



# Genomic interrogation of familial short stature contributes to the discovery of the pathophysiological mechanisms and pharmaceutical drug repositioning

Henry Sung-Ching Wong<sup>1</sup>, Ying-Ju Lin<sup>2,3</sup>, Hsing-Fang Lu<sup>1,4</sup>, Wen-Ling Liao<sup>5,6</sup>, Chien-Hsiun Chen<sup>7</sup>, Jer-Yuarn Wu<sup>7</sup>, Wei-Chiao Chang<sup>1,8,9,10\*</sup> and Fuu-Jen Tsai<sup>2,3,11,12\*</sup>

## Abstract

**Background:** Genetic factors, dysregulation in the endocrine system, cytokine and paracrine factors are implicated in the pathogenesis of familial short stature (FSS). Nowadays, the treatment choice for FSS is limited, with only recombinant human growth hormone (rhGH) being available.

**Methods:** Herein, starting from the identification of 122 genetic loci related to FSS, we adopted a genetic-driven drug discovery bioinformatics pipeline based on functional annotation to prioritize crucial biological FSS-related genes. These genes were suggested to be potential targets for therapeutics.

**Results:** We discovered five druggable subnetworks, which contained seven FSS-related genes and 17 druggable targets.

**Conclusions:** This study provides a valuable drug repositioning accompanied by corresponding targetable gene clusters for FSS therapy.

**Keywords:** Genome-wide association study, Familial short stature, Single-nucleotide polymorphism, Pharmacogenomics, Drug repositioning/repurposing

## Background

Individuals whose body height is in the 3rd percentile or greater below the mean of the population (of the same gender and chronologic age) are defined as short stature (SS). Several mechanisms including endocrine regulation (growth hormone, insulin-like growth factor-1, androgens, and thyroid hormone), proinflammatory cytokines, and paracrine factors have been identified as regulating linear growth [1–3]. Genetic factors account for ~80% of variations in human body height [4]. A systematic evaluation of human height genetics through a genome-wide association study (GWAS) uncovered 697 variants, located in 423 loci [5]. Subsequently, those discoveries

were extended to rare and very rare variants (with minor allele frequencies [MAFs] of 0.1%~4.8%) [6]. In addition, many genetic loci were found to be associated with human height across different populations [7–15], revealing the intricate polygenic architecture that determines human height.

Familial short stature (FSS), also known as “genetic SS”, is found in 23%~37% of individuals with SS [16, 17] and is characterized by patients with an SS family history, but normal growth. FSS is one of the most common types of SS and is solely affected by inheritance, thus making it a suitable candidate for identifying genetic loci associated with SS. We can rule out other pathologic causes of growth failure that may potentially confound genetic studies. Based on this idea, an association study of FSS-associated genetic variants in a Taiwanese population was conducted [17]. In that study, six

\* Correspondence: [wcc@tmu.edu.tw](mailto:wcc@tmu.edu.tw); [d0704@mail.cmuh.org.tw](mailto:d0704@mail.cmuh.org.tw)

<sup>1</sup>Department of Clinical Pharmacy, School of Pharmacy, Taipei Medical University, Taipei, Taiwan

<sup>2</sup>Genetic Center, Department of Medical Research, China Medical University Hospital, Taichung, Taiwan

Full list of author information is available at the end of the article



FSS risk genes, including *ZBTB38*, *ZNF638*, *LCORL*, *CABLES1*, *CDK10*, and *TSEN15*, were reported.

Recombinant human growth hormone (rhGH) is currently the only available treatment for SS. However, the efficacy of using rhGH for normal SS remains inconclusive, with some studies showing positive results [18, 19], while others did not [20, 21]. Accordingly, new therapeutics for SS are needed, and new approaches are warranted to expedite treatment. Nowadays, tremendous unveiled genetic loci have been united in tandem with various biological resources and functional annotation methodologies to identify novel drug targets and provide insights for drug repositioning [22, 23]. Hence, genetic loci characterized as being associated with FSS may ultimately be a good starting point for implementation of drug repositioning for SS patients.

In this study, we inquired into the biological and functional links of 122 FSS-associated single-nucleotide polymorphisms (SNPs) in a Taiwanese population and framed an annotation-based analytical pipeline to prioritize FSS-related genes that have the potential to be exploited as drug targets, and appraised the capacity of those drugs to be repurposed.

## Methods

### GWAS analysis of FSS cases and controls

Samples who fulfilled the diagnostic criteria of FSS were recruited from Children Hospital, China Medical University. The FSS was diagnosed by clinicians with the following criteria, including body height less than 3rd percentile to the population with corresponding age, and with a family history of short stature. In addition, only samples with ordinal annual growth rate and coincide bone and chronologic age will be included in this study. The controls in this study were selected from Taiwan Biobank based on their body height, i.e., >75th of all samples. We obtained informed consent from all study participants and guardians. This study was performed in accordance with approved guidelines and regulations.

In sample-level quality control (QC) step, for the 827 FSS patients, we removed 30 duplicated samples, two samples with data quality center (DQC) < 0.82, and 7 samples with call rate < 97%. For the remaining 788 samples, 52 were filtered in kinship QC step and left 736 samples for association analysis. For the controls from Taiwan Biobank, after removing samples with DQC < 0.82, failed plate QC, failed sample QC, missing gender and age information and failed kinship check, resulting in 464 remained for downstream analysis.

In marker-level QC step, for the 628,132 autosomal SNPs, we excluded the SNPs with MAF < 5%, SNP call rate < 98% in either case or control groups, Hardy-Weinberg equilibrium test *p*-value < 0.0001 (based on controls), and with batch effect. The remaining 530,030

(84.38%) SNPs were subjected to association analysis under additive inheritance model.

### Functional annotation of FSS-related SNPs

The region of FSS-associated SNPs (human genome hg19) was annotated using ANNOVAR [24]. The region of variants was categorized as either exonic, intronic, non-coding (nc) RNA intronic, the 5' untranslated region (UTR), the 3' UTR, intergenic, upstream, or downstream. For variants located in an exonic region, we further characterized their functional type, i.e., synonymous or non-synonymous.

### Identifying SNPs in linkage disequilibrium (LD) with FSS-related variants

For the 122 FSS-associated variants identified from a GWAS of a Taiwanese population, SNPs that were in high LD to these variants were identified using the 1000 Genome [25] phase 3 database (dbSNP Build 137). SNPs with an  $r^2$  value (a measure of LD) of > 0.8 and within a 100-kilobase (kb) window of FSS-associated variants based on an East Asian (EAS) super-population were selected using the R *proxysnps* package.

### Conspectus of the drug repositioning analysis for FSS

In this study, we proposed a bioinformatics pipeline called SNP-heuristic and expression-based functional unifying network (Shefun) algorithm embodied by two major portions: (1) an SNP-heuristic part and (2) an expression-based functional unifying network part.

The first part is centralized on SNPs. By SNP-based annotations, we could obtain functional states (non-coding/non-synonymous/synonymous), chromatin state, and *cis*-regulation data of each SNP. These data provided two aspects of information for the second part of the Shefun algorithm: resolution of tissue-specificity and determination of "seed" genes. For tissue specificity, based on the enrichment of FSS-associated SNPs with an active chromatin state, we resolved the tissue type(s) for a coexpression analysis. In addition, genes with *cis*-expression quantitative trait locus (eQTL) annotation and/or with non-synonymous variant(s) located in it could be utilized as "seed" genes for network construction.

The second part of Shefun, which mainly focuses on genes, includes several consecutive analytical modus operandi as follows: the construction of tissue-specific expression-based networks; a subnetwork enrichment analysis to establish gene-phenotype relationships; drug repurposing by inferring drug-phenotype relationships; an over-representation analysis; and primary target annotation. All of these functional analyses are unified into a network scene.

### Non-synonymous, chromatin state segmentation and *cis*-eQTL annotations

FSS-associated SNPs (and SNPs in high LD with FSS-related SNPs) were queried in HaploReg (vers. 4.1) [26] using the 1000 Genome Phase 1 database and an Asian (ASN) population. The functional state, chromatin state segmentation (25-state), and *cis*-eQTL information were extracted from the output sheet of HaploReg.

SNPs with a chromatin state of 1~19 were defined as “active”; 20~25 as “inactive”, and the remaining as “not available” (n.a.). For each cell type, we calculated the number of SNPs with an active chromatin state, and calculated one-sided *p* values ( $Z = (N - \text{mean}(N))/SD(N)$ , where *N* is the number of SNPs with state 1~19 in the given cell type, and *SD* is the standard deviation) by comparing to the mean of the number of “active SNPs” across cell types (mean no. = 84.73).

For the *cis*-eQTL part, given the results from chromatin state segmentation, we selected only SNPs with *cis*-eQTL annotation in the following tissue types: whole blood, adipose (subcutaneous) tissues, adipose (visceral omentum) tissues, breast mammary tissue, skin (sun-exposed; lower leg), cells (transformed fibroblasts), muscle (skeletal), skin (not sun-exposed; suprapubic), osteoblasts (prostaglandin E2 (PGE2)), osteoblasts (bone morphogenetic protein 2 (BMP2)), osteoblasts (Dex.) and osteoblasts (untreated). We further merged tissue types into seven categories: adipose, blood, bone, breast, fibroblast, skeletal muscle, and skin.

The SNPs were categorized based on non-coding/non-synonymous/synonymous, the active/inactive chromatin state, and *cis*-eQTL, and visualized them by a radar chart using the R *fmsb* package.

### Genotype-tissue expression (GTEx) transcriptomic dataset pre-processing

GTEx expression data (five tissue types including adipose, breast, fibroblast, skeletal muscle, and skin) were downloaded from *recount2* (<https://jhubiostatistics.shinyapps.io/recount/>) and processed using the R *recount* package. Samples with an RNA integrity number (RIN) of < 6.0 were filtered. Next, gene expression values were aggregated by the average, and then log<sub>2</sub>-scaled (scaled  $E = \log_2(E + 1)$ , where *E* represents the gene expression value). Then, lowly expressed genes were removed by preserving genes with a scaled expression of > 1 in 80% of the samples in at least one tissue type. Finally, we performed a principal component analysis (PCA) adjustment for latent covariates, also known as surrogate variables, using the R *sva* package.

### Bone tissue dataset pre-processing

As GTEx did not include bone expression data, we thus downloaded a bone biopsy transcriptomic dataset (E-

MEXP-1618) of postmenopausal females from ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/experiments/E-MEXP-1618/>). The raw gene expression values were normalized using the R *gcrma* package.

### Expression-based network construction

The expression-based network (six tissue types, excluding “whole blood”) was consociated with two levels of information: (1) messenger (m) RNA coexpression and (2) protein-protein interactions (PPIs). To do this, for each selected tissue type, FSS-related genes (“seed” genes), constituted by tissue-specific eGenes (from *cis*-eQTL annotation) and genes that contained non-synonymous SNPs, served as input genes for a coexpression network analysis. For each input gene, genes with the top 10/15/20/25/30 highest Pearson’s product-moment correlation coefficient were included to build a subnetwork. Then, the subnetworks were further expanded using PPI information adopted from the Human Protein Reference Database (HPRD, vers. Release9\_041310) [27]. Furthermore, self-loops and redundant links were removed from each subnetwork for the sake of conciseness. Different subnetworks were fused into a bigger subnetwork if they contained at least one identical gene.

### Gene set enrichment analysis (GSEA)

The “pathways” for GSEA were the merged expression-based subnetworks, and the gene-level statistics were beta-coefficients (related to “height”) acquired from Taylor et al. (human skeletal muscle biopsies) [28]. The GSEA was conducted using the R *fgsea* package with 99,999 permutations. The significance threshold was set to a false discovery rate (FDR) of < 0.1. The subnetworks that reached a significant threshold were defined as “height-related subnetworks”. For each height-related subnetwork, genes within it were assigned a value of + 1 if the subnetwork was positively enriched (representing a positive “gene-phenotype relationship”) and - 1 if the subnetwork was negatively enriched (representing a negative “gene-phenotype relationship”).

### Ligand/drug repositioning

Ligand-target (gene) interaction data were queried from the Guide to PHARMACOLOGY website (<http://www.guidetopharmacology.org/download.jsp>, vers. 2019.3). Data were first filtered by the following criteria: (1) human species; (2) non-endogenous agents; (3) a clear type/action of the mechanism for each ligand-target pair; and (4) distinct target (gene symbol) information. We further removed the drug-gene pair of the actions of “binding”, “mixed”, and “neutral”. Next, we assigned a value of + 1 to the ligand-target pair of action of the mechanism of “activation”, “agonist”, “biased agonist”, “full agonist”, “partial agonist”, and “positive”; and also the type of mechanism of “activator” and “agonist”.

Similarly, we assigned a value of  $-1$  to ligand-target pairs with an action mechanism of “antagonist”, “feedback inhibition”, “inhibition”, “inverse agonist”, “irreversible inhibition”, “negative”, “pore blocker”, “slows inactivation”, and “voltage-dependent inhibition”; and mechanism types of “antagonist”, “channel blocker”, “gating inhibitor” and “inhibitor”. Consequently,  $+1$  or  $-1$  represents a positive or negative drug-gene relationship, respectively.

For each gene in the height-related subnetworks, the drug-phenotype relationship was inferred by multiplying the assigned values of “drug-gene relationship” and “gene-phenotype relationship”. There were four possibilities to show the logic of how we inferred the drug/ligand effect, i.e., “drug-gene relationship”  $\times$  “gene-phenotype relationship” = “drug-phenotype relationship”: (1)  $+1 \times +1 = +1$ ; (2)  $+1 \times -1 = -1$ ; (3)  $-1 \times +1 = -1$ ; and (4)  $-1 \times -1 = +1$ . A final value of  $+1$  suggests that the drug may enhance or exacerbate the phenotype of interest, and a final value of  $-1$  suggests that the drug may alleviate, diminish, or inhibit the phenotype of interest. The repositioning analysis revolved around genes in height-related subnetworks, and drugs/ligands were selected which possibly targeted those genes with a calculated value (drug-phenotype relationship) of  $+1$  only, as this meant that the selected drugs/ligands possibly enhanced the phenotype of interest (i.e., height) and therefore was a potential candidate for repurposing to FSS.

#### Gene ontology (GO) biological process (BP) terms and Kyoto encyclopedia of genes and genomes (KEGG) pathway over-representation analysis (ORA)

Height-related subnetwork genes were subjected to a GO analysis [29] to assess their enrichment in BP terms. The enrichment test was performed using “weight01” implemented in the R *topGO* package. Moreover, the KEGG ORA test was performed using the R *clusterProfiler* package. The Benjamini-Hochberg (BH) method was applied for multiple test corrections.

#### Statistical and bioinformatics analysis

All in-house statistical and bioinformatics scripts for drug repositioning analysis were written in R language (<https://www.r-project.org/>). Gene symbols from different sources were unified using the R *HGNChelper* package. The conversion between gene symbols, Entrez Gene ID, and Ensembl Stable ID was performed using the R *clusterProfiler* package. The networks were illustrated using the R *igraph* package utilizing the Fruchterman-Reingold (FR) algorithm.

## Results

### Genome-wide association and genotyping approaches reveal a total of 122 FSS-associated SNPs

To determine novel susceptible genetic loci of FSS, FSS patients ( $n = 788$ , male = 51.91%) from Children’s Hospital,

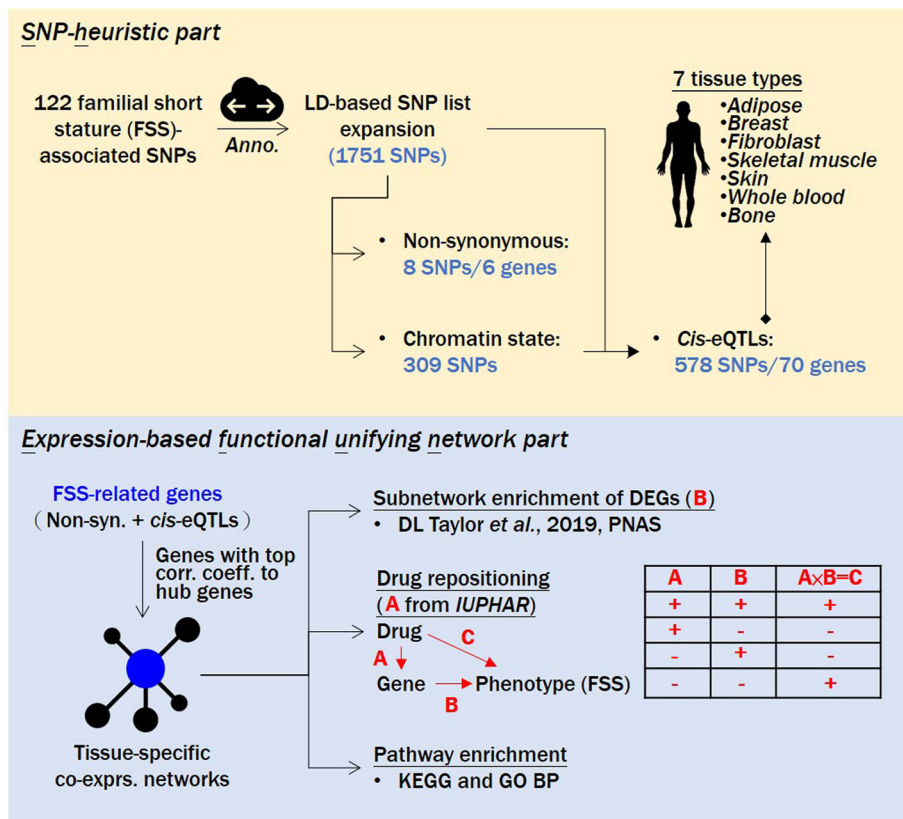
China Medical University were enrolled. The diagnosis of these patients (cases) was made by clinicians according to the diagnostic criteria of FSS (Additional file 1: Fig. S1). The patients with growth hormone deficiency were excluded from this study. The controls ( $n = 435$ , male = 42.67%) were from Taiwan Biobank that whose height was above the 75th (Q3) of the total population. Both cases and controls were Han Chinese population residing in Taiwan. After sample-level and marker-level quality control, 530,030 SNPs were subjected to initial genome-wide association screening under the additive inheritance model. Multidimensional scaling (MDS) was performed and no significant population stratification was found (Additional file 2: Fig. S2). As shown in Additional file 3: Fig. S3, significant associations between genetic loci and FSS were observed. In total, we identified 14 genome-wide significant ( $p < 5 \times 10^{-8}$ ) SNPs in the genome-wide screening of FSS cases and controls (Additional file 6: Table S1), including rs822611 (Chr 1), rs6731651 (Chr 2), rs16828530 (Chr 3), rs9290657 (Chr 3), rs10028040 (Chr 3), rs1863593 (Chr 8), rs16900402 (Chr 8), rs28786672 (Chr 9), rs7852806 (Chr 9), rs2172912 (Chr 12), rs12826453 (Chr 12), rs9520911 (Chr 13), rs17732181 (Chr 17), and rs4815179 (Chr 20). In present study, we also identified the top 88 genetic loci (Additional file 6: Table S1 with  $p < 10^{-4}$ ). These 88 novel genetic loci were located in the 44 closest genes. Among these 44 closest genes, eight genes have at least two SNPs within the same gene. These eight closest genes included *AGO4*, *SESTD1*, *PARD3B/ICOS*, *RFC1*, *UNC5C*, *IL7*, *BCL11B*, and *MIAT/MN1*. Among them, *BCL11B*, *IL-7*, *MN1*, and *UNC5C* are involved in embryonic, connective tissue, organ development, and developmental disorders.

Moreover, our previous study suggested 34 SNPs that were also associated with an FSS risk [17]. These 34 human height-related SNPs were located in the 13 closest genes. These 13 closest genes included *TSEN15*, *EFEMP1*, *ZNF638*, *CEP63*, *ZBTB38*, *LCORL*, *HHIP*, *ANAPC10*, *GSDMC*, *QSOX2*, *ADAMTSL3*, *CDK10*, and *CABLES1* that also involved in embryonic, organismal, and tissue development.

### Functional annotations of 122 FSS-associated SNPs

To identify input genes for the downstream analyses, we consolidated several SNP annotation criteria to map the SNPs to genes (Fig. 1 [top]). In the 122 FSS-associated SNPs, most were located in intronic ( $n = 53$ , 43.44%) and intergenic ( $n = 58$ , 47.54%) regions (Additional file 7: Table S2). Among 122 SNPs, four SNPs were located in an exonic region (Additional file 8: Table S3).

As GWAS and genotyping approaches selected the genotyped SNPs using an LD-tagging method, it could potentially miss causal SNPs that are linked to FSS. Therefore, we expanded the SNP list by querying SNPs in high LD ( $r^2 > 0.8$  within a 100-kb window) with our



**Fig. 1** Schematic showing an overview of the drug repositioning pipeline in this study

SNP list using the 1000 Genome (phase 3, vers. 5a) EAS database, resulting in 1751 SNPs (121 FSS-associated SNPs and 1630 SNPs in LD with FSS-associated SNPs, where rs10086016 was excluded due to a lack of gene annotation). With the expanded SNP list, we next queried their (1) exonic function, (2) chromatin state segmentation (25-state), and (3) *cis*-eQTL information using HaploReg (vers. 4.1) (Fig. 2).

As a result, we identified six genes (*CALCOCO2*, *MUC16*, *TSEN15*, *DCAF16*, *GSDMC*, and *ADAMTSL3*) in which eight non-synonymous SNPs were located (Fig. 2 [left] and Additional file 9: Table S4). In addition, among 1751 SNPs, we found 309 (17.65%) SNPs with at least one active chromatin state segmentation (states 1~19) annotation. These SNPs were enriched ( $p < 0.1$ ) in different cell types including adipocytes, skeletal muscle cells, bone marrow-derived cells, skin melanocytes, mammary epithelial cells, and bone-related cells such as osteoblasts and chondrocytes (in total 16 cell types, with brain-related cell types excluded; Fig. 2 [middle], Additional file 4: Fig. S4, and Additional file 10: Table S5).

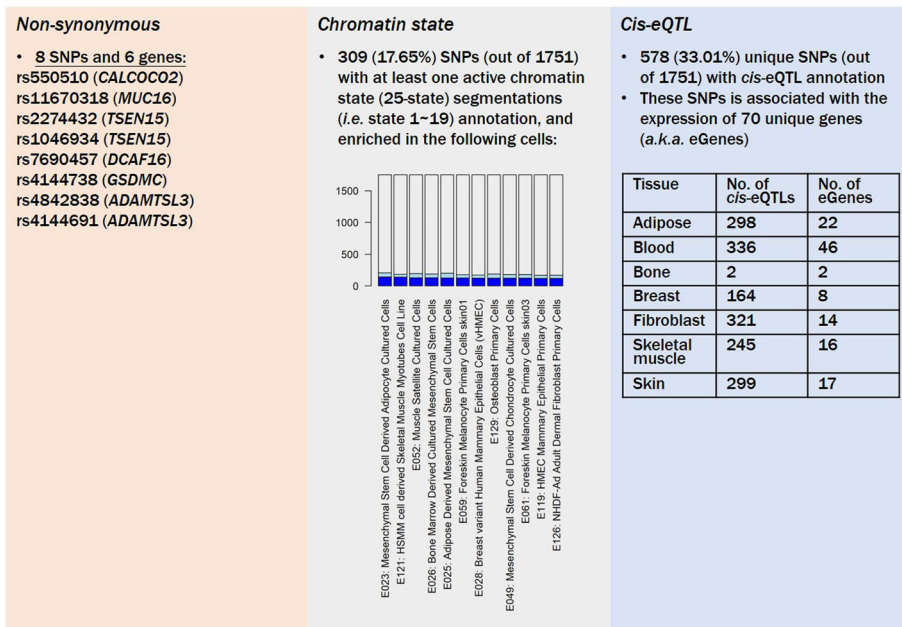
Based on these findings, we focus on seven tissues including adipose, blood, bone, breast, fibroblast, skeletal muscle, and skin to seek SNPs with *cis*-eQTL annotation, and identified 298 (17.08%), 336 (19.19%), 2

(0.11%), 164 (9.37%), 321 (18.33%), 245 (13.99%), and 299 (17.08%) *cis*-eQTLs, respectively. In total, these 578 (33.01% of 1751) *cis*-eQTLs were correlated to 70 unique eGenes. In greater detail, the numbers of eGenes in each tissue type were 22, 46, 2, 8, 14, 16, and 17, respectively (Fig. 2 [right] and Additional file 5: Fig. S5). However, the number of eGenes shared among different tissues was relatively low (Fig. 3), suggesting the uniqueness of the SNP-gene regulation machinery.

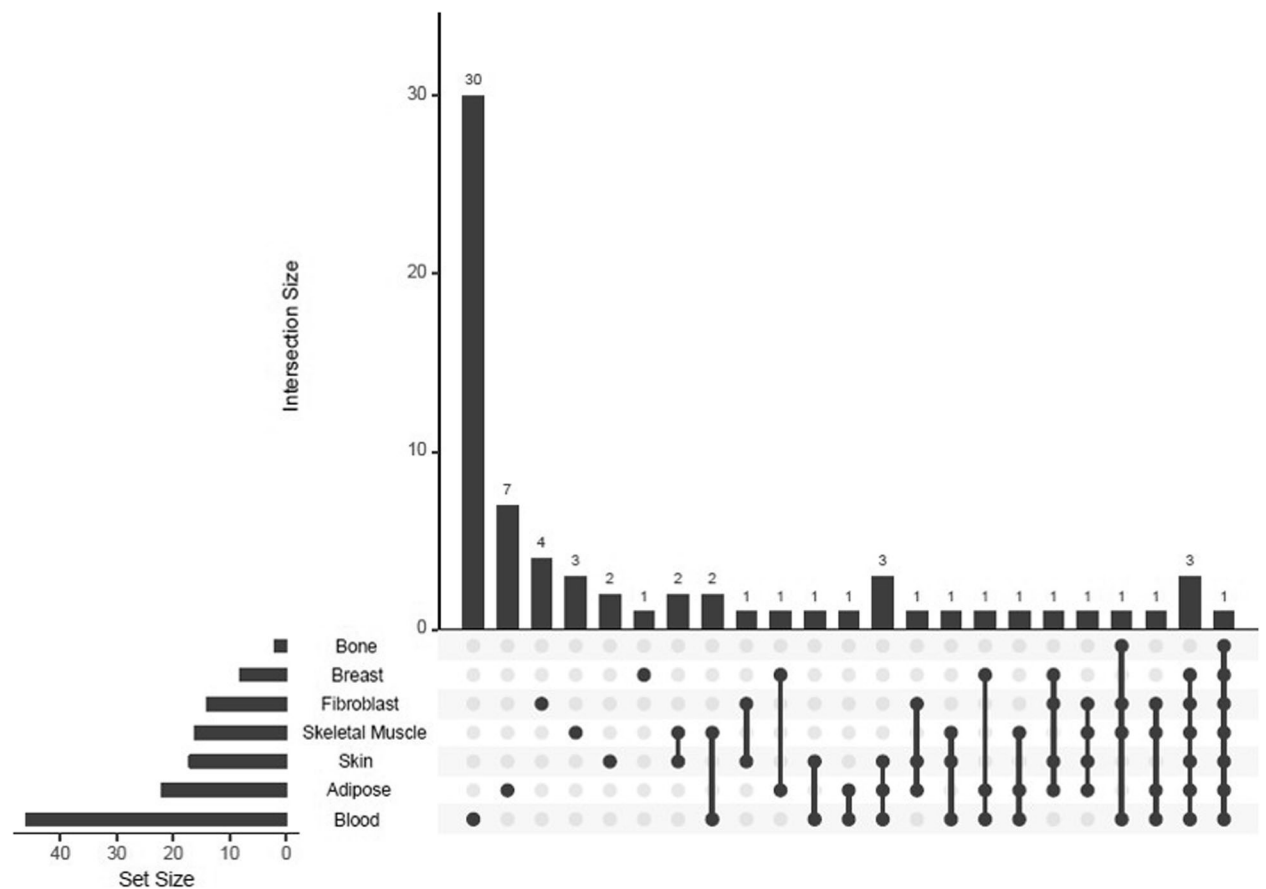
Overall, we categorized the SNPs based on annotations, including the functional state (non-coding/non-synonymous/synonymous), chromatin state segmentation (25 states), and *cis*-regulation (Fig. 4).

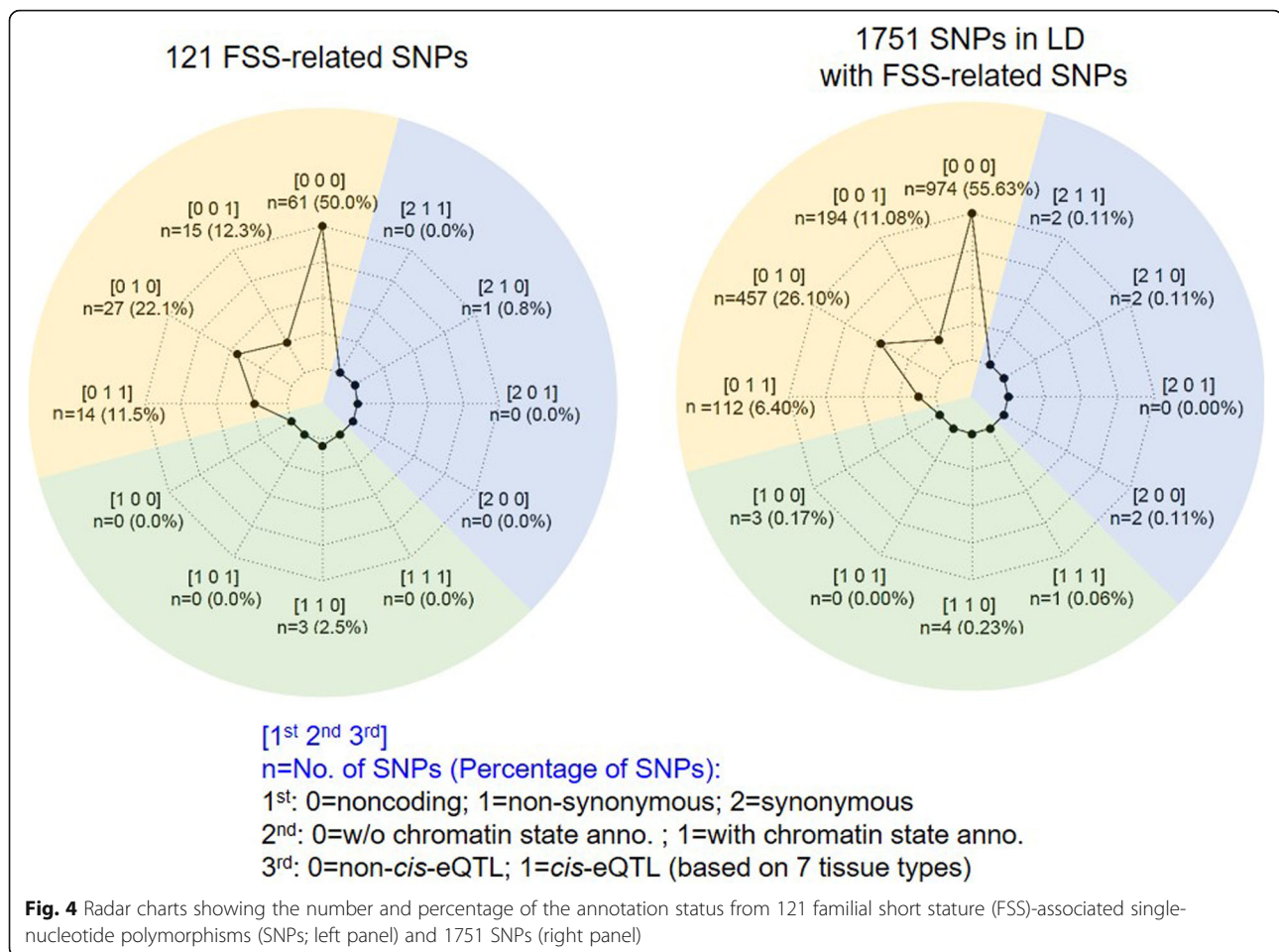
#### Construction of expression (mRNA-coexpression and PPI)-based networks

Given the hypothesis that genes collaborate together to form functional units and to regulate a specific phenotype/pathology (in this case, FSS), we next utilized two published transcriptomic datasets (GTEx [vers. 7] for adipose, breast, fibroblast, skeletal muscle, and skin tissues; and E-MEXP-1618 for bone tissue) to capture the cooperating unit by constructing a so-called “expression-based network”.



**Fig. 2** Schematic showing results of the single-nucleotide polymorphism (SNP)-heuristic part analysis. The barplot in the middle panel shows cell types that were significantly enriched in SNPs with an active chromatin state. Blue color indicates SNPs with chromatin state segmentation of 1~19; the light-blue color indicates SNPs with chromatin state segmentation of 20~25; while the remaining have no available annotations





To do this, FSS-related genes (composed of tissue-specific eGenes and genes with a non-synonymous annotation) served as “seed” genes for network construction. For each tissue type, we created a network by calculating Pearson’s product-moment correlation coefficients between each of the “seed” genes and the other genes. To focus on the most relevant coexpression links and also to take network robustness into consideration, we identified the top 10/15/20/25/30 coexpressed genes with the highest correlation to each “seed” gene. In addition, the networks were further expanded using HPRD (vers. Release9\_041310) PPI information. We investigated genes with PPIs with each “seed” gene and included them in the network. In total, we generated  $6 \times 5 = 30$  expression-based networks (Fig. 1 **bottom**).

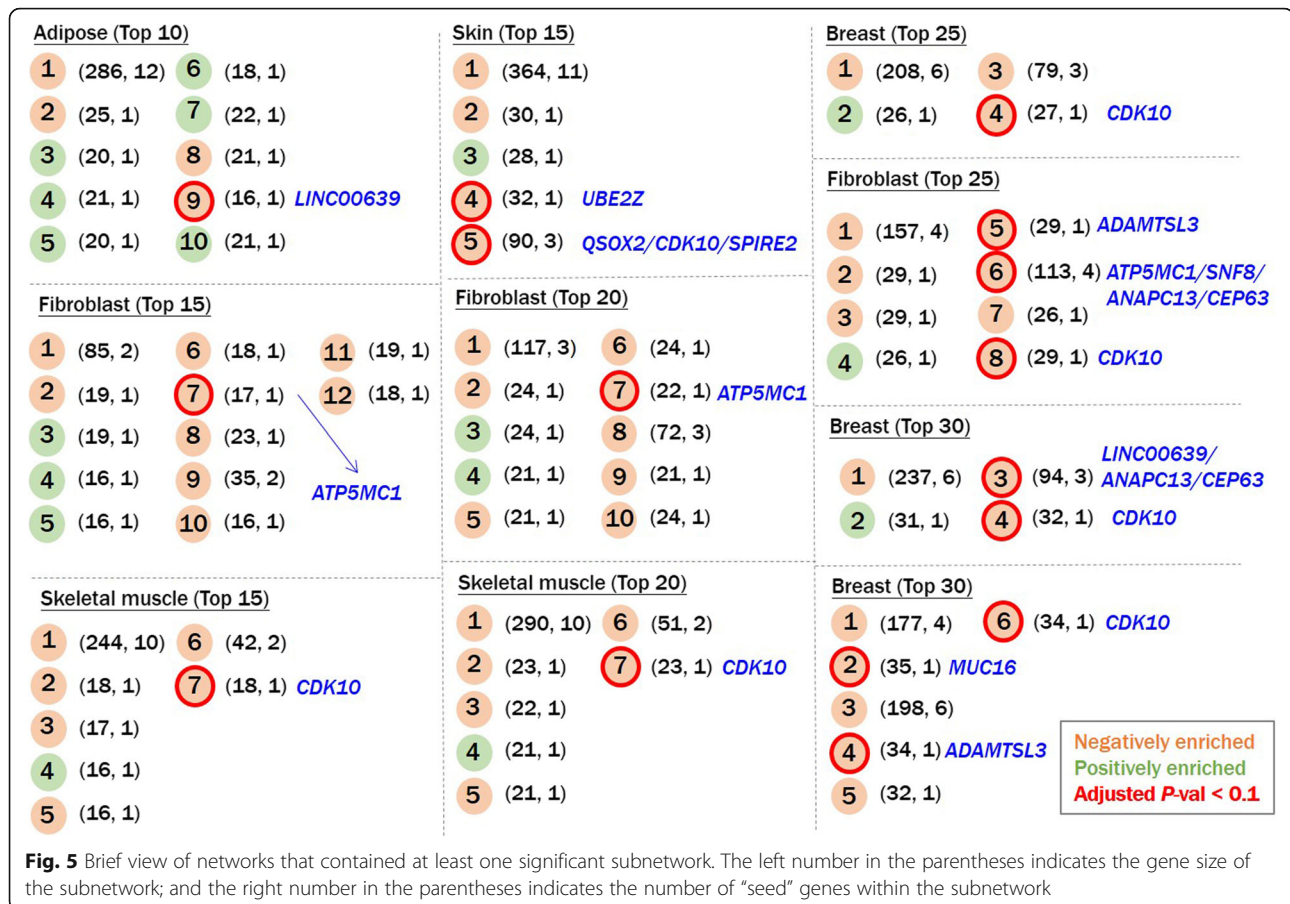
#### Identification of subnetworks that were positively or negatively enriched in height-related genes

To clarify the gene (integrated as a network)-phenotype relationship, we leveraged differentially expressed data related to the height from Taylor et al. [28] and performed a subnetwork-based GSEA. In the tissue-specific networks, each “seed” gene was linked with coexpression

genes and/or PPI genes to form a subnetwork, which was possibly merged into a larger subnetwork if it contained at least one identical gene member with another subnetwork. For each amalgamated subnetwork, we conducted the GSEA (permutation no. = 99,999) by incorporating differential expression information, i.e., genes’ beta-coefficient statistics to the height. Significantly enriched (adjusted  $p < 0.1$ ) subnetworks were defined as “height-related subnetworks”. 16 height-related subnetworks across 10 (33.3%) of 30 networks were identified, with network sizes ranging 16–113, and the number of “seed” genes ranging one to four. Notably, all identified height-related subnetworks were inversely correlated (negatively enriched) with expressions of genes that were positively associated with height (Fig. 5).

#### Drug repositioning to FSS by targeting height-related subnetworks

To integrate the direction of a drug’s effect on FSS into our pipeline, in other words, to elucidate drug-phenotype relationships, we incorporated (1) interaction data for ligands and targets (drug-gene relationship) from the Guide to PHARMACOLOGY database (vers. 2019.3) and (2)



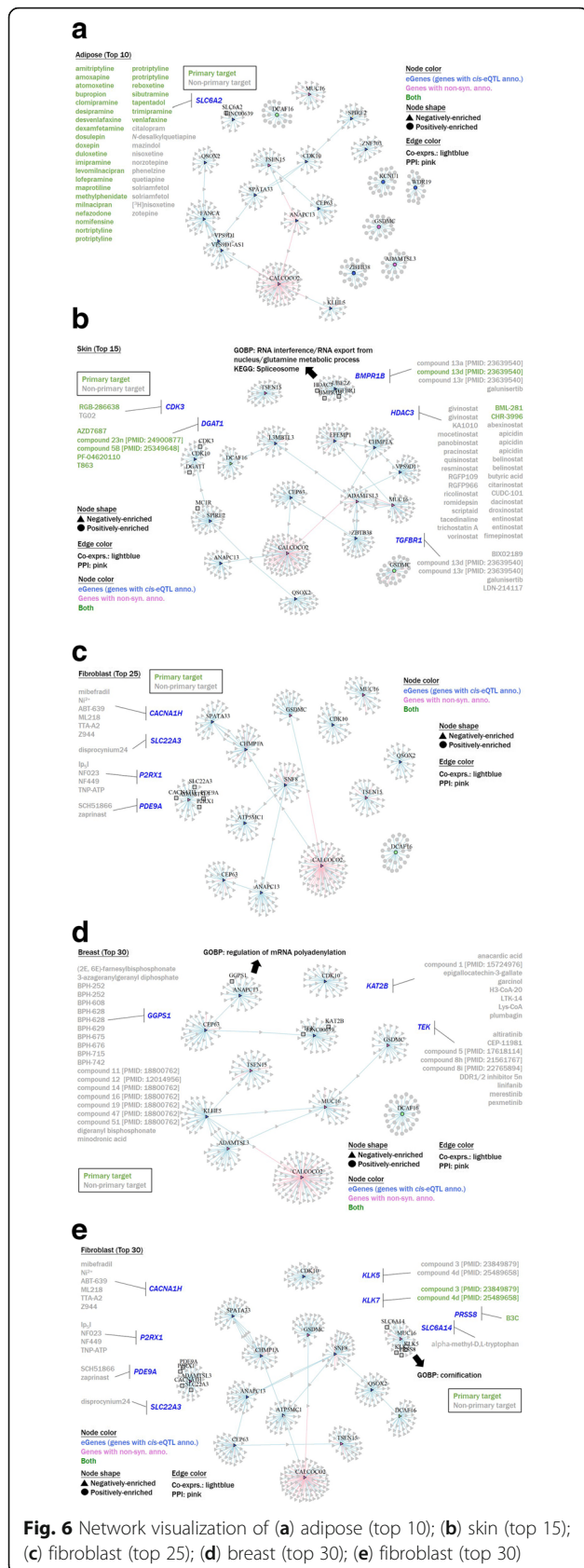
predefined gene-phenotype relationships (Fig. 1 [bottom]). Given the Shefun pipeline, we determined that five of 30 networks (with seven different subnetworks spanning four tissue types) possessed repurposing potential, including (1) adipose (top 10) containing 39 ligand-gene pairs (Fig. 6a). In this network, *SLC6A2*, a norepinephrine transporter (NET) gene was identified as a potential drug target for SS repositioning. (2) Skin (top 15) containing 58 ligand-gene pairs (Fig. 6b). Two drug-targeted subnetworks were identified: one containing the drug-targeted genes *CDK3* and *DGATI* and the other containing *BMPRI1B*, *HDAC3*, and *TGFBR1*. (3) Fibroblast (top 25) containing 13 ligand-gene pairs (Fig. 6c). *CACNAIH*, *SLC22A3*, *P2RX1*, and *PDE9A* were identified as drug-targeted genes in this network. (4) Breast (top 30) containing 40 ligand-gene pairs (Fig. 6d) and drug-targeted genes such as *GGPS1*, *KAT2B* and *TEK*. (5) And, fibroblast (top 30) containing 19 ligand-gene pairs (Fig. 6e). In this network, two subnetworks were found to be potential candidates for drug repurposing, with one subnetwork containing the drug-targeted genes *KLK5*, *KLK7*, *PRSS8*, and *SLC6A14* and the other subnetwork containing *CACNAIH*, *P2RX1*, *PDE9A*, and *SLC22A3*. Therefore, these drugs/ligands could be candidates for further investigation. Given that some of the genes from the ligand-gene pairs that we

identified might not be the primary target of the specific ligands, and might thus indicate possible safety issues, we therefore annotated information of “primary target” or “non-primary target” for each ligand-gene pair. This information may assist in the future prioritization of drugs/ligands for FSS repositioning.

#### Pathways and biological processes over-representing drug-targeted subnetworks

For height-related subnetworks that contained the drug-targeted gene(s), we conducted GO BP terms and KEGG pathway ORA (Additional file 11: Table S6). The significant (with an FDR of < 0.1) BP terms and pathways are illustrated in Fig. 6a-e. For the skin (top 15), a subnetwork centered on *UBE2Z* (a “seed” gene) showed significant enrichment in RNA interference, RNA export from nuclei, glutamine metabolic process terms, and the spliceosome pathway (Fig. 6b). Another subnetwork (centered on *ANAPC13*) of the breast (top 30) also showed significant enrichment in the regulation of mRNA polyadenylation (Fig. 6d). In addition, a *MUC16*-centered subnetwork in the fibroblast (top 30) network showed significant enrichment in the cornification term (Fig. 6e).





**Fig. 6** Network visualization of (a) adipose (top 10); (b) skin (top 15); (c) fibroblast (top 25); (d) breast (top 30); (e) fibroblast (top 30)

## Discussion

In this work, we integrated several biological resources to prioritize FSS-related genetic variants and identified candidate druggable genes for FSS. Using a bioinformatics pipeline, we first annotated FSS-related variants and mapped those variants to genes (in the SNP-heuristic part). Next, we conducted gene-based annotations and prioritized genes in a network-based manner (in the expression-based functional unifying network part). As a result of this study, we reported five candidate networks for drug repositioning comprised of seven unique FSS-related genes (“seed” genes) including *LINC00639*, *CDK10*, *SPIRE2*, *QSOX2*, *ADAMTSL3*, *ANAPC13*, and *CEP63*. Overall, we identified 17 unique druggable genes.

Some of the determined druggable genes were reported to be directly associated with SS according to the Human Phenotype Ontology (HPO; the identity of SS: HP:0004322) and Gene-Disease Associations (GAD) databases, as exemplified by *SLC6A2* [30], a member of the Na<sup>+</sup>:neurotransmitter symporter family, which is targeted by some antipsychotic agents. Likewise, *BMPRI1B*, a member of the bone morphogenetic protein (BMP) receptor family of transmembrane serine/threonine kinases, which belongs to the transforming growth factor (TGF)-β superfamily, was reported to be associated with acromesomelic dysplasia [31]. It is noteworthy that the BMP and TGF-β signaling pathways were suggested to play central roles in human growth, and hence are linked to the mechanism of the development of SS [32, 33]. *TGFBRI*, a gene that forms a heteromeric complex with the *TGFB2* protein, was also identified as a drug target of several TGF-β inhibitors for FSS repositioning in this study.

Additionally, we identified a number of druggable genes that may interact with known SS-related genes, despite they themselves are lacking of known associations with FSS, including *CDK3* (which interacts with *CABLES1*), *TGFBRI* (which interacts with *TGFB3*), *PDE9A* (which interacts with *HPRT1*), *TEK* (which interacts with *PIK3R1*), and *KLK7* (which interacts with *CDSN*). These genes were considered to be “indirectly” linked to FSS and might have potential to serve as targets for repurposing.

Furthermore, our results demonstrated several biologically meaningful gene clusters in drug repositioning for FSS: two groups of GOBP were related to the development biology pathway: one is a subnetwork in the network of “breast” (top 30), which contains *GGPS1*, *KAT2B*, and *TEK*. Specifically, *TEK* may interact with the SS-related gene, *PIK3R1*, which codes an enzyme that phosphorylates the 3’ position of the inositol ring of phosphatidylinositol [34]. *KAT2B*, a gene that associated with p300/CBP, mediates *PLK4* acetylation and thus

acts as a negative regulator of centrosome amplification [35]. Notably, *PLK4* is also an SS-related gene. Importantly, we identified several acetyltransferase inhibitors that may target *KAT2B*, including anacardic acid, garcinol, plumbagin, and so on. The other gene cluster was located in the network of “fibroblast” (top 30), which contains *KLK5*, *KLK7*, *PRSS8*, and *SLC6A14*. In addition, *GGPS1*, a member of the prenyltransferase family, which encodes an enzyme that catalyzes the synthesis of geranylgeranyl diphosphate from farnesyl diphosphate and isopentenyl diphosphate, was associated with osteogenesis imperfecta. In addition, *GGPS1* was also reported to be correlated with the bone mineral density [36] and atypical femoral fractures [37]. In this study, we identified bisphosphonates that may target *KAT2B*. In addition, B3C, an activator of the epithelial sodium channel ENa, may target *PRSS8*. In short, we revealed several promising drugs, providing reasonable druggable gene clusters for FSS based on this genomic interrogation platform.

Nevertheless, we discovered two similar subnetworks in the “fibroblast” (top 25) and “fibroblast” (top 30), which contained druggable genes (*CACNA1H*, *SLC22A3*, and *P2RX1*) that implicated in cation (calcium) homeostasis regulation, however, these genes have no clear connection to SS or FSS. Interestingly, a gene belonging to the above-mentioned subnetworks, *PDE9A*, is able to interact with *HPRT1*, which encodes an enzyme that is crucial for the generation of purine nucleotides through the purine salvage pathway, and is thus associated with SS. Therefore, our analysis may unearth previously unknown mechanisms/pathways of FSS which in turn, provides new insights for drug repositioning. Obviously, the findings need further rigorous experiments for validation.

The genome-wide scale association analysis that scanned the entire genome without bias provided an unprecedented opportunity for drug repurposing by linking disease indications with druggable genes, i.e., “genetics-driven genomic drug discovery” [22, 38, 39], which is exemplified by the identification of *PCSK9* for the treatment of hypercholesterolemia [40]. We thus postulated that our “FSS-associated variants” should be subjected to a drug-repositioning analysis. We, therefore, leveraged the Guide to PHARMACOLOGY database to identify potential therapeutic agents that were initially developed for other diseases that may be repurposed to alleviate FSS. In addition, we showed the plausibility of drug target identification by using genomic approaches.

However, we noted several limitations. First, in GWAS part, false positives associations may not be excluded due to small power of current study. Second, further functional investigations are needed to validate the candidate drug targets identified by our annotation-

based analytical pipeline. Third, the affinity and specificity of drugs that target SS-related genes may differ. Further experiments are required to select suitable drugs. Fourth, some druggable genes (e.g., *SLC6A2*, *CDK3*, and *TEK*) were the targets of antipsychotic/anticancer agents, which may generally lead to more-severe adverse events. Therefore, in order to balance the risk and benefits, we emphasize that the genes targeted by safer agents should initially be prioritized to assess their clinical potential for repositioning to FSS.

## Conclusions

In summary, we prioritized seven candidate FSS-related genes (*LINC00639*, *CDK10*, *SPIRE2*, *QSOX2*, *ADAMTSL3*, *ANAPC13*, and *CEP63*) and 17 genes (*SLC6A2*, *CDK3*, *DGAT1*, *BMPRI1B*, *HDAC3*, *TGFBR1*, *CACNA1H*, *SLC22A3*, *P2RX1*, *PDE9A*, *GGPS1*, *KAT2B*, *TEK*, *KLK5*, *KLK7*, *PRSS8*, and *SLC6A14*) for drug repurposing. Among them, drugs targeting *DGAT1*, *HDAC3*, *PDE9A*, *GGPS1*, *KAT2B*, *KLK5*, *KLK7*, *PRSS8*, and *SLC6A14* were recommended for repurposing not only due to the consideration of plausible mechanistic explanations but also after taking safety issues into evaluation. This study provides insights for understanding the pathophysiology of FSS and thereby conferring new approaches for drug discovery. Finally, our study demonstrated the power of comprehensive genomic interrogation in drug discovery for human diseases.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12929-019-0581-2>.

**Additional file 1: Fig. S1.** Gender-specific (male [top] and female [bottom]) age distribution curves of FSS patients (in this study; red color with solid line) and normal Taiwanese population (grey and black colors with solid and dashed lines). (DOCX 165 kb)

**Additional file 2: Fig. S2.** Multidimensional scaling (MDS) results. MDS plots of cases (CA) and controls (CN) in this study with (top) or without (bottom) other population from the 1000 Genome Database. (DOCX 77 kb)

**Additional file 3: Fig. S3.** Genome-wide association screening results. Manhattan plot of single-nucleotide polymorphisms (SNPs) on autosomal chromosomes under the additive inheritance model. The red line shows the threshold of the genome-wide association screening ( $p < 10^{-4}$ ). (DOCX 119 kb)

**Additional file 4: Fig. S4.** 309 of 1751 (17.65%) unique single-nucleotide polymorphisms (SNPs) with at least one active chromatin state segmentation (states 1–19) in the following 12 cells (with brain-related cells excluded): mesenchymal stem cell-derived adipocyte cultured cells, adipose-derived mesenchymal stem cell cultured cells, HSMM cell-derived skeletal muscle myotubes cell line, muscle satellite cultured cells, bone marrow-derived cultured mesenchymal stem cells, foreskin melanocyte primary cells skin 01, foreskin melanocyte primary cells skin 03, NHDF-Ad adult dermal fibroblast primary cells, breast variant human mammary epithelial cells (vHMEC), HMEC mammary epithelial primary cells, osteoblast primary cells, mesenchymal stem cell-derived chondrocyte cultured cells. (DOCX 554 kb)

**Additional file 5: Fig. S5.** *Cis*-expression qualitative trait loci (eQTLs) associated with the expression of 70 unique genes (*a.k.a.* eGenes). (DOCX 91 kb)

**Additional file 6: Table S1.** Top 88 genetic loci identified from the initial genome-wide association study (GWAS) screening of Taiwanese familial short stature (FSS). (DOCX 19 kb)

**Additional file 7: Table S2.** Single-nucleotide polymorphism (SNP)-based regional annotation. (DOCX 12 kb)

**Additional file 8: Table S3.** Summary of four single-nucleotide polymorphisms (SNPs) located in exonic regions. (DOCX 12 kb)

**Additional file 9: Table S4.** Non-synonymous single-nucleotide polymorphisms (SNPs). (DOCX 14 kb)

**Additional file 10: Table S5.** Tissues with enrichment ( $p < 0.1$ ) of chromHMM-annotated single-nucleotide polymorphisms (SNPs) are shown (Note: brain-related tissues were excluded from the downstream analysis). (DOCX 14 kb)

**Additional file 11: Table S6.** Subnetwork statistics. (DOCX 13 kb)

### Abbreviations

ASN: Asian; BH: Benjamini-Hochberg; BMP: Bone morphogenetic protein; BP: Biological process; EAS: East Asian; eQTL: Expression quantitative trait locus; FDR: False discovery rate; FR: Fruchterman-Reingold; FSS: Familial short stature; GAD: Gene-Disease Associations; GO: Gene ontology; GSEA: Gene set enrichment analysis; GTEX: Genotype-Tissue Expression; GWAS: Genome-wide association study; HPO: Human Phenotype Ontology; HPRD: Human Protein Reference Database; KEGG: Kyoto Encyclopedia of Genes and Genomes; LD: Linkage disequilibrium; MAF: Minor allele frequency; NET: Norepinephrine transporter; ORA: Over-representation analysis; PCA: Principal component analysis; PPI: Protein-protein interaction; QC: Quality control; rhGH: Recombinant human growth hormone; RIN: RNA integrity number; SD: Standard deviation; SS: Short stature; TGF: Transforming growth factor; UTR: Untranslated region

### Acknowledgments

None.

### Authors' contributions

Henry Sung-Ching Wong, Hsing-Fang Lu, and Wei-Chiao Chang conceived the basic concept of this study. Chien-Hsiun Chen, Jer-Yuan Wu, and Henry Sung-Ching Wong performed the computational analysis. Henry Sung-Ching Wong and Wei-Chiao Chang wrote the manuscript. Fuu-Jen Tsai and Wei-Chiao Chang supervised and coordinated this research. All authors discussed and interpreted the results of this manuscript. All authors read the final version of the manuscript.

### Funding

This work was supported by grants from the Ministry of Science and Technology (MOST 105-2628-B-038 - 001-MY4), Taipei Medical University (12310-0223), and the National Health Research Institutes (NHRI-107A1-MGCP-1817202), all of Taiwan. This study was also supported by grants from the China Medical University Hospital (DMR-108-113 and DMR-108-114), and the Ministry of Science and Technology, Taiwan (MOST 108-2314-B-039 -044 -MY3).

### Availability of data and materials

None.

### Ethics approval and consent to participate

This research protocol was approved by the institutional review board (IRB) and the ethics committee of the Human Studies Committee of China Medical University Hospital (CMUH). Our study was conducted in accordance with the *Declaration of Helsinki* principles and meet the requirements of the institute. The samples of familial short stature were recruited by the Children's Hospital, China Medical University, Taichung, Taiwan. We obtained written informed consent from all participants and their parents (or legal guardians).

### Consent for publication

Written informed consent was obtained from the participant for publication of their individual details and accompanying images in this manuscript. The consent form is held by the authors and is available for review by the Editor-in-Chief. We accept the conditions of submission and the BioMed Central Copyright and License Agreement.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Department of Clinical Pharmacy, School of Pharmacy, Taipei Medical University, Taipei, Taiwan. <sup>2</sup>Genetic Center, Department of Medical Research, China Medical University Hospital, Taichung, Taiwan. <sup>3</sup>School of Chinese Medicine, China Medical University, Taichung, Taiwan. <sup>4</sup>Laboratory of Bone and Joint Diseases, RIKEN Center for Integrative Medical Sciences, Tokyo, Japan. <sup>5</sup>Graduate Institute of Integrated Medicine, China Medical University, Taichung, Taiwan. <sup>6</sup>Center for Personalized Medicine, China Medical University Hospital, Taichung, Taiwan. <sup>7</sup>Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. <sup>8</sup>Master Program for Clinical Pharmacogenomics and Pharmacoproteomics, School of Pharmacy, Taipei Medical University, Taipei, Taiwan. <sup>9</sup>Department of Medical Research, Shuang Ho Hospital, Taipei Medical University, New Taipei City, Taiwan. <sup>10</sup>Pain Research Center, Wan Fang Hospital, Taipei Medical University, Taipei, Taiwan. <sup>11</sup>Children's Hospital of China Medical University, Taichung, Taiwan. <sup>12</sup>Department of Biotechnology and Bioinformatics, Asia University, Taichung, Taiwan.

Received: 27 May 2019 Accepted: 9 October 2019

Published online: 07 November 2019

### References

1. Nilsson O, Weise M, Landman EB, Meyers JL, Barnes KM, Baron J. Evidence that estrogen hastens epiphyseal fusion and cessation of longitudinal bone growth by irreversibly depleting the number of resting zone progenitor cells in female rabbits. *Endocrinology*. 2014;155(8):2892–9.
2. Emons JA, Boersma B, Baron J, Wit JM. Catch-up growth: testing the hypothesis of delayed growth plate senescence in humans. *J Pediatr*. 2005; 147(6):843–6.
3. Sederquist B, Fernandez-Vojvodich P, Zaman F, Savendahl L. Recent research on the growth plate: impact of inflammatory cytokines on longitudinal bone growth. *J Mol Endocrinol*. 2014;53(1):T35–44.
4. Lettre G. Recent progress in the study of the genetics of height. *Hum Genet*. 2011;129(5):465–72.
5. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat Genet*. 2014;46(11):1173–86.
6. Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, Fine RS, Lu Y, Schurmann C, Highland HM, et al. Rare and low-frequency coding variants alter human adult height. *Nature*. 2017;542(7640):186–90.
7. Lanktree MB, Guo Y, Murtaza M, Glessner JT, Bailey SD, Onland-Moret NC, Lettre G, Ongen H, Rajagopalan R, Johnson T, et al. Meta-analysis of dense Genecentric association studies reveals common and uncommon variants associated with height. *Am J Hum Genet*. 2011;88(1):6–18.
8. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*. 2010;467(7317):832–8.
9. Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, Yoon D, Lee MH, Kim DJ, Park M, et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet*. 2009;41(5):527–34.
10. Estrada K, Krawczak M, Schreiber S, van Duijn K, Stolk L, van Meurs JB, Liu F, Penninx BW, Smit JH, Vogelzangs N, et al. A genome-wide association study of northwestern Europeans involves the C-type natriuretic peptide signaling pathway in the etiology of human height variation. *Hum Mol Genet*. 2009; 18(18):3516–24.
11. Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S, et al. Many

- sequence variants affecting diversity of adult human height. *Nat Genet.* 2008;40(5):609–15.
12. Kim JJ, Lee HI, Park T, Kim K, Lee JE, Cho NH, Shin C, Cho YS, Lee JY, Han BG, et al. Identification of 15 loci influencing height in a Korean population. *J Hum Genet.* 2010;55(1):27–31.
  13. Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C, et al. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat Genet.* 2008;40(5):584–91.
  14. Okada Y, Kamatani Y, Takahashi A, Matsuda K, Hosono N, Ohmiya H, Daigo Y, Yamamoto K, Kubo M, Nakamura Y, et al. A genome-wide association study in 19 633 Japanese subjects identified LHX3-QSOX2 and IGF1 as adult height loci. *Hum Mol Genet.* 2010;19(11):2303–12.
  15. Liu JZ, Medland SE, Wright MJ, Henders AK, Heath AC, Madden PA, Duncan A, Montgomery GW, Martin NG, McArae AF. Genome-wide association study of height and body mass index in Australian twin families. *Twin research and human genetics : the official journal of the International Society for Twin Studies.* 2010;13(2):179–93.
  16. Sisley S, Trujillo MV, Khoury J, Bäckeljauw P. Low incidence of pathology detection and high cost of screening in the evaluation of asymptomatic short children. *J Pediatr.* 2013;163(4):1045–51.
  17. Lin YJ, Liao WL, Wang CH, Tsai LP, Tang CH, Chen CH, Wu JY, Liang WM, Hsieh AR, Cheng CF, et al. Association of human height-related genetic variants with familial short stature in Han Chinese in Taiwan. *Sci Rep.* 2017;7(1):6372.
  18. Albertsson-Wikland K, Aronson AS, Gustafsson J, Hagenas L, Ivarsson SA, Jonsson B, Kristrom B, Marcus C, Nilsson KO, Ritzen EM, et al. Dose-dependent effect of growth hormone on final height in children with short stature without growth hormone deficiency. *J Clin Endocrinol Metab.* 2008;93(11):4342–50.
  19. Leschek EW, Rose SR, Yanovski JA, Troendle JF, Quigley CA, Chipman JJ, Crowe BJ, Ross JL, Cassorla FG, Blum WF, et al. Effect of growth hormone treatment on adult height in peripubertal children with idiopathic short stature: a randomized, double-blind, placebo-controlled trial. *J Clin Endocrinol Metab.* 2004;89(7):3140–8.
  20. Rekers-Mombarg LT, Massa GG, Wit JM, Matranga AM, Buckler JM, Butenandt O, Chaussain JL, Frisch H, Leiberman E, Yturriaga R, et al. Growth hormone therapy with three dosage regimens in children with idiopathic short stature. European study group participating investigators. *J Pediatr.* 1998;132(3 Pt 1):455–60.
  21. van Gool SA, Kamp GA, Odink RJ, de Muinck Keizer-Schrama SM, Delemarre-van de Waal HA, Oostdijk W, Wit JM. high-dose GH treatment limited to the prepubertal period in young children with idiopathic short stature does not increase adult height. *Eur J Endocrinol.* 2010;162(4):653–60.
  22. Okada Y. From the era of genome analysis to the era of genomic drug discovery: a pioneering example of rheumatoid arthritis. *Clin Genet.* 2014;86(5):432–40.
  23. Sakae S, Okada Y. Human genetics contributes to the understanding of disease pathophysiology and drug discovery. *J Orthop Sci.* 2017;22(6):977–81.
  24. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
  25. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA et al. a global reference for human genetic variation. *Nature* 2015, 526(7571):68–74.
  26. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012;40(Database issue):D930–4.
  27. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, et al. Human protein reference database—2009 update. *Nucleic Acids Res.* 2009;37(Database issue):D767–72.
  28. Taylor DL, Jackson AU, Narisu N, Hemani G, Erdos MR, Chines PS, Swift A, Idol J, Didion JP, Welch RP, et al. Integrative analysis of gene expression, DNA methylation, physiological traits, and genetic variation in human skeletal muscle. *Proc Natl Acad Sci U S A.* 2019;116(22):10883–8.
  29. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al. Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium Nat Genet* 2000, 25(1):25–29.
  30. Becker KG, Barnes KC, Bright TJ, Wang SA. The genetic association database. *Nat Genet.* 2004;36(5):431–2.
  31. Stange K, Desir J, Kakar N, Mueller TD, Budde BS, Gordon CT, Horn D, Seemann P, Borck G. A hypomorphic BMPRI1B mutation causes du Pan acromesomelic dysplasia. *Orphanet J Rare Dis.* 2015;10:84.
  32. Baron J, Savendahl L, De Luca F, Dauber A, Phillip M, Wit JM, Nilsson O. Short and tall stature: a new paradigm emerges. *Nat Rev Endocrinol.* 2015;11(12):735–46.
  33. Pogue R, Lyons K. BMP signaling in the cartilage growth plate. *Curr Top Dev Biol.* 2006;76:1–48.
  34. Dymont DA, Smith AC, Alcantara D, Schwartzenruber JA, Basel-Vanagaite L, Curry CJ, Temple IK, Reardon W, Mansour S, Haq MR, et al. Mutations in PIK3R1 cause SHORT syndrome. *Am J Hum Genet.* 2013;93(1):158–66.
  35. Fournier M, Orpinell M, Grauffel C, Scheer E, Garnier JM, Ye T, Chavart V, Joint M, Esashi F, Dejaegere A, et al. KAT2A/KAT2B-targeted acetylation reveals a role for PLK4 acetylation in preventing centrosome amplification. *Nat Commun.* 2016;7:13227.
  36. Choi HJ, Choi JY, Cho SW, Kang D, Han KO, Kim SW, Kim SY, Chung YS, Shin CS. Genetic polymorphism of geranylgeranyl diphosphate synthase (GGSP1) predicts bone density response to bisphosphonate therapy in Korean women. *Yonsei Med J.* 2010;51(2):231–8.
  37. Peris P, Gonzalez-Roca E, Rodriguez-Garcia SC, Del Mar L-CM, Monegal A, Guanabens N. Incidence of mutations in the ALPL, GGP1, and CYP11A1 genes in patients with atypical femoral fractures. *JBMR Plus.* 2019;3(1):29–36.
  38. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, Kochi Y, Ohmura K, Suzuki A, Yoshida S, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature.* 2014;506(7488):376–81.
  39. Wong HS, Juan YS, Wu MS, Zhang YF, Hsu YW, Chen HH, Liu WM, Chang WC. Integrative bioinformatic analyses of an oncogenomic profile reveal the biology of endometrial cancer and guide drug discovery. *Oncotarget.* 2016;7(5):5909–23.
  40. Petrides F, Shearston K, Chatelais M, Guilbaud F, Meilhac O, Lambert G. The promises of PCSK9 inhibition. *Curr Opin Lipidol.* 2013;24(4):307–12.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

