDR (Benjamini–Hochberg) adjusting for 14 tests, giving a P-value threshold of 0.036.

There was no significant correlation with levels of plasma 2 hour blood glucose ($P_{SC\_LOG10} = 0.37$, $\beta_{SC\_LOG10} = -0.0093$ and $P_{VC\_LOG10} = 0.15$, $\beta_{VC\_LOG10} = 0.017$), fasting plasma insulin ($P_{SC\_LOG10} = 0.16$, $\beta_{SC\_LOG10} = 0.019$ and $P_{VC\_LOG10} = 0.030$, $\beta_{VC\_LOG10} = 0.033$), interleukin 6 ($P_{SC\_LOG10} = 0.41$, $\beta_{SC\_LOG10} = 0.040$ and $P_{VC\_LOG10} = 0.011$, $\beta_{VC\_LOG10} = 0.14$) (Fig. 2).

Mean adipocyte size correlated positively with levels of triglycerides (mmol/mol) ($P_{SC\_LOG10} = 3.53 \cdot 10^{-7}$, $\beta_{SC\_LOG10} = 0.032$ and $P_{VC\_LOG10} = 0.013$, $\beta_{VC\_LOG10} = 0.019$), and the triglycerides to HDL cholesterol ratio ($P_{SC\_LOG10} = 1.32 \cdot 10^{-9}$, $\beta_{SC\_LOG10} = 0.049$ and $P_{VC\_LOG10} = 5.15 \cdot 10^{-5}$, $\beta_{VC\_LOG10} = 0.039$). We further observed a negative correlation with HDL cholesterol ($P_{SC} = 2.97 \cdot 10^{-7}$, $\beta_{SC} = -0.052$ and $P_{VC} = 8.28 \cdot 10^{-9}$, $\beta_{VC} = -0.069$) (Fig. 3). These results are significant after controlling for FDR (Benjamini–Hochberg) adjusting for 10 tests, giving a P-value threshold of 0.03.

There was no significant difference in levels of LDL cholesterol ($P_{SC} = 0.38$, $\beta_{SC} = 0.027$ and $P_{VC} = 0.31$, $\beta_{VC} = 0.036$) nor with total cholesterol ($P_{SC} = 0.22$, $\beta_{SC} = 0.040$ and $P_{VC} = 0.96$, $\beta_{VC} = 0.0020$) (Fig. 3).

We performed the same analyses stratified by sex, and did not observe any statistically significant differences in effect size between males and females (Supplementary Fig. 14-15).

We also performed these analyses using the upper 95% quantile of adipocyte size instead of mean adipocyte size (Supplementary Fig. 16-17), and using the ratio between mean adipocyte size in visceral and subcutaneous adipose tissue (Supplementary Fig. 18-19).

Adipocyte hypertrophy all cohorts

We found that mean adipocyte size, in visceral depots, increases ($\mu = 0.12$, $P_{\mu} = 0.00014$, $\tau^2 = 0.0020$, $P_{\tau} = 0.09$, $I^2 = 51\%$) with age, with significant heterogeneity between the cohorts. Mean adipocyte size did not significantly increase with age in subcutaneous depots ($\mu = 0.07$, $P = 0.15$, $\tau^2 = 0.0068$, $P_{\tau} = 0.0005$, $I^2 = 79.9\%$) (Supplementary Fig. 20-21). Mean adipocyte size, in visceral depots, increased ($\mu = 0.39$, $P_{\mu} = 1.89 \cdot 10^{-7}$, $\tau^2 = 0.021$, $P_{\tau} = < 0.0001$, $I^2 = 88.5\%$) with BMI, with significant heterogeneity between the cohorts. Mean adipocyte size also increases in subcutaneous depots with BMI ($\mu = 0.40$, $P = 8.28 \cdot 10^{-21}$, $\tau^2 = 0.0057$, $P_{\tau} = 0.0053$, $I^2 = 72.8\%$), with significant heterogeneity between the cohorts (Supplementary Fig. 22-23). Mean adipocyte size, in visceral depots, was smaller in females than in males ($\mu = -0.32$, $P_{\mu} = 6.77 \cdot 10^{-15}$, $\tau^2 = 0$, $P_{\tau} = 0.49$), but this was not detected in subcutaneous depots ($\mu = 0.12$, $P_{\mu} = 0.41$, $\tau^2 = 0.061$, $P_{\tau} = < 0.0001$, $I^2 = 91.3\%$), with very

high significant heterogeneity between the cohorts (Supplementary Fig. 24-25). Mean adipocyte size was larger in type 2 diabetes cases compared to controls in visceral depots ($\mu = 0.11$, $P_\mu = 0.021$, $\tau^2 = 0$, $P_\tau = 0.69$), with no significant heterogeneity between cohorts. (Supplementary Fig. 26-27). Lastly, mean adipocyte size in visceral depots is significantly smaller ($\mu = -0.51$, $P = 1.57 \cdot 10^{-7}$, $\tau^2 = 0.037$, $P_\tau = < 0.0001$, $I^2 = 86.2\%$) than in subcutaneous depots, with very high significant heterogeneity between the cohorts (Supplementary Fig. 28).

These results replicate those in (14), except the finding of different adipocyte sizes in type 2 diabetes cases compared to controls.

Effect of subcutaneous adipose tissue sampling location

Mean adipocyte size in subcutaneous tissue is larger in GTEX (biopsy sampled from the knee area) compared to all the other cohorts ($P_{ALL} = 3.71 \cdot 10^{-35}$, $\beta_{ALL} = 771.77$). Mean adipocyte size in subcutaneous tissue is also larger in GTEX compared to each specific cohort ($P_{Leipzig} = 2.30 \cdot 10^{-23}$, $\beta_{Leipzig} = 773.61$), ($P_{Munich} = 1.19 \cdot 10^{-9}$, $\beta_{Munich} = 728.21$), ($P_{Hohenheim} = 6.97 \cdot 10^{-24}$, $\beta_{Hohenheim} = 2300.19$) and ($P_{Endox} = 2.93 \cdot 10^{-24}$, $\beta_{Endox} = 1004.69$).

GWAS

We observed no inflation across our GWAS across the five adipocyte size traits, in the whole cohort, and when stratifying by sex. We detected four independent signals, reaching significance ($P < 5.0 \cdot 10^{-8}$) (Supplementary Fig. 29-34).

Two loci associating with 95%-quantile of adipocyte size

We detect an association for an intron variant rs73184721 (chr3:175280793), located in the *NAALADL2* gene associated with an increased 95%-quantile of adipocyte size in visceral adipose tissue in the meta-analysis of all five cohorts ($\beta = 930.18$, $\beta_{SD} = 0.42$, $P = 8.82 \cdot 10^{-9}$) (Fig. 5). There was still an association when adjusted for BMI, though attenuated ($\beta_{SD} = 0.27$, $P = 1.19 \cdot 10^{-4}$). This variant did not show sex heterogeneity ($P = 0.81$). The variant was also associated with increased mean adipocyte size in visceral adipose tissue ($\beta_{SD} = 0.31$, $P = 4.38 \cdot 10^{-5}$).

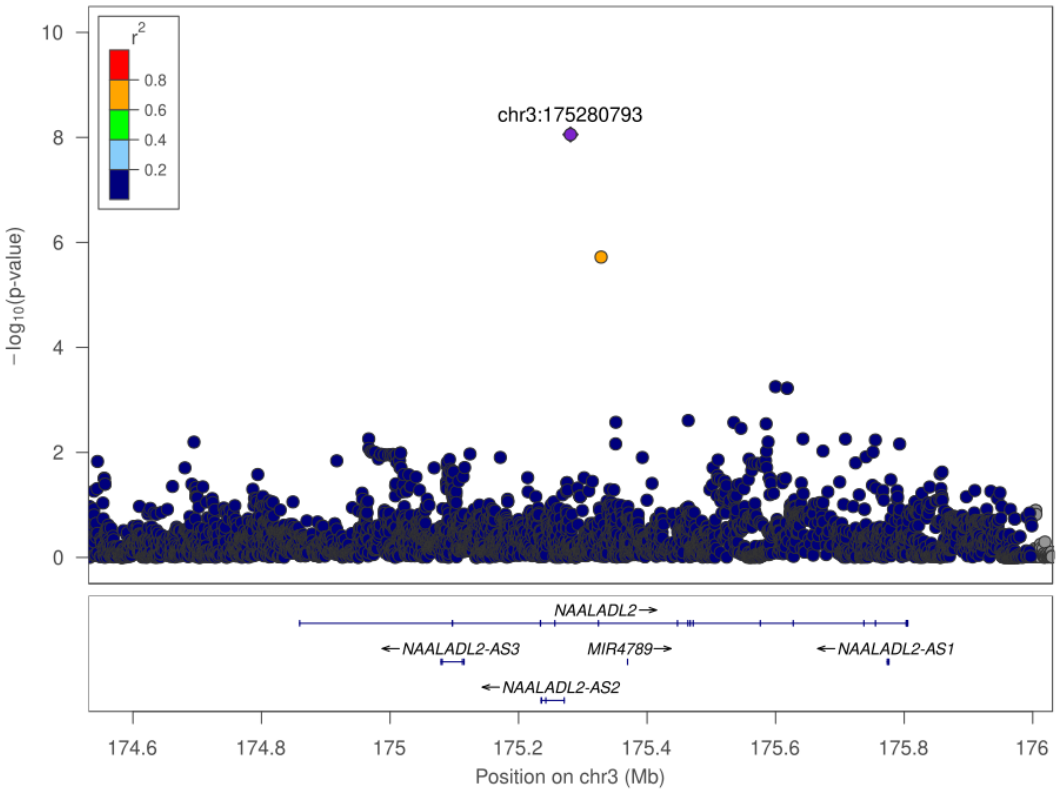


Figure 5. LocusZoom plot of rs73184721 and its association with 95%-quantile of adipocyte size in visceral adipose tissue. A region of +/- 750 kb has been plotted around the variant. The colour of the dots indicates the degree of LD between rs73184721 and that variant, the R^2 values are based on the GTEx sequence data.

We also detect an association for the small indel intron variant rs200047724 (chr14:78949603), located in the *NRXN3* gene on chromosome 14. It is associated (β = 1586.21, β_{SD} = 0.60, P = $7.60 \cdot 10^{-8}$) with an increased 95%-quantile of adipocyte size in subcutaneous adipose tissue in a meta-analysis of all five cohorts (Fig. 6).

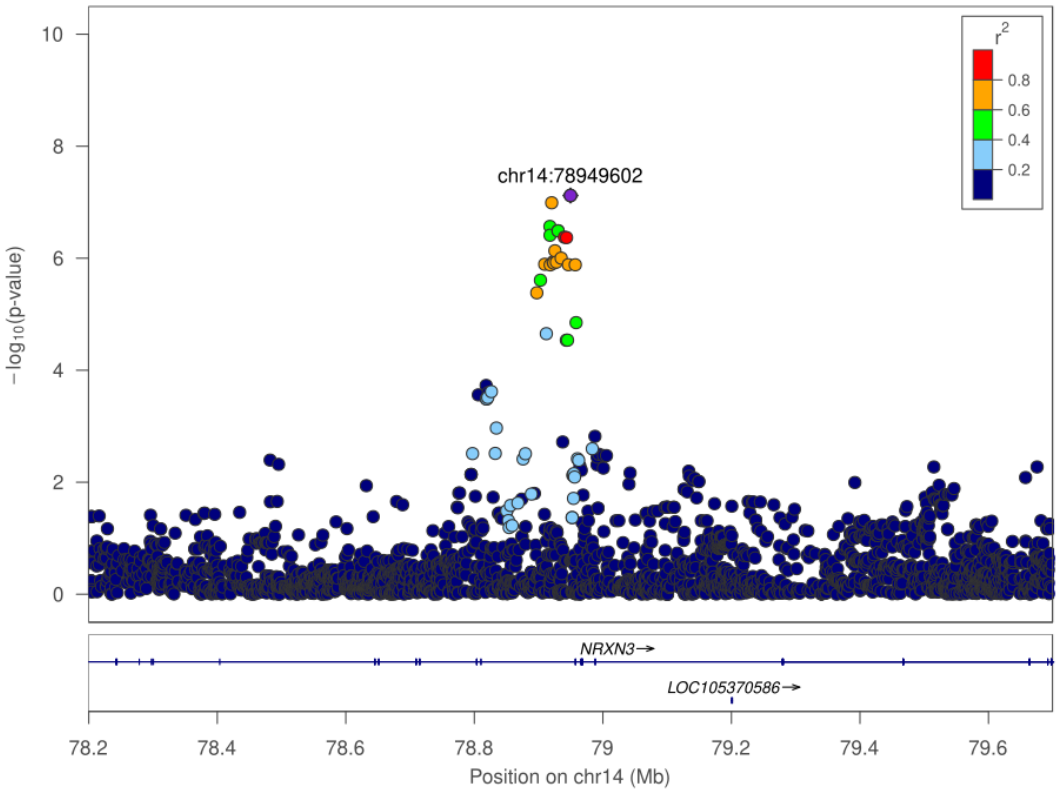


Figure 6. LocusZoom plot of rs200047724 and its association with 95%-quantile of adipocyte size in subcutaneous adipose tissue. A region of +/- 750 kb has been plotted around the variant. The colour of the dots indicates the degree of LD between rs200047724 and that variant, the R^2 values are based on the GTEx sequence data.

The variant rs72811236 mentioned in (14) only associates nominally with 95%-quantile of adipocyte size in subcutaneous tissue after adjusting for BMI.

Two loci associating with adipocyte size specific to females

We detected an association for an intergenic variant rs140503338 (chr14:68826371) with increased mean adipocyte size in females in subcutaneous adipose tissue in a meta-analysis of all five cohorts ($\beta = 956.06$, $\beta_{SD} = 0.85$, $P = 2.91 \cdot 10^{-8}$), limited to female donors (Supplementary Fig. 35). This variant was also associated with an increased 95%-quantile of adipocyte size in females in subcutaneous adipose tissue ($\beta_{SD} = 0.78$, $P = 3.21 \cdot 10^{-7}$). This association was still present when adjusted for BMI ($\beta_{SD} = 0.66$, $P = 4.78 \cdot 10^{-6}$). The variant did show sex heterogeneity ($P = 1.04 \cdot 10^{-4}$). We observed a nominal association in the meta-analysis for mean adipocyte size ($\beta_{SD} = 0.32$, $P = 0.042$) and 95%-quantile of adipocyte size ($\beta_{SD} = 0.35$, $P = 0.026$) in women in visceral adipose tissue. Furthermore this variant is associated plasma glycated haemoglobin levels ($\beta_{SD} = 0.65$, $P = 0.0081$) in our PheWAS (Supplementary Table 1) and nominally with BMI ($\beta_{SD} = 0.42$, $P = 0.016$) and 2 h plasma glucose ($\beta_{SD} = 1.25$, $P = 0.046$).

We also identified an association of an intron variant rs11656704 (chr17:19789496) in the gene *ULK2* with a decreased 95%-quantile of adipocyte size in women in visceral adipose tissue ($\beta = -700.64$, $\beta_{SD} = -0.31$, $P = 4.50 \cdot 10^{-8}$). This signal remained significant after adjusting for BMI ($\beta_{SD} = -0.26$, $P = 7.36 \cdot 10^{-7}$). This variant also showed sex heterogeneity (P

= 0.00025). An association in the meta-analysis is seen for mean adipocyte size (β_{SD} = -0.29, P = $4.62 \cdot 10^{-7}$) in visceral adipose tissue. Furthermore this variant is associated with WHR (β_{SD} = -0.54, P = 0.0056) and plasma glycated haemoglobin levels (β_{SD} = -0.26, P = 0.0055) in our PheWAS (Supplementary Table 1).

Another variant rs188936910 that associated with mean size of adipocytes in subcutaneous adipose tissue in females only at the suggestive level in (14), associated nominally in our study with mean size of adipocytes in subcutaneous adipose tissue in females (P = 0.013).

We did not identify any genome-wide significant associations in a sex-dimorphic test (31) genome-wide. Lastly, we identified a nominal significant signal on chromosome 22 for mean size of adipocytes in visceral adipose tissue (top variant: rs7292070, β_{SD} = 0.25, P = $2.21 \cdot 10^{-7}$) which cololocalised with a signal for HDL cholesterol (P_{H4} = 0.90) (32). The top variant of this region for HDL cholesterol is rs2070512 (β_{SD} = -0.036, P = $1.39 \cdot 10^{-100}$).

Discussion

Our novel approach of WSI segmentation enables us to detect and quantify very large adipocytes, which Glastonbury et al., 2020 failed to do as shown in Fig. 4. Our approach enables us to detect the metabolic impact of very large adipocytes, and to investigate their genetic architecture.

We show that larger mean adipocyte size is associated with a metabolically adverse profile, with an increased risk of type 2 diabetes and increased plasma fasting glucose consistent with literature (8), the correlation for fasting plasma glucose is stronger in visceral adipose tissue ($\beta_{SC\_LOG10} = 0.013$ and $\beta_{VC\_LOG10} = 0.015$). Furthermore, increased mean adipocyte size is associated with increased levels of glycated haemoglobin, consistent with (33) where average adipose cell area in visceral adipose tissue was associated with higher odds of having persistently elevated glycated haemoglobin at 12 months post-Roux-en-Y gastric bypass, and inconsistent with (34) where there was a significant negative correlation of glycated haemoglobin with mean size of adipocytes from the subcutaneous adipose tissue.

The decrease in adiponectin levels with increased mean adipocyte size is consistent with previous findings (8). There is a significant association for interleukin 6 and mean adipocyte size in visceral but not in subcutaneous adipose tissue, this is inconsistent with (35), where levels of interleukin 6 were higher in diabetic and non-diabetic obese patients compared to control subjects in subcutaneous adipose tissue (Fig. 2). The increased mean size of adipocytes in men, and with increased age in visceral adipose tissue is consistent with the meta analysis in (8). However, the finding of no increase in mean size of adipocytes with age in subcutaneous adipose tissue is not. Additionally the decreased mean adipocyte size in visceral depots is also consistent with literature (8) (Supplementary Fig. 20-28).

We do not detect a statistically significant increase in fasting plasma insulin levels with increased mean size of adipocytes, which is not consistent with literature (8) (Fig. 2). Nor do we detect any differences in the impact of adipocyte hypertrophy on metabolism between males and females (Supplementary Fig. 14-15).

These results are yet another indication that obesity is much more complex than just BMI, but that adipose tissue and adipocytes play a key role in metabolic dysfunction and subsequent disease. Adipocyte hypertrophy strongly associates with an adverse metabolic profile independent of BMI. But this relationship is complex and subject to a lot of noise, we hypothesise that is why we are not able to replicate all the findings regarding adipocyte size and metabolic health from literature.

We could only find two previous studies (12,14) investigating the genetics of adipocyte morphology; both of them fail to report genome-wide significant associations ($P < 5.0 \cdot 10^{-8}$).

rs73184721 is an intron variant in the large *NAALADL2* gene, whose function is currently unknown. However, another gene close to the associated variant is *NLGN1*, which acts as a splice site-specific ligand for β-neurexins. Intriguingly, this variant has been nominally associated with increased pancreas fat percentage and increased subcutaneous adipose tissue volume (36). *NLGN1* is associated with predicted visceral adipose tissue mass (37), BMI (38), and type 2 diabetes (39). We therefore speculate that the effect on increased 95%-quantile of adipocyte size in visceral tissue might be mediated through modification of

the expression of the *NLGN1* gene. Additionally, this variant associates with decreased WHR, but not BMI in the Leipzig cohort (Supplementary Table 1). However, when all the cohorts are meta analysed the association is attenuated after adjustment for BMI. We therefore conclude that further investigation is needed to establish the molecular mechanism of this association.

NRXN3 is a neurexin gene, it encodes proteins that function in the nervous system as receptors and cell adhesion molecules. It is mainly expressed in the brain. Intriguingly, the *NRXN3* gene has previously been associated with BMI (40), and waist circumference (41). Even more intriguingly altered methylation-expression in extremely obese individuals has been associated with *NRXN3 (42).*

The association of rs200047724 in the *NRXN3* gene with an increased 95% quantile of size of adipocytes that is independent of BMI hints at a new biological mechanism that impacts the distribution of adipocytes, independent of BMI.

The female specific associated variant rs140503338, is located close to a cluster of genes. *RAD51B* is a gene upstream of this variant that is a member of the *RAD51* protein family. *RAD51* family members are evolutionarily conserved proteins essential for DNA repair by homologous recombination, it is previously associated with BMI-adjusted waist-hip ratio (43) and waist-hip ratio (44). Interestingly, it is also most highly expressed in the uterus, hinting at a sex specific underlying mechanism (25). *GALNT16* is a gene that enables polypeptide N-acetylgalactosaminyltransferase activity, it is previously associated with BMI (40). We see a nominal association with BMI in the Leipzig cohort (Supplementary Table 1), it has previously been nominally associated with increased waist hip ratio adjusted for BMI in both males and females (45). We therefore speculate that the effect might be mediated through the *RAD51B* gene, however further investigation is needed, to establish this.

The other female specific variant is rs11656704 an intron variant in the gene *ULK2*, it is located close to the gene *AKAP10* previously associated with BMI (46). However, this variant does not associate with BMI in the Leipzig cohort but it does associate with decreased WHR and decreased plasma glycated haemoglobin levels (Supplementary Table 1). We can speculate that the *AKAP10* gene is involved, however further investigation is needed to establish this association.

rs73184721 displays an unusual LD pattern where our lead variant only tags one other variant. Worried that this might be due to an imputation or genotyping error, we confirmed the same unusual LD pattern for the 1000 genomes data (47) and the TOPMed panel (48). For rs140503338 we did the same and checked the LD pattern against the 1000 genomes data (47) and the TOPMed panel (48) and found the pattern to be consistent with those.

Interestingly for the colocalisation analysis the same direction of effects (increased mean size of adipocytes in visceral adipose tissue as HDL cholesterol levels decrease), as in the epidemiological analyses of hypertrophy (Fig. 3). Their top variant (rs2070512) is an eQTL for the gene *CCDC116* in subcutaneous and visceral adipose tissue and *UBE2L3* and *YDJC* in subcutaneous adipose tissue only (25) (GTEX version 8). However, none of these genes have obvious links to adipose tissue, so further investigation is needed. But this does indicate a relationship between adipocyte size and HDL cholesterol.

For the first time to the best of our knowledge, we have reported genetic variants that attain genome-wide significance in their association with adipocyte size. However, due to the lack of other genetic studies on adipocyte size, and heterogeneity of the phenotype used in (12), replication of the signals is not possible. We do believe that especially the signals of rs200047724, rs73184721 and rs140503338 are interesting and merit further investigation, either with functional studies, or GWAS with increased sample size.

Availability and implementation

The method is implemented in python making use of the PyTorch library. The code for training the model can be found at [https://github.com/Nellaker-group/PyTorchUnet], the code for running prediction with the model can be found here [https://github.com/Nellaker-group/UNetHAPPY].

Acknowledgements

We would like to thank Rona Strawbridge and Ingrid Dahlman for contributing with summary statistics from the GENiAL cohort (12). Furthermore, we would like to thank the Core Facility for Comparative Experimental Pathology at Klinikum Rechts der Isar, Technical University of Munich for their technical support obtaining the whole slide images.

Author contributions

E.J., C.N., C.M.L. and M.C. conceived the study and designed the experiments. E.J, P.D.C. and C.N. wrote the code. E.J. and J.H. created ground truth annotations. E.J. performed the experimental analysis. J.H. and P.K. prepared and processed the adipose tissue histology samples. H.H., M.B., and M.C. collected the adipose tissue biopsies. E.J. conducted the genetic association studies of adipocyte morphology. E.J. performed the quality control and imputation of the genetic data. E.J. and P.K. conducted the epidemiological studies of adipocyte morphology. E.J., P.K, P.D.C., C.N., and M.C. prepared the manuscript with input, revisions, and approval from H.H., M.B., T.H., C.M.L.. The supervision of the research was done by P.D.C. C.M.L., C.N., and M.C..

Funding statements

E.J and C.N. funded by an Alliance Grant from The Novo Nordisk Foundation Center for Basic Metabolic Research an independent research centre at the University of Copenhagen, partially funded by an unrestricted donation from the Novo Nordisk Foundation (NNF18CC0034900 and NNF23SA0084103). C.M.L. is supported by the Li Ka Shing Foundation, NIHR Oxford Biomedical Research Centre, Oxford, NIH (1P50HD104224-01), Gates Foundation (INV-024200), and a Wellcome Trust Investigator Award (221782/Z/20/Z).

1. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol. 2018 Feb;14(2):88–98.

2. Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. Aliment Pharmacol Ther. 2011 Aug;34(3):274–85.

3. Lotta LA, Gulati P, Day FR, Payne F, Ongen H, Van De Bunt M, et al. Integrative

genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. 2017;49(1):17.

4. Longo M, Zatterale F, Naderi J, Parrillo L, Formisano P, Raciti GA, et al. Adipose Tissue Dysfunction as Determinant of Obesity-Associated Metabolic Complications. Int J Mol Sci [Internet]. 2019 May 13;20(9). Available from: http://dx.doi.org/10.3390/ijms20092358

5. Gao H, Mejhert N, Fretz JA, Arner E, Lorente-Cebrián S, Ehrlund A, et al. Early B cell factor 1 regulates adipocyte morphology and lipolysis in white adipose tissue. Cell Metab. 2014 Jun 3;19(6):981–92.

6. Spalding KL, Arner E, Westermark PO, Bernard S, Buchholz BA, Bergmann O, et al. Dynamics of fat cell turnover in humans. Nature. 2008 Jun 5;453(7196):783–7.

7. Liu F, He J, Wang H, Zhu D, Bi Y. Adipose morphology: A critical factor in regulation of human metabolic diseases and adipose tissue dysfunction. Obes Surg. 2020 Dec 6;30(12):5086–100.

8. Ye RZ, Richard G, Gévry N, Tchernof A, Carpentier AC. Fat cell size: Measurement methods, pathophysiological origins, and relationships with metabolic dysregulations. Endocr Rev. 2022 Jan 12;43(1):35–60.

9. Bouchard C, Despres JP, Mauriege P. Genetic and nongenetic determinants of regional fat distribution. Endocr Rev. 1993 Feb;14(1):72–93.

10. Herrera BM, Lindgren CM. The genetics of obesity. Curr Diab Rep. 2010 Dec;10(6):498–505.

11. Kulyté A, Aman A, Strawbridge RJ, Arner P, Dahlman IA. Genome-wide association study identifies genetic loci associated with fat cell number and overlap with genetic risk loci for type 2 diabetes. Diabetes. 2022 Jun 1;71(6):1350–62.

12. Lundbäck V, Kulyté A, Arner P, Strawbridge RJ, Dahlman I. Genome-wide association study of diabetogenic adipose morphology in the GENetics of Adipocyte Lipolysis (GENiAL) cohort. Cells. 2020 Apr 27;9(5):1085.

13. Glastonbury CA, Couto Alves A, El-Sayed Moustafa JS, Small KS. Cell-type heterogeneity in adipose tissue is associated with complex traits and reveals disease-relevant cell-specific eQTLs. Am J Hum Genet. 2019 Jun 6;104(6):1013–24.

14. Glastonbury CA, Pulit SL, Honecker J, Censin JC, Laber S, Yaghootkar H, et al. Machine Learning based histology phenotyping to investigate the epidemiologic and genetic basis of adipocyte morphology and cardiometabolic traits. PLoS Comput Biol. 2020 Aug;16(8):e1008044.

15. Galarraga M, Campión J, Muñoz-Barrutia A, Boqué N, Moreno H, Martínez JA, et al. Adiposoft: automated software for the analysis of white adipose tissue cellularity in histological sections. J Lipid Res. 2012 Dec;53(12):2791–6.

16. McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, et al. CellProfiler 3.0: Next-generation image processing for biology. PLoS Biol. 2018 Jul;16(7):e2005970.

17. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Springer; 2015. 234-241 p. (International Conference on Medical image

computing and computer-assisted intervention).

18. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library [Internet]. arXiv [cs.LG]. 2019. Available from: http://arxiv.org/abs/1912.01703

19. Kingma DP, Ba J. Adam: A Method for Stochastic Optimization [Internet]. arXiv [cs.LG]. 2014. Available from: http://arxiv.org/abs/1412.6980

20. Buslaev A, Parinov A, Khvedchenya E, Iglovikov VI, Kalinin AA. Albumentations: fast and flexible image augmentations [Internet]. arXiv [cs.CV]. 2018. Available from: http://arxiv.org/abs/1809.06839

21. Vanea C, Džigurski J, Rukins V, Dodi O, Siigur S, Salumäe L, et al. HAPPY: A deep learning pipeline for mapping cell-to-tissue graphs across placenta histology whole slide images [Internet]. bioRxiv. 2023 [cited 2024 Mar 13]. p. 2022.11.21.517353. Available from: https://www.biorxiv.org/content/biorxiv/early/2023/02/27/2022.11.21.517353

22. Leutenegger ST, Institute for Computer Applications in Science and Engineering. STR: A Simple and Efficient Algorithm for R-Tree Packing. Institute for Computer Applications in Science and Engineering, NASA Langley Research Center; 1997. 34 p.

23. Polsby D, Popper R. The Third Criterion: Compactness as a Procedural Safeguard Against Partisan Gerrymandering. 2015 Oct 16 [cited 2024 Mar 18]; Available from: https://digitalcommons.law.yale.edu/ylpr/vol9/iss2/6

24. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The genotype-tissue expression (GTEx) project. 2013;45(6):580.

25. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020 Sep 11;369(6509):1318–30.

26. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet. 2016 Oct;48(10):1284–7.

27. Loh PR, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016 Nov;48(11):1443–8.

28. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012 Jun 17;44(7):821–4.

29. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics. 2010 Sep 1;26(17):2190–1.

30. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 2014 May 15;10(5):e1004383.

31. Winkler TW, Kutalik Z, Gorski M, Lottaz C, Kronenberg F, Heid IM. EasyStrata: evaluation and visualization of stratified genome-wide association meta-analysis data. Bioinformatics. 2015 Jan 15;31(2):259–61.

32. Graham SE, Clarke SL, Wu KHH, Kanoni S, Zajac GJM, Ramdas S, et al. The power of genetic diversity in genome-wide association studies of lipids. Nature. 2021 Dec 9;600(7890):675–9.

33. Keshavjee SH, Schwenger KJP, Yadav J, Pickel L, Ghorbani Y, Sung HK, et al. Adipose tissue and plasma markers associated with HbA1c pre- and post-bariatric surgery: A cross-sectional and cohort study. Obes Surg. 2023 Aug 28;33(8):2443–51.

34. Fang L, Guo F, Zhou L, Stahl R, Grams J. The cell size and distribution of adipocytes from subcutaneous and visceral fat is associated with type 2 diabetes mellitus in humans. Adipocyte. 2015 Oct;4(4):273–9.

35. Bahceci M, Gokalp D, Bahceci S, Tuzcu A, Atmaca S, Arikan S. The correlation between adiposity and adiponectin, tumor necrosis factor alpha, interleukin-6 and high sensitivity C-reactive protein levels. Is adipocyte size associated with inflammation in adults? J Endocrinol Invest. 2007 Mar;30(3):210–4.

36. Liu Y, Basty N, Whitcher B, Bell JD, Sorokin EP, van Bruggen N, et al. Genetic architecture of 11 organ traits derived from abdominal MRI using deep learning. Elife [Internet]. 2021 Jun 15 [cited 2024 Sep 5];10. Available from: https://elifesciences.org/articles/65554

37. Karlsson T, Rask-Andersen M, Pan G, Höglund J, Wadelius C, Ek WE, et al. Contribution of genetics to visceral adiposity and its relation to cardiovascular and metabolic disease. Nat Med. 2019 Sep;25(9):1390–5.

38. Huang J, Huffman JE, Huang Y, Do Valle Í, Assimes TL, Raghavan S, et al. Genomics and phenomics of body mass index reveals a complex disease network. Nat Commun. 2022 Dec 29;13(1):7973.

39. Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. Nat Genet. 2020 Jul;52(7):680–91.

40. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015 Feb 12;518(7538):197–206.

41. Heard-Costa NL, Zillikens MC, Monda KL, Johansson A, Harris TB, Fu M, et al. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. PLoS Genet. 2009 Jun;5(6):e1000539.

42. McAllan L, Baranasic D, Villicaña S, Brown S, Zhang W, Lehne B, et al. Integrative genomic analyses in adipocytes implicate DNA methylation in human obesity and diabetes. Nat Commun. 2023 May 15;14(1):2784.

43. Lotta LA, Wittemans LBL, Zuber V, Stewart ID, Sharp SJ, Luan J ’an, et al. Association of Genetic Variants Related to Gluteofemoral vs Abdominal Fat Distribution With Type 2 Diabetes, Coronary Disease, and Cardiovascular Risk Factors. JAMA. 2018 Dec 25;320(24):2553–63.

44. Sieh W, Rothstein JH, Klein RJ, Alexeeff SE, Sakoda LC, Jorgenson E, et al. Identification of 31 loci for mammographic density phenotypes and their associations with breast cancer risk. Nat Commun. 2020 Oct 9;11(1):5116.

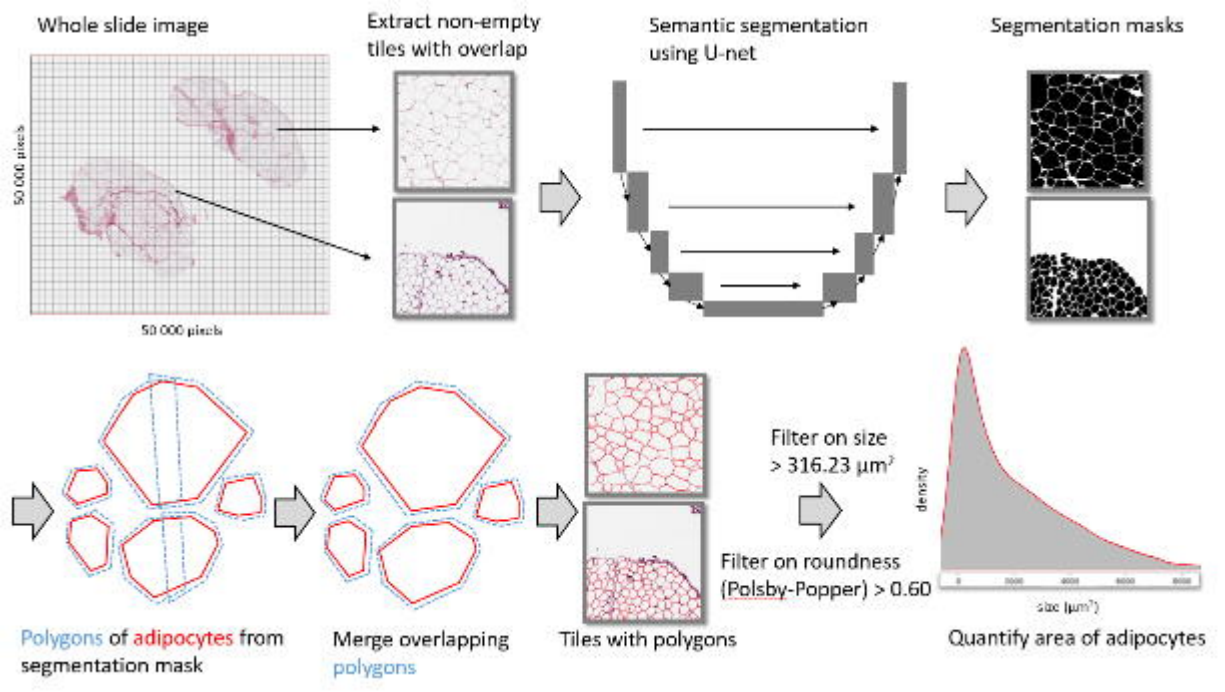
45. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. Hum Mol Genet. 2019 Jan 1;28(1):166–74.

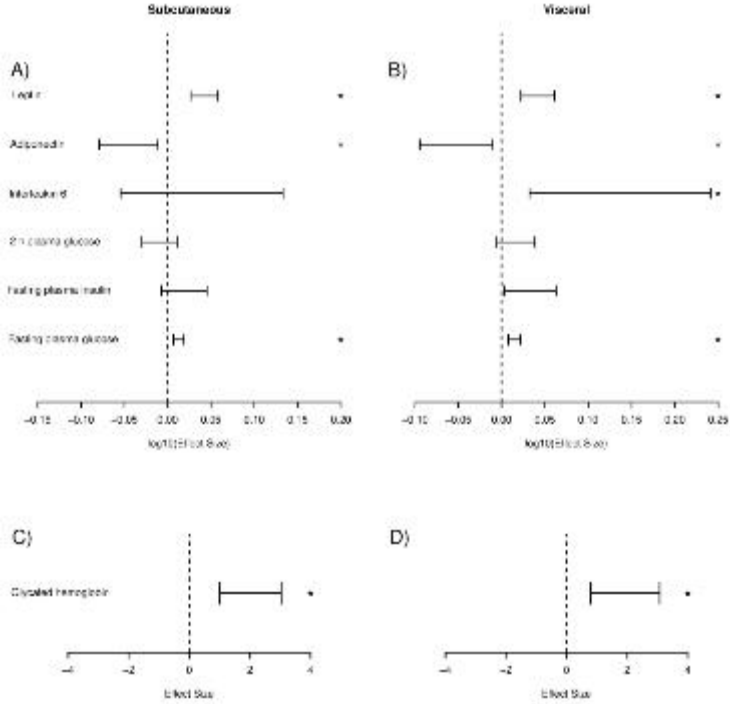
46. Turcot V, Lu Y, Highland HM, Schurmann C, Justice AE, Fine RS, et al. Protein-altering

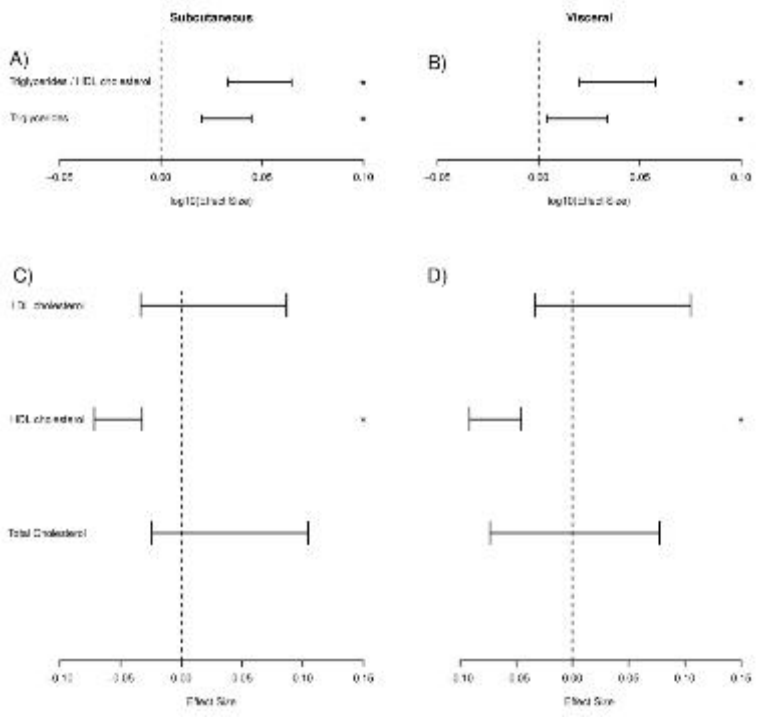
variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat Genet. 2018 Jan;50(1):26–41.

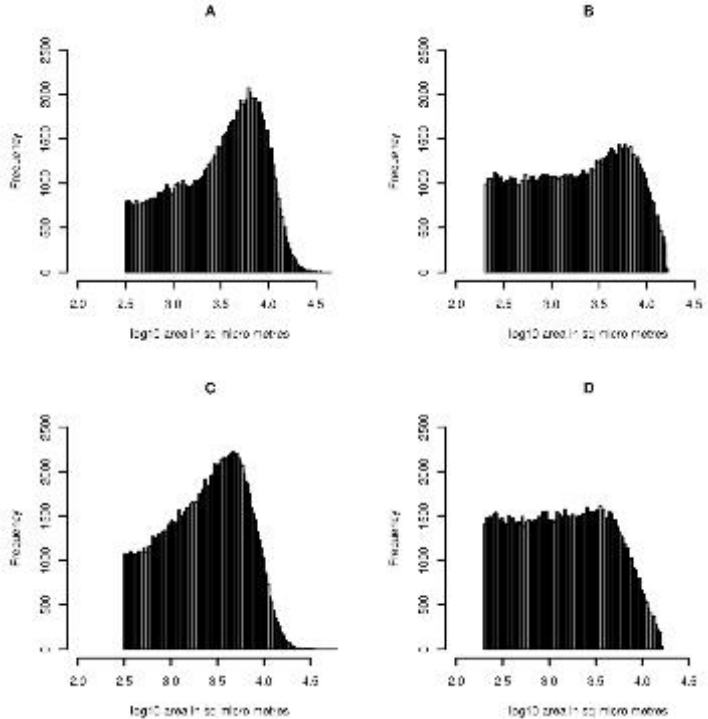
47. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015 Oct 1;526(7571):68–74.

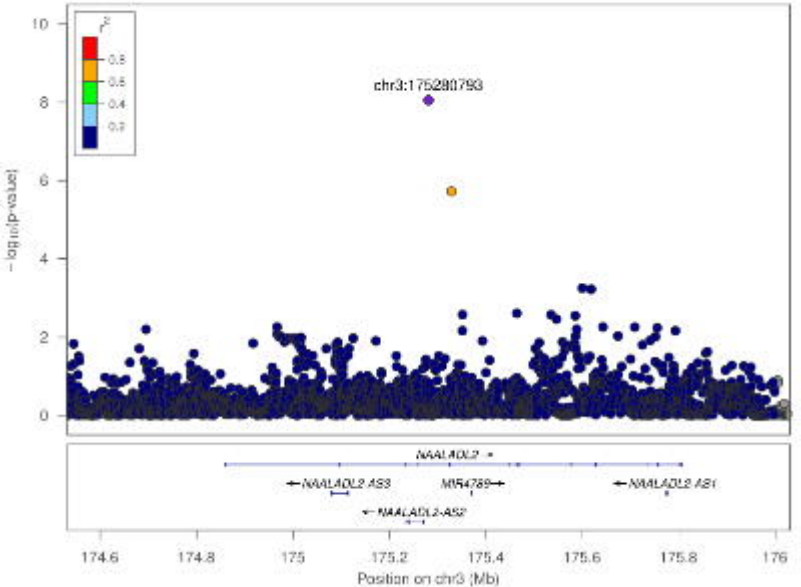
48. Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. Nature. 2021 Feb;590(7845):290–9.

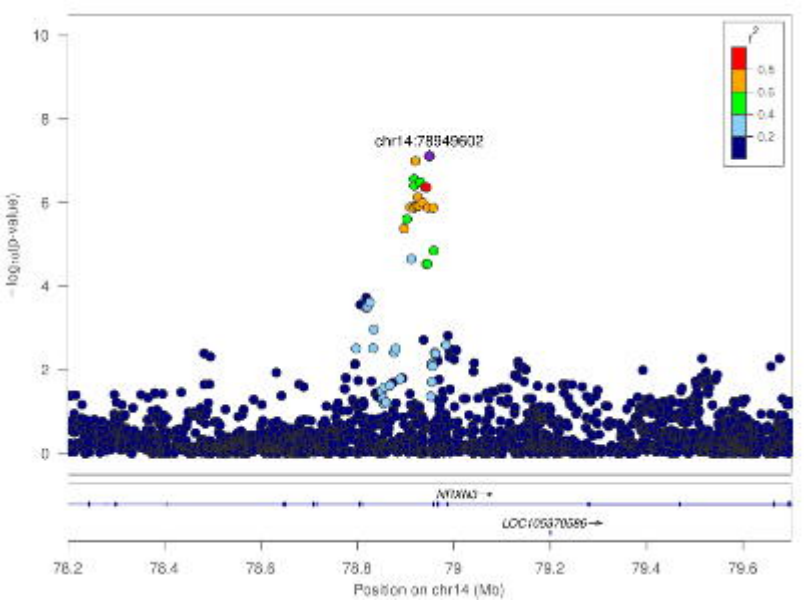

Whole slide image
50 000 pixels
50 000 pixels
Extract non-empty tiles with overlap
Semantic segmentation using U-net
Segmentation masks
Polygons of adipocytes from segmentation mask
Merge overlapping polygons
Tiles with polygons
Filter on size > 316.23 µm²
Filter on roundness (Polsby-Popper) > 0.60
density
size (µm²)
Quantify area of adipocytes


Subcutaneous
Visceral
A)
B)
C)
D)
Leptin
Adiponectin
Interleukin 6
2 h plasma glucose
Fasting plasma insulin
Fasting plasma glucose
Glycated hemoglobin
−0.15
−0.10
−0.05
0.00
0.05
0.10
0.15
0.20
0.25
−4
−2
0
2
4
Effect Size


Subcutaneous
Visceral
A)
B)
C)
D)
Triglycerides / HDL cholesterol
Triglycerides
HDL cholesterol
Total Cholesterol
-0.05
0.00
0.05
0.10
0.15
Effect Size


A
B
C
D
Frequency
0
500
1000
1500
2000
2500
2.0
2.5
3.0
3.5
4.0
4.5
log10 area in sq. micro metres


r²
0.8
0.6
0.4
0.2
chr3:175280793
- log10(p-value)
10
8
6
4
2
0
NAALADL2
NAALADL2-AS3
MIR4789
NAALADL2-AS1
NAALADL2-AS2
174.6
174.8
175
175.2
175.4
175.6
175.8
176
Position on chr3 (Mb)


chr14:78849602
r²
0.8
0.6
0.4
0.2
-log10(p-value)
10
8
6
4
2
0
NYAP2
LOC105370586
78.2
78.4
78.6
78.8
79
79.2
79.4
79.6
Position on chr14 (Mb)