A tissue- and organ-based cell biological atlas of obesity-related human genes and cellular pathways

Iman Tavassoly[1*], Valentina Barbieri[1*], Coen van Hasselt[1,2] Pedro Martinez[1], Eric A. Sobie[1], Jens Hansen[1], Evren Azeloglu[3], Joseph Goldfarb[1], Philippe Sanseau[4], Deepak K Rajpal[5,6] and Ravi Iyengar[1**]

1. Department of Pharmacological Sciences and Systems Biology Center New York, Icahn School of Medicine at Mount Sinai, New York NY 10029, USA

2. Leiden Academic Centre for Drug Research, LACDR/Pharmacology, Einsteinweg 55, 2333 CC Leiden, The Netherlands

3. Department of Medicine – Division of Nephrology, Icahn School of Medicine at Mount Sinai, New York NY 10029, USA

4. GSK, Medicines Research Center, Gunnels Wood Road, Stevenage SG1 2NY, UK

5. GSK,1250 South Collegeville Road, Collegeville, PA 19426, USA

6. Current Affiliation: Bioinformatics, Translational Sciences, Sanofi, Cambridge, MA 02139, USA

*Equal contributions as first authors.

**Corresponding author:

Dr. Ravi Iyengar, Department of Pharmacological Sciences, Box 1215 Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue, Room 12-70, New York, 10029, NY, USA; ravi.iyengar@mssm.edu

Abstract

Over the last decades, several features of obesity have been identified at behavioral, physiological, endocrine and genomic levels, and they have revealed the complexity of the disease; obesity results from a combination of genetic predisposition, endocrine disorders, and dysregulation of both food intake and energy expenditure. This complexity makes the development of new therapeutic regimens challenging and bariatric surgery is still the treatment of choice for many obese patients. Given the need for noninvasive therapeutic intervention strategies, we sought to systematically study the biological manifestations of obesity in peripheral organs. We analyzed publicly available datasets of genes, genomic determinants, and levels of obesity-related hormones in the blood, using a combination of methodologies, including graph theory and dynamical modeling, that allow for the integration of different types of datasets. The analysis revealed tissue- and organ-specific metabolic impairments and potential new drug targets. All the data are organized into a tissue/organ-based subcellular-function atlas for human obesity. The data show that the complexity of the obesity arises due to the multiplicity of subcellular processes in different peripheral organs.

Introduction

Over the past decade, the prevalence of obesity and severe obesity among adults has risen to almost 40% of the world population [1]. According to the World Health Organization (WHO), obesity is a health issue not only in high income, but also in low- and mid-income countries, and is a risk factor for many chronic diseases including diabetes, cardiovascular diseases and cancer [2]. Obesity has been extensively studied at the behavioral, physiological, endocrine, and, more recently, the genomic level [3-6]. All these different biological levels and the interactions among them play a role in both the etiology and the pathophysiology of obesity and contribute to its complexity. Obesity is physiologically defined as a body mass index (BMI) of over $30kg/m^2$; it manifests itself as a combination of genetic predisposition, dysregulation of food intake and energy expenditure, and as an endocrine disorder making it challenging to develop therapeutic approaches to maintain body weight in a healthy range [3, 4]. After many years of trying, several large pharmaceutical companies have discontinued programs for the development of anti-obesity drugs [7]. Although there have been recent studies of peripheral drug targets for the treatment of obesity (incidental data have shown that GLP-1 agonists can be effective in moderate weight loss [8]), most FDA-approved drugs have primary targets in the Central Nervous System (CNS) [7]. Currently, the most effective treatment for obesity is bariatric surgery [9]. Given the need for new therapeutics that can control obesity without surgery, we hypothesized that we could identify potential novel therapeutic targets by systematically studying biological expressions of obesity in peripheral organs. We decided to utilize publicly available data because an enormous amount of information exists that can be analyzed using multiple computational systems biology approaches allowing for the integration of diverse types of datasets and models.

In principle, it is straightforward to argue that changes in genes and genomic determinants lead to alterations in cell biological pathways, and these, in turn, affect tissue and organ function, and finally whole-body physiology. However, it has been difficult, in practice, to connect the genomic information to the dynamics of physiological functions [10]. A recent review demonstrates the utility of integrated analyses of genomic determinants such as Single Nucleotide Polymorphisms (SNPs) with biological and behavioral determinants in obesity [11]. We sought to determine the relationships between genomic and molecular determinants and the dynamics associated with lean and obese physiological states. Data from patients after bariatric surgery provided a sample of subjects who had been obese but who had reduced their weight and BMI toward normal levels. As a measure of the physiological states, we used orexigenic and anorexigenic hormone levels in the blood and their changes after feeding. We hypothesized a) that genetic predisposition could be linked to endocrine impairment in energy homeostasis; b) that different tissues and organs could respond differently to genetic predisposition and endocrine impairment at the level of gene expression; c) that this variability could produce differences in catabolism and anabolism and confer distinct physiological profiles in lean, obese, and post-bariatric-surgery individuals. Because the effects of the same genetic and genomic determinants on cellular pathways may vary among different tissues and organs, we decided to map the genes and their associated cell biological processes to six peripheral tissues/organs that are known to be involved in nutrient processing and energy storage: stomach, intestine, pancreas, liver, adipose tissue, and skeletal muscle [12]. Upon performing this analysis, we organized the information as an atlas for human obesity, containing data from multiple biological levels.

The basic framework for the construction of the atlas is shown in **Figure 1**. The flow chart used for the analyses is shown in **Supplementary Figure S1,** and the databases used in the study are listed in **Supplementary Table S1**. The major goal of our mapping was to identify key genes and subcellular pathways by integrating the genomic and endocrine information using several computational methods. Genomic and genetic data were obtained from public databases [13-16] and then expanded utilizing the human interactome to take into consideration the role of possible intermediate genes. Because differences in endocrine regulation following food intake have been shown in lean, obese and post-bariatric surgery individuals [12], we determined whether such differences could be linked to genetic predisposition and how they would affect the genomic and genetic data. We used published data on changes in plasma levels of orexigenic and anorexigenic hormones to analyze physiological differences between lean and obese individuals and the effects of bariatric surgery on the plasma levels of hormones related to the regulation of food intake and obesity. We then developed a dynamical model that allowed us to identify those kinetic parameters that regulate the hormone levels in the systemic circulation and are more likely to be different in lean, obese, and post-bariatric surgery subjects. These findings allowed us to rank these kinetic parameters and the genes that are associated to them; the genes were then mapped onto the genomic and genetic network and weighted according to their significance in physiological functions; because we had three different physiological profiles, three endocrine-weighted genetic and genomic networks were obtained. To organize the information in terms of anatomy and biological functions, we filtered the ranked networks according to organ level transcriptomic patterns in humans [17]. We then used a combination of the FDA Adverse Events Database (FAERS) and DrugBank data [18, 19] to identify targets of drugs that have side effects of either weight gain or weight loss and drugs used for weight control whose targets include

5

peripheral tissues and mapped these targets to the appropriate organs. The organ-selective gene lists were then used to identify relevant subcellular metabolic processes, leading to the creation of an atlas of obesity-related genes and subcellular processes in the stomach, pancreas, liver, adipose tissue, intestine, and muscle of the human body. The atlas illustrates how catabolic and anabolic subcellular processes differ at the tissue/organ level in lean, obese, and post-bariatric surgery individuals.

Results

Genetic and genomic datasets

We searched the NCBI Gene database and the EBI GWAS catalog (**Supplementary Table S1**) [13-16] to identify genetic and genomic variability associated with BMI in humans. The NCBI Gene database contains genes that have been associated with specific traits in published studies (2006-2018). The GWAS catalog includes all the SNPs and mapped genes associated with diseases or traits in published studies from 2006 to 2018. The overall strategy for identifying relevant genes is shown in **Supplementary Figure S2**. As of April 2018, we identified 2527 genes associated with either "obesity", "body mass index", "overweight", "free fatty mass", or "leanness" from the NCBI Gene database (**Supplementary Figure S2**). The complete list of query terms and the number of associated genes is shown in **Supplementary Table S2**. The list of genes and associated publications are shown in **Supplementary File 1**. Genomic variability was incorporated as SNPs associated with obesity, leanness, and obesity-related traits with a p-value $\leq 5*10^{-8}$ from the EBI GWAS catalog (**Supplementary Figure S3**). The SNPs were mapped to a total of 324 regions that were utilized for subsequent analysis (**Supplementary Figure S3, Supplementary Table S3, Supplementary File 2**). Common to both sets were 117 genes, yielding 2734 unique genes (**Figure 2A**) that are distributed among

all chromosomes except the sex chromosomes (**Figure 2B)**. To further identify all the genes related to pathways involved in obesity, we built a protein-protein interaction (PPI) network, using proteins encoded by the 2734 unique genes from the two datasets as seed nodes. Using the nearest neighbor expansion method, we then identified proteins in the human interactome that were one intermediate node between seed nodes. This process resulted in a network of 7268 gene products that, at this stage of the study, were considered to have equal statistical significance and constituted the main dataset for subsequent analyses (**Figure 2C and Supplementary Figure S4**). To identify the major biological processes that these genes participate in, we performed an enrichment analysis for biological processes utilizing the Molecular Biology of the Cell Ontology (MBCO) algorithm [20]. This algorithm allows for the identification of interactions between subcellular processes (SCPs) and classifies them into hierarchical levels according to their function. The analysis showed that the genes from the unfiltered PPI network dataset seemed to be significantly related to cellular communication, lipid metabolism, cellular response to stress, extracellular matrix homeostasis, metabolism, and transport of glucose and lipids (**Figure 2D and Supplementary Figures S5, S6, and S7**).

Endocrine dynamics datasets

Impairment in hormonal regulation of hunger and satiation is a pivotal physiological marker of obesity [12]. One reasonable explanation for the observed differences between lean, obese, and post-bariatric surgery states is that kinetic regulation of hormones could be different in these three physiological states. To assess this, we developed a dynamical model that describes the basal secretion, stimulus-evoked secretion, utilization, and degradation of 3 orexigenic and 8 anorexigenic peripheral hormones in human subjects [8, 21-32]. Synthesis and storage of the hormones were not considered explicitly but were subsumed as part of the secretion rate. Our

goal in developing this dynamical model was to identify the kinetic parameters to which hormone levels were most sensitive. These kinetic parameters, in turn, can be associated with cellular pathways and genes related to obesity. Our general strategy was to choose parameter values in the model by fitting model time courses to experimentally measured blood hormone levels, before and after a meal. As we do not know if these are the true biologically relevant parameters, we then conducted extensive parameter variations followed by multivariable regression analyses to identify the most sensitive kinetic parameters associated with the levels of each hormone.

We collated plasma level concentrations at seven time points from published studies [8, 24-32] to include a fasting state (-120min-0min), a feeding point (at 0 min) and a post-prandial state (0-120min). Four hours is the usual timeframe for blood sample collection in published studies, so we ran our simulations accordingly. The model reproduces the biological response to the ingestion of glucose as follows: glucose stimulates ghrelin (GHRL) secretion from the stomach. Absorption by the intestine stimulates the secretion of glucagon-like peptide-1 (GLP-1), cholecystokinin (CCK), oxyntomodulin (OXM), gastric inhibitor peptide (GIP), and peptide YY (PYY). Once absorbed, glucose delivery to the pancreas via the mesenteric circulation stimulates the secretion of insulin (INS), glucagon (GCG), and amylin (AMY), and delivery to adipocytes via the systemic circulation stimulates secretion of leptin (LEP) and adiponectin (ADIPOQ) (**Figure 3A**). According to the literature, in the lean subjects, orexigenic hormones (GHRL, ADIPOQ, and GCG) show higher fasting levels and lower post-prandial levels. The anorexigenic hormones (INS, GLP-1, CCK, OXM, GIP, PYY, and LEP) present lower fasting levels and higher postprandial levels. In obese subjects, hormone responsiveness to nutrients was reported to be totally (GHRL, ADIPOQ, GLP-1, PYY, OXM) or partially impaired (GCG, LEP, INS,

AMY, CCK, GIP). Plasma levels of some hormones in post-bariatric surgery patients lie between the lean-curves and the obese-curves (GHRL, GCG, LEP, AMY, GIP), while others appear higher than in either lean or obese individuals (INS, CCK, GLP-1, PYY, and OXM) (**Figure 3B**). We ran simulated time courses for lean, obese and post-bariatric surgery subjects in a fasting and a postprandial state to estimate reliable initial values for the kinetic parameters such that the simulated curves closely resembled published experimental data [8, 25, 27, 28, 30-33] (**Figure 3B and Supplementary File 3**).

The magnitude of the endocrine response to nutrients can be assessed by the difference in plasma levels of the hormones before and after the meal. To quantify it, we chose to calculate the ratio between the 2-hour area under the curve (AUC) of the postprandial plasma concentration and the 2-hour AUC of the fasting plasma concentration for each hormone (**Figure 3C and Supplementary Figure S8**). To identify which kinetic parameters are more likely to affect the endocrine response represented by the AUC, we performed a multivariable regression analysis (MRA) [34]. The logic of this analysis was to randomly vary the initial values of the set of kinetic parameters used in the model a number of times (**Supplementary Figure S9**), utilizing each new set of values to run a simulation that by varying each parameter affects the output and expresses it in terms of a numerical Partial Least Square Regression (PLSR) coefficient: the higher the coefficient, the more sensitive is the response to the parameter; the lower the coefficient, the more robust is the response to changes in the parameter. The kinetic parameters for lean, obese, and post-surgery populations were then ranked according to the coefficients. The methodology is schematically shown in **Figure 3D and Supplementary Figure S10** and fully explained in online methods.

bioRxiv preprint doi: https://doi.org/10.1101/2020.03.16.993824; this version posted March 18, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We focused on the top ten most sensitive kinetic parameters from simulations of lean, obese, and post-bariatric surgery subjects because they account for 90% of the total sensitivity (**Figure 3E**). This analysis can indicate both the processes that have the greatest effects on hormone levels and how their relative importance can be different between physiological states. For instance, most of the top-ranking parameters involved the degradation and utilization of anorexigenic hormones, except for ghrelin, which ranked highly in all three physiological states. In lean subjects specifically, oxyntomodulin and amylin also seem to play an important role, while insulin ranked lower. In obese subjects, leptin secretion scored the highest, but cholecystokinin and amylin were among the top five. Glucagon-like peptide 1 seems to play a key role as well, providing an unbiased physiological reason for its therapeutic use for obesity [35]. In post-surgery subjects, utilization of hormones secreted by the intestine appeared to play a key role, and it could, in part, be related to the changes in the anatomy of the patients due to the surgery.

Mapping of the physiological responses onto the genetic and genomic determinants

We determined whether there were correlations between the genetic and genomic dataset (**Figure 2C**) and the physiological regulation results (**Figure 3E**). To identify genes that are associated in the literature with each kinetic parameter, we text-mined the PubMed database utilizing the MBCO algorithm [20]. For each parameter, we obtained a list of genes ranked by their association to it based on published data (**Figure 4A**). Each gene was then assigned an average rank defined as the mean of the MBCO rank and the MRA rank for the associated parameter (**Figure 4A**). Because the same gene can have a different average rank in lean, obese, and post-surgery subjects due to different MRA ranking of the parameter in the three populations, we obtained population-specific gene ranks that incorporated their predicted

endocrine sensitivity (**Figure 4B**). The complete list of genes and parameters is in **Supplementary File 4.**

We looked for the genes identified in the endocrine sensitivity analysis in the unfiltered genetic and genomic PPI network **(Figure 4B)** and weighted them using the average rank previously calculated. Because these rank values are population-dependent, we generated lean-, obese-, and post-surgery PPI networks (**Figure 4B**). The genes that are present in the endocrine dataset but not in the genomic and genetic dataset were weighted according to their distance from the ranked genes. This mapping allowed us to classify the genetic and the genomic dataset according to its connection with the endocrine ranked genes and to create population-specific networks: genes that are present in both datasets ranked at the top; genes that are present in the endocrine dataset but not in the genomic and genetic dataset ranked lower.

Because the same gene products can differ in expression, regulation, and function in different tissues which have significant roles in weight control, we investigated which genes in the 3 PPI networks were expressed in peripheral tissues/organs of interest. We obtained the tissue/organ-specific mRNA expression data for each gene from the GTEX portal for adipose tissue, intestine, liver, skeletal muscle, pancreas, and stomach (**Supplementary Figure S11**). We divided the genes of the three subnetworks according to their expression in the 6 tissues/organs. The analysis generated 18 PPI subnetworks, one for lean, obese, and post-surgical subjects in each of the 6 tissues.

While most weight-loss drugs that have been marketed have primary targets in the CNS, it may be advantageous to seek peripherally active therapeutic approaches. We determined whether any FDA-approved drugs that have a high risk of weight gain or weight loss as side effects

have as secondary targets, gene products in any of the subnetworks (**Supplementary File 5**). We selected drugs that have exclusively peripheral targets and drugs that have both peripheral and Central Nervous System (CNS) targets. Drugs with exclusively CNS targets, as lorcaserin and phentermine, were excluded. The respective targets for each drug were obtained from Drugbank [19] and DGIdb [36]. The top 10 drugs associated with weight gain and weight loss, and their targets are presented in **Figure 6**. We mapped the targets onto the tissue-specific subnetworks (**Supplementary Figure S12 and S13**). 8 weight gain-associated targets and 11 weight loss-associated targets are distributed among the 6 tissues/organs of interest so that each contains at least one target **(Figure 5)**. The implications of this result are two-fold: first, the presence of known secondary targets in the subnetworks generated by our analysis indicates that one reason behind the complex pathophysiology of obesity may be that determinants can be identified in several levels. Second, because the CNS-associated side effects of anti-obesity drugs are undesirable, it suggests the possibility that we can identify new peripheral targets to treat obesity.

Catabolic and anabolic differences in lean, obese and post-surgery subjects

To evaluate differences in catabolism and anabolism in our PPI models of lean, obese and post-surgery subjects, we identified, utilizing the MBCO algorithm, the subcellular processes that the gene products in the 18 subnetworks participate in [20] (**Supplementary Figure S14**). We classified these metabolic processes as catabolic and anabolic processes, according to standard textbook criteria [37]. Ratios of catabolic and anabolic processes within the top 100 subcellular processes within each of the six tissues and organs of interest and each physiological condition were calculated.

Interestingly, lean and post-surgery subjects show a higher number of subcellular catabolic processes, while obese subjects have similar numbers of each (**Figure 7**). When the subcellular processes are analyzed at the organ level, results vary considerably, and the apparent similarity of lean and post-surgery subjects is lost. In lean subjects, catabolic processes seem to be favored in pancreas, intestine, and muscle, but with ratios close to 1 in liver, stomach and adipose tissue. In post-surgery subjects, catabolic processes predominate in liver, stomach, and muscle, ratios are close to 1 in the intestine and adipose tissue, and anabolic processes predominate in the pancreas. These differing profiles merit further study. In obese subjects, catabolic processes predominate only in the liver, anabolic in stomach, pancreas, and adipose tissue, and ratios approach 1 in muscle and intestine. The large change in the catabolic/anabolic ratio in the stomach in post-surgery versus obese or lean subjects could reflect changes in the gastrointestinal anatomy due to surgical removal of a portion of the stomach. Our results suggest that anabolic and catabolic processes in the pancreas, intestine, and muscle could be considered as potential peripheral targets for future combination therapeutic treatments, since in obese subjects the catabolic/anabolic ratio in these organs is less than 1 meaning the anabolism is dominant while in the lean subjects this ratio is higher than 1 in these organs meaning that catabolism is dominant. These organs can be targeted to change the ratio and balance between the catabolism and anabolism in obese subjects. (**Figure 7**)

Discussion

This atlas, to the best of our knowledge, is among the first to contain most obesity-related genetic, genomic, transcriptomic and physiological data and to show their interactions focusing on non-CNS targets. Our aim in developing the atlas is to obtain new knowledge on the pathophysiology of obesity and to help catalyze new drug-based therapeutic approaches for

13

treating obesity. Most published atlases have focused on molecular data such as gene expression profiles [38, 39] or relationships between gene expression and pathology [40, 41]. Typically, the physiology of obesity has been approached by analyzing one or a few hormones levels in the blood at a time [8, 21-29], not adequately taking into consideration the complexity of the disease. A key finding from our atlas is that multiple tissues/organs are likely to have significant involvement in obesity, and it may not be possible to rank them and the potential drug targets within them in a simple manner.

Our bottom-up integrative analysis offers multiple insights of the biological organization combining genetic and genomic determinants with different types of physiological dynamics. Several features of our method illustrate how it can be transferable to the study of other systemic diseases. First, it utilizes publicly available knowledge and allows scaling the information by systematic statistical analysis. Second, it allows the integration of datasets: static and dynamic, unbiased, and specified. Third, it permits prediction for tissue/organ-selective new combinatorial drug targets that can be further experimentally tested. We took data from several biological levels, from molecular whole genomic data to cell-level transcriptomic data, to organ-level endocrine data, in order to exhaustively characterize 3 different phenotypic states: lean, obese, and post-bariatric surgery. The integration achieved by connecting dynamical models to network models enables us to identify subcellular processes within the organs of interest. It should be noted that PPI networks allow for the identification of intermediate genes or gene products that could be potentially overlooked [42]. The computational model of the hormone levels and the regression analysis allow for the study of a system from a dynamic perspective and for ranking cell-level information. The use of transcriptomic expression profiles for various organs (**Figure 5**) provides a facile approach to parse the physiological state-selective molecular

14

networks into tissue/organ-based physiological state-selective networks. Enrichment analysis of subcellular biological processes permits us to identify the interactions between the different datasets from a functional perspective. This ability to resolve the molecular networks provides insight into the different cellular pathways related to obesity that are operational in different tissues and organs. While the role of the central nervous system as a target for control of body weight [43] has been well studied in terms of hunger circuits [44], similar information regarding other tissues and organs have not been available until now. Our analysis unravels the complexity underlying obesity using a systems biology approach: genomic determinants widely distributed on all the non-sex chromosomes (**Figure 2**) and cellular pathways that are operational in many tissues and organs (**Figure 7**, **Supplementary Figure S14**). From the perspective of physiological states, our findings provide a clear representation of the divergence of the plasma level of several metabolic hormones and of the catabolic and anabolic subcellular processes among lean, obese and post-bariatric surgery patients in six tissues and organs. In this context, skeletal muscle, pancreas, intestine and adipose tissue emerged as the most perturbed (compared to the lean state) by the obese state. Stomach, liver, pancreas and skeletal muscle seem to be the most perturbed (compared to the lean state) in the post-bariatric surgery state. Thus, it appears that the post-bariatric state is different from the lean state in terms of organ function, although they are phenotypically similar. We believe this represents the first identification of the perturbation of pathways implicated in energy production and expenditure in obesity by data integration. Additionally, from the atlas, we can deduce by unbiased analysis that in lean people catabolic processes predominate while in obese people anabolic processes predominate. Although this conclusion is intuitively obvious, have it established by unbiased information and analysis of multiple datasets, our findings provide a clear representation of the

divergence of the level of several energy-regulating hormones and of the divergence in the catabolic and anabolic subcellular processes among lean, obese and post-bariatric surgery patients. According to our analysis, these alterations stem primarily from perturbed transport and mobilization of cholesterol and LDL, glycolysis, and glycogen synthesis (**Supplementary Figure S14**). Despite this relatively simple overall conclusion, the distribution of these anabolic and catabolic processes among multiple organs highlights the underlying complexity in going from physiological knowledge to drug development. Adipose tissue, intestine and, rather unexpectedly, stomach seem to be the most promising organs for new potential drug targets (Table 2), although we cannot fully discount whether the use of data from bariatric surgery influences our ranking of the stomach. According to our results, PPARG (peroxisome proliferator-activated receptor gamma) seems to be significant in adipose tissue and intestine, and PPARA (peroxisome proliferator-activated receptor alpha) in muscle, intestine, and liver. This is consistent with their roles in glucose metabolism and fatty acid storage. ADBR2 (the beta-2-adrenergic receptor) which is known to participate in glycogenesis and gluconeogenesis in the liver and skeletal muscle, seems to play a role in adipose tissue as well. In the intestine, the presence of IL7R (interleukin7 receptor) and PTGDR2 (prostaglandin DP2 receptor) could be an indication of inflammatory states, which have been shown to be correlated with obesity [3]. ALDH3A1 (aldehyde dehydrogenase 3 family member A1) is involved in lipid peroxidation and could represent an interesting target for further investigation. While the presence of several genes of the cytochrome family is to be expected in the liver, our results suggest that APOF (apolipoprotein F) involved in the transport of cholesterol, could be considered for future therapeutic approaches.

Overall, our analysis indicates that drug treatment of obesity aimed at non-CNS targets is most likely to be successful if it is pleiotropic in nature, affecting cellular regulation at multiple sites.

Online Methods

Human genetic dataset

The human genetic dataset is composed of genes selected from the NCBI Gene Database by text mining. Using "*trait* + [orgn] homo sapiens" as a query term we retrieved genes that were associated with 12 terms related to obesity according to published literature as of May 2018: "obesity", "obese", "bariatric surgery", "Body Mass Index OR BMI", "fat", "fat mass", "fat free mass OR FFM", "lean OR leanness", "overweight". A cut-off of at least 2 publications per gene was set. We retrieved 2734 overall unique genes. The trait-specific and complete list of genes can be consulted in the **Supplementary File 1**.

Human genomic dataset

To create this dataset we collected genomic data available in the Genome Wide Association Studies Catalog [3, 45] from the May 2017 version. The GWAS data are currently mapped to Genome Assembly GRCh38.p7 and dbSNP Build 147. From the list of selectable diseases and traits in the catalog we selected 11 terms of interest for our study with their respective list of SNPs and associated genes: adiposity, body fat percentage, body mass (lean), body mass index, fat body mass, obesity, visceral adipose tissue/subcutaneous adipose tissue ratio, waist circumference, waist-hip ratio, weight, weight loss (gastric bypass surgery). The respective numbers of SNPs, intergenic and intragenic regions are listed in **Supplementary Table S3**. The results were filtered to exclude studies that didn't provide replication; only adult subject cohort

studies were considered; finally, only SNPs that were associated with traits with p-value $\leq 5*10^{-8}$, the standard genomic statistical significance threshold, were selected. A total of 751 SNP-trait associations, 220 mapped genes, and 30 intergenic regions were found. To obtain a list of unique SNPs, if a SNP was reported in more than one study, only the most significant SNP-trait association score (smallest p-value) was considered. The GWAS catalog already provides genome mapping for the listed genes. Nevertheless, the mapping was further manually curated utilizing Genome Browser [15] to assess any new potential mapping though ENCODE [14]. The final curated list of SNP-trait associations for each trait of interest is stored in **Supplementary File 2**.

Protein-Protein interaction network analysis

We utilized X2K software [46] to create a protein-protein interaction (PPI) network. X2K provides a large-scale protein-protein interaction network made from 18 sources that connect seed lists of genes/proteins. The user can map a list of seed genes to a human interactome. We utilized the 2631 genes in the genetic and genomic datasets as seeds to build a PPI network. A path length of 2 was chosen meaning that the path between two seed nodes requires two edges to connect the nodes, so there is only one intermediate protein between them, i.e., a protein that interacts with both seed nodes. BIND [47], HPRD [48], MINT [49], Biocarta [50], InnateDB [51], MIPS [52], SNAVI [53], BioGRID [54], INTact [55], DIP [56], PDZbase [57], and KEGG [58] databases were selected to extrapolate the intermediate genes from the human interactome. 4637 intermediates were identified.

Multicompartmental dynamical model

We built multicompartmental dynamical models using coupled ordinary differential equations to describe the endocrinological response to glucose ingestion in a lean, obese, and post-bariatric surgery human subject by modeling the plasma levels of eleven hormones associated with the regulation of satiation and appetite. We built a model for each hormone and used the same glucose stimulus for each. Our model differs from the previously published ones [59, 60] because (a) it describes endogenous instead of exogenous plasma levels of the hormones; (b) it provides the kinetic behavior of eleven hormones; (c) it takes into account multiple compartments and organs; (d) it describes three different physiological states; (e) the aim of the model is to identify a set of kinetic parameters that best describe the temporal progression of the experimentally observed values. These parameters would then allow us to determine statistically significant differences between lean, obese and post-bariatric surgery endocrinology. These parameters can then be varied in an unbiased manner to identify the fragile and robust parameters of the system as whole.

Collation of the experimental data of hormone plasma levels

Data points for each individual time course were obtained from published data [8, 22-32]. Inclusion criteria for the published data were: a) The selected hormone must be associated to the regulation of appetite level and satiation level in human subjects; b) The plasma level of the hormones must be endogenous; c) Time points must cover at least one hour of a fasting state and two hours of a postprandial state; d) Subjects should not be affected by any comorbidity; d) Post-surgery subjects must have undergone Roux Y Gastric Bypass and should preferably be one year post-surgery.

Data points for lean, obese and post-RYGB surgery were collected for 3 orexigenic hormones (ghrelin, adiponectin, and glucagon) and eight anorexigenic hormones (leptin, insulin, amylin, cholecystokinin, peptide YY, glucagon-like peptide 1, gastric inhibitory peptide-1, oxyntomodulin) (**Supplementary Table 4**). Each point in every time course is the mean ± S.E.M. of time points from different published data (A complete list of the time points is provided in **Supplementary File 3**). All the different units of concentration of hormones were converted to pmol/L utilizing the molecular weight of each hormone.

Compartmental dynamic ordinary differential equation (ODE) model

Our model describes the kinetics of eleven different hormones secreted by the different organs in the digestive system in response to meal ingestion. Glucose digestion, absorption, production, utilization and degradation and each hormone's secretion, first pass through the liver, utilization and degradation were considered. The model comprises seven explicit compartments (stomach, intestine, pancreas, adipocytes, plasma, muscle and liver) and one black-box compartment (degradation). A scheme of the interactions in the model is shown in **Figure3A**. Given the complexity of the system, and the availability of only plasma glucose and hormone concentrations to find the parameters, the parameter estimation methodologies were used to build a reliable model [61-64]. The model parameters were numerically identified by nonlinear least-squares as implemented in MATLAB 2018a.

Glucose Kinetics.

Endogenous Glucose Production (EGP):

The functional description of EGP in terms of glucose and insulin signals comprises a direct glucose signal and both delayed and anticipated insulin signals

$$EGP = k_{p1} - k_{p2}[Glu]_{plasma} - k_{p3}[INS]_{delay2} - k_{p4}[INS]_{portal}$$

$$EGP(0) = EGP_{basal}$$

where $[Glu]_{plasma}$ is the concentration of the plasma glucose, $[INS]_{portal}$ is the concentration of insulin in the portal vein, $[INS]_{delay2}$ $[INS]_{delay}$ is a delayed insulin signal realized with a chain of two compartments:

$$\frac{d[INS]_{delay1}}{dt} = k_i([INS]_{plasma}(t) - [INS]_{delay1}(t))$$

$$\frac{d[INS]_{delay2}}{dt} = k_i([INS]_{delay1}(t) - [INS]_{delay2}(t))$$

k_{p1} (mg/kg/min) is the basal rate of EGP at zero glucose and insulin, k_{p2} (min^{-1}) is the parameter for liver glucose effectiveness in reducing plasma glucose, k_{p3} (mg/kg/min/(mol/l)) is the parameter governing amplitude of insulin action on the liver, k_{p4} (mg/kg/min/(mol/kg)) is the parameter for amplitude of portal insulin action on the liver and k_i is the parameter accounting for delay between the insulin signal and insulin action. EGP is also constrained to be non-negative.

Glucose Ingestion and Absorption: Glucose transits through the stomach and intestine by assuming the stomach to be represented by two compartments (one for solid, $[Glu]_{stomach1}$, and one for triturated phase, $[Glu]_{stomach2}$), while a single compartment is used to describe the intestine. Glucose concentration in the stomach as a function of time is given by:

21

$$\frac{d[Glu]_{stomach1}}{dt} = -k_{stomach} * [Glu]_{stomach1}$$

$$\frac{d[Glu]_{stomach2}}{dt} = k_{stomach1} * [Glu]_{stomach1} - k_{gut} * [Glu]_{stomach2}$$

Glucose concentration in the intestine is given by:

$$\frac{d[Glu]_{intestine}}{dt} = k_{gut} * [Glu]_{stomach2} - k_{abs} * [Glu]_{intestine}$$

Glucose Utilization:

Hormone-independent utilization represents glucose uptake by the brain and erythrocytes $F_{cns}(t)$. Peripheral tissue hormone-dependent utilization depends non-linearly on glucose in the tissues and it is described by a Michaelis–Menten kinetics:

$$U_{Glu\_tissue} = \frac{V_m\big(X(t)\big) * [Glu]_{tissue}(t)}{K_m\big(X(t)\big) + [Glu]_{tissue}(t)}$$

$$V_m\big(X(t)\big) = V_{m0} + V_{mx} * X(t)$$

$$K_m\big(X(t)\big) = K_{m0} + K_{mx} * X(t)$$

As we have 11 hormones:

$$X = \sum_{1}^{11} [hormone]_{Tissue}$$

Total glucose utilization U_{Glu}:

$$U_{Glu}(t) = U_{Glu_{tissue}}(t) + F_{cns}(t)$$

Hormone Kinetics

Glucose ingestion triggers the secretion of orexigenic and anorexigenic hormones from different peripheral organs and tissue.

Hormone Secretion:

22

Hormone secretion into plasma by specific organs occurs if glucose concentration in that organ exceeds a certain threshold. Hormone-specific secretions are described below.

Stomach Secretion of Ghrelin (GHRL):

If $\beta_{GHRL}\left(\frac{[Glu]_{stomach}}{V_{Glu}} - h_{GHRL}\right) \geq S_{GHRL_{Basal}}$ then ghrelin secretion is triggered:

$$\frac{d[GHRL]_{free}}{dt} = -\alpha_{Ghrl} * ([GHRL]_{free}(t) + \alpha_{GHRL} * \beta_{GHRL}(\frac{[Glu]_{stomach}}{V_{Glu}} - h_{GHRL})])$$

otherwise

$$\frac{d[GHRL]_{free}}{dt} = -\alpha_{GHRL} * [GHRL]_{free}(t) - \alpha_{GHRL} * S_{GHRL\_Basal}$$

Where α_{GHRL} (min^{-1}) is the delay between the glucose signal and ghrelin secretion, β_{GHRL} (pmol/kg/min per mg/dL is the stomach "X/A cells" responsivity to glucose, and h_{GHRL} (mg/dL) is the threshold level of glucose above which the "X/A cells" initiate to produce ghrelin (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Intestine Secretion of Glucagon-like Peptide 1 (GLP-1):

If $\beta_{GLP1}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{GLP1}) \geq S_{GLP1\_Basal}$ then GLP-1 secretion is triggered:

$$\frac{d[GLP1]_{free}}{dt} = -\alpha_{GLP1} * ([GLP1]_{free}(t) + \alpha_{GLP1} * \beta_{GLP1}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{GLP1})])$$

otherwise

$$\frac{d[GLP1]_{free}}{dt} = -\alpha_{GLP1} * [GLP1]_{free}(t) - \alpha_{GLP1} * S_{GLP1\_Basal}$$

23

Where α_{GLP1} (min^{-1}) is the delay between glucose signal and GLP-1 secretion, β_{GLP1} (pmol/kg/min per mg/dL) is the lower intestine L cells responsivity to glucose, and h_{GLP1} (mg/dL) is the threshold level of glucose above which the L cells begin to produce GLP1 (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Intestine Secretion of Peptide YY (PYY):

If $\beta_{PYY}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{PYY}) \geq S_{PYY\_Basal}$ then PYY secretion is triggered:

$$\frac{d[PYY]_{free}}{dt} = -\alpha_{PYY} * ([PYY]_{free}(t) + \alpha_{PYY} * \beta_{PYY}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{PYY})])$$

otherwise

$$\frac{d[PYY]_{free}}{dt} = -\alpha_{PYY} * [PYY]_{free}(t) - \alpha_{PYY} * S_{PYY\_Basal}$$

Where α_{PYY} (min^{-1}) is the delay between glucose signal and PYY secretion, β_{PYY} (pmol/kg/min per mg/dL) is the lower intestine L cells responsivity to glucose, and h_{PYY} (mg/dL) is the threshold level of glucose above which the L cells begin to produce PYY (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Intestine Secretion of Gastric Inhibitor Peptide (GIP):

If $\beta_{GIP}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{GIP}) \geq S_{GIP\_Basal}$ then GIP secretion is triggered:

$$\frac{d[GIP]_{free}}{dt} = -\alpha_{GIP} * ([GIP]_{free}(t) + \alpha_{GIP} * \beta_{GIP}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{GIP})])$$

otherwise

$$\frac{d[GIP]_{free}}{dt} = -\alpha_{GIP} * [GIP]_{free}(t) - \alpha_{GIP} * S_{GIP\_Basal}$$

24

Where α_{GIP} (min^{-1}) is the delay between glucose signal and GIP secretion, β_{GIP} (pmol/kg/min per mg/dL) is the lower intestine L cells responsivity to glucose, and h_{GIP} (mg/dL) is the threshold level of glucose above which the L cells begin to produce GIP (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Intestine Secretion of Oxyntomodulin (OXM):

If $\beta_{OXM}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{OXM}) \geq S_{OXM\_Basal}$ then oxyntomodulin secretion is triggered:

$$\frac{d[OXM]_{free}}{dt} = -\alpha_{OXM} * ([OXM]_{free}(t) + \alpha_{OXM} * \beta_{OXM}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{OXM})])$$

otherwise

$$\frac{d[OXM]_{free}}{dt} = -\alpha_{OXM} * [OXM]_{free}(t) - \alpha_{OXM} * S_{OXM\_Basal}$$

Where α_{OXM} (min^{-1}) is the delay between glucose signal and OXM secretion, β_{OXM} (pmol/kg/min per mg/dL) is the lower intestine L cells responsivity to glucose, and h_{OXM} (mg/dL) is the threshold level of glucose above which the L cells begin to produce OXM (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Intestine Secretion of Cholecystokinin (CKK):

If $\beta_{CKK}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{CKK}) \geq S_{CKK\_Basal}$ then cholecystokinin secretion is triggered:

$$\frac{d[CKK]_{free}}{dt} = -\alpha_{CKK} * ([CKK]_{free}(t) + \alpha_{CKK} * \beta_{CKK}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{CKK})])$$

otherwise

$$\frac{d[CKK]_{free}}{dt} = -\alpha_{CKK} * [CKK]_{free}(t) - \alpha_{CKK} * S_{CKK\_Basal}$$

25

Where α_{CKK} (min^{-1}) is the delay between glucose signal and CKK secretion, β_{CKK} (pmol/kg/min per mg/dL) is the lower intestine I cells responsivity to glucose, and h_{CKK} (mg/dL) is the threshold level of glucose above which the I cells begin to produce CKK (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Pancreatic secretion of insulin (INS).

If $\beta_{INS}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{INS}) \geq S_{INS\_Basal}$ then insulin secretion is triggered:

$$\frac{d[INS]_{free}}{dt} = -\alpha_{INS} * ([INS]_{free}(t) + \alpha_{INS} * \beta_{INS}(\frac{[Glu]_{intestine}}{V_{Glu}} - h_{INS})])$$

otherwise

$$\frac{d[INS]_{free}}{dt} = -\alpha_{INS} * [INS]_{free}(t) - \alpha_{INS} * S_{INS\_Basal}$$

Where α_{INS} (min^{-1}) is the delay between glucose signal and insulin secretion, β_{INS} (pmol/kg/min per mg/dL) is the pancreatic β cells responsivity to glucose, and h_{INS} (mg/dL) is the threshold level of glucose above which the β cells begin to produce insulin (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Pancreatic Secretion of Amylin (AMY):

If $\beta_{AMY}(\frac{[Glu]_{pancreas}}{V_{Glu}} - h_{AMY}) \geq S_{AMY\_Basal}$ then amylin secretion is triggered:

$$\frac{d[AMY]_{free}}{dt} = -\alpha_{AMY} * ([AMY]_{free}(t) + \alpha_{AMY} * \beta_{AMY}(\frac{[Glu]_{pancreas}}{V_{Glu}} - h_{AMY})])$$

otherwise

$$\frac{d[AMY]_{free}}{dt} = -\alpha_{AMY} * [AMY]_{free}(t) - \alpha_{AMY} * S_{AMY\_Basal}$$

Where α_{AMY} (min^{-1}) is the delay between glucose signal and amylin secretion, β_{AMY} (pmol/kg/min per mg/dL) is the pancreatic β cells responsivity to glucose, and h_{AMY} (mg/dL) is the threshold level of glucose above which the β cells begin to produce amylin (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Pancreatic Secretion of Glucagon (GCG):

If

$$\beta_{GCG}(\frac{[Glu]_{pancreas}}{V_{Glu}} - h_{GCG}) \geq S_{GCG\_Basal}$$

Glucagon secretion is triggered:

$$\frac{d[GCG]_{free}}{dt} = -\alpha_{GCG} * ([GCG]_{free}(t) + \alpha_{GCG} * \beta_{GCG}(\frac{[Glu]_{pancreas}}{V_{Glu}} - h_{GCG})])$$

otherwise

$$\frac{d[GCG]_{free}}{dt} = -\alpha_{GCG} * [GCG]_{free}(t) - \alpha_{GCG} * S_{GCG\_Basal}$$

Where α_{GCG} (min^{-1}) is the delay between glucose signal and glucagon secretion, β_{GCG} (pmol/kg/min per mg/dL) is the pancreatic α cells responsivity to glucose, and h_{GCG} (mg/dL) is the threshold level of glucose above which the α cells begin to produce glucagon (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Adipose Tissue Secretion of Leptin (LEP):

If $\beta_{LEP}(\frac{[Glu]_{adipocytes}}{V_{Glu}} - h_{LEP}) \geq S_{LEP\_Basal}$ then leptin secretion is triggered:

27

$$\frac{d[LEP]_{free}}{dt} = -\alpha_{LEP} * ([LEP]_{free}(t) + \alpha_{LEP} * \beta_{LEP}(\frac{[Glu]_{adipocytes}}{V_{Glu}} - h_{LEP})])$$

otherwise

$$\frac{d[LEP]_{free}}{dt} = -\alpha_{LEP} * [LEP]_{free}(t) - \alpha_{LEP} * S_{LEP\_Basal}$$

Where α_{LEP} (min^{-1}) is the delay between glucose signal and leptin secretion, β_{LEP} (pmol/kg/min per mg/dL) is the adipocytes responsivity to glucose, and h_{LEP} (mg/dL) is the threshold level of glucose above which the adipocytes begin to produce leptin (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Adipose Tissue Secretion of Adiponectin (ADIPOQ):

If $\beta_{ADIPOQ}(\frac{[Glu]_{adipocytes}}{V_{Glu}} - h_{ADIPOQ}) \geq S_{ADIPOQ\_Basal}$ then adiponectin secretion is triggered:

$$\frac{d[ADIPOQ]_{free}}{dt} = -\alpha_{ADIPOQ} * ([ADIPOQ]_{free}(t) + \alpha_{ADIPOQ} * \beta_{ADIPOQ}(\frac{[Glu]_{adipocyte}}{V_{Glu}} - h_{ADIPOQ})])$$

otherwise

$$\frac{d[ADIPOQ]_{free}}{dt} = -\alpha_{ADIPOQ} * [ADIPOQ]_{free}(t) - \alpha_{ADIPOQ} * S_{ADIPOQ\_Basal}$$

Where α_{ADIPOQ} (min^{-1}) is the delay between glucose signal and adiponectin secretion, β_{ADIPOQ} (pmol/kg/min per mg/dL) is the adipocytes responsivity to glucose, and h_{ADIPOQ} (mg/dL) is the threshold level of glucose above which the adipocytes initiate to produce adiponectin (has been set to the basal glucose concentration to guarantee system steady state in basal condition).

Degradation and Plasma Level of the Hormones:

Degradation occurs both in the liver (first pass) and in the periphery and it follows the kinetics listed below.

28

Ghrelin Degradation and Plasma Level:

$$\frac{d[GHRL]_{liver}}{dt} = -(m_{1GHRL} + m_{3GHRL}(t)) * [GHRL]_{liver}(t) + m_{2GHRL} * [GHRL]_{plasma}(t) * S_{GHRL_{Basal}}$$

Which determines the temporal changes of the plasma level of GHRL:

$$\frac{d[GHRL]_{plasma}}{dt} = -(m_{2GHRL} + m_{4GHRL}) * [GHRL]_{plasma}(t) + m_{1GHRL} * [GHRL]_{liver}(t)$$

m_{1GHRL}, m_{2GHRL}, m_{3GHRL} (min^{-1}) are rate parameters. m_{4GHRL} is peripheral degradation. Hepatic

extraction (HE) that determines liver degradation of ghrelin is described by:

$$HE_{GHRL}(t) = -m_{5GHRL} * S_{GHRL\_Basal} + m_{6GHRL}$$

$$m_{3GHRL}(t) = \frac{HE_{GHRL} * m_{1GHRL}}{1 - HE_{GHRL}(t)}$$

At basal steady state:

$$m_{6GHRL} = m_{5GHRL} * S_{GHRL_{Basal}} * HE_{GHRL_{Basal}}$$

$$m_{3GHRL}(0) = \frac{HE_{GHRL_{Basal}} * m_{1GHRL}}{1 - HE_{GHRL_{Basal}}}$$

$$S_{GHRL_{Basal}} = m_{3GHRL}(0) * [GHRL]_{Liver_{basal}} + m_{4GHRL} * [GHRL]_{Plasma_{basal}} = D_{GHRL_{Basal}}$$

where $D_{GHRL_{Basal}}$ is the basal total degradation.

$$m_{2GHRL} = \left(\frac{S_{GHRL_{Basal}}}{[GHRL]_{Plasma_{basal}}} - \frac{m_{4GHRL}}{1 - HE_{GHRL_{Basal}}}\right) * \left(\frac{1 - HE_{GHRL_{Basal}}}{HE_{GHRL_{Basal}}}\right)$$

$$m_{4_{GHRL}} = 0.4 * \left(\frac{S_{GHRL_{Basal}}}{[GHRL]_{Plasma_{basal}}} \right) * (1 - HE_{GHRL_{Basal}})$$

GLP-1 Degradation and Plasma Level:

$$\frac{d[GLP1]_{liver}}{dt} = -(m_{1GLP1} + m_{3GLP1}(t)) * [GLP1]_{liver}(t) + m_{2GLP1} * [GLP1]_{plasma}(t) * S_{GLP1_{Basal}}$$

Which determines the temporal change in the plasma level of GLP1:

$$\frac{d[GLP1]_{plasma}}{dt} = -(m_{2GLP1} + m_{4GLP1}) * [GLP1]_{plasma}(t) + m_{1GLP1} * [GLP1]_{liver}(t)$$

m_{1GLP1}, m_{2GLP1}, m_{3+GLP1} (min^{-1}) are rate parameters. m_{4GLP1} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of ghrelin is described by:

$$HE_{GLP1}(t) = -m_{5GLP1} * S_{GLP1\_Basal} + m_{6GLP1}$$

$$m_{3_{GLP1}}(t) = \frac{HE_{GLP1} * m_{1\_GLP1}}{1 - HE_{GLP1}(t)}$$

At basal steady state:

$$m_{6GLP1} = m_{5GLP1} * S_{GLP1_{Basal}} * HE_{GLP1_{Basal}}$$

$$m_{3_{GHRL}}(0) = \frac{HE_{GLP1_{Basal}} * m_{1GLP1}}{1 - HE_{GLP1_{Basal}}}$$

$$S_{GLP1_{Basal}=}m_{3GLP1}(0) * [GIP]_{Liver_{basal}} + m_{4GLP1} * [GLP1]_{Plasma_{basal}} = D_{GLP1_{Basal}}$$

where D $_{GLP1basal}$ is the basal total degradation.

$$m_{2_{GLP1}} = \left(\frac{S_{GLP1_{Basal}}}{[GLP1]_{Plasma_{basal}}} - \frac{m_{4_{GLP1}}}{1 - HE_{GLP1_{Basal}}} \right) * \left(\frac{1 - HE_{GLP1_{Basal}}}{HE_{GLP1_{Basal}}} \right)$$

$$m_{4_{GLP1}} = 0.4 * \left(\frac{S_{GLP1_{Basal}}}{[GLP1]_{Plasma_{basal}}} \right) * (1 - HE_{GLP1_{Basal}})$$

PYY Degradation and Plasma Level:

$$\frac{d[PYY]_{liver}}{dt} = -(m_{1PYY} + m_{3PYY}(t)) * [PYY]_{liver}(t) + m_{2PYY} * [PYY]_{plasma}(t) * S_{PYY_{Basal}}$$

Which determines the temporal changes in the plasma level of PYY:

$$\frac{d[PYY]_{plasma}}{dt} = -(m_{2PYY} + m_{4PYY}) * [PYY]_{plasma}(t) + m_{1PYY} * [PYY]_{liver}(t)$$

m_{1PYY}, m_{2PYY} and m_{4PYY} (min^{-1}) are rate parameters. m_{4PYY} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of PYY is described by:

$$HE_{PYY}(t) = -m_{5PYY} * S_{PYY_{Basal}} + m_{6PYY}$$

$$m_{3PYY}(t) = \frac{HE_{PYY} * m_{1PYY}}{1 - HE_{PYY}(t)}$$

At basal steady state:

$$m_{6PYY} = m_{5PYY} * S_{PYY_{Basal}} * HE_{PYY_{Basal}}$$

$$m_{3PYY}(0) = \frac{HE_{PYY_{Basal}} * m_{1PYY}}{1 - HE_{PYY_{Basal}}}$$

$$S_{PYY_{Basal}} = m_{3PYY}(0) * [PYY]_{Liver_{basal}} + m_{4PYY} * [PYY]_{Plasma_{basal}} = D_{PYY_{Basal}}$$

where $D_{PYY_{Basal}}$ is the basal total degradation.

$$m_{2_{PYY}} = \left(\frac{S_{PYY_{Basal}}}{[PYY]_{Plasma_{basal}}} - \frac{m_{4PYY}}{1 - HE_{PYY_{Basal}}} \right) * \frac{1 - HE_{PYY_{Basal}}}{HE_{PYY_{Basal}}}$$

$$m_{4_{PYY}} = 0.4 * \left(\frac{S_{PYY_{Basal}}}{[PYY]_{Plasma_{basal}}} \right) * 1 - HE_{PYY_{Basal}}$$

GIP Degradation and Plasma Level:

$$\frac{d[GIP]_{liver}}{dt} = -(m_{1GIP} + m_{3GIP}(t)) * [GIP]_{liver}(t) + m_{2GIP} * [GIP]_{plasma}(t) * S_{GIP_{Basal}}$$

Which determines the temporal changes in the plasma level of GIP:

$$\frac{d[GIP]_{plasma}}{dt} = -(m_{2GIP} + m_{4GIP}) * [GIP]_{plasma}(t) + m_{1GIP} * [GIP]_{liver}(t)$$

m_{1GIP}, m_{2GIP}, m_{4GIP} (min^{-1}) are rate parameters. m_{4GIP} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of GIP is described by:

$$HE_{GIP}(t) = -m_{5GIP} * S_{GIP\_Basal} + m_{6GIP}$$

$$m_{3GIP}(t) = \frac{HE_{GIP} * m_{1GIP}}{1 - HE_{GIP}(t)}$$

At basal steady state:

$$m_{6GIP} = m_{5GIP} * S_{GIP_{Basal}} * HE_{GIP_{Basal}}$$

$$m_{3GIP}(0) = \frac{HE_{GIP_{Basal}} * m_{1GIP}}{1 - HE_{GIP_{Basal}}}$$

$$S_{GIP_{Basal}} = m_{3GIP}(0) * [GIP]_{Liver_{basal}} + m_{4GIP} * [GIP]_{Plasma_{basal}} = D_{GIP_{Basal}}$$

32

where $D_{GIP_{Basal}}$ is the basal total degradation.

$$m_{2_{GIP}} = \left(\frac{S_{GIP_{Basal}}}{[GIP]_{Plasma_{basal}}} - \frac{m_{4GIP}}{1 - HE_{GIP_{Basal}}} \right) * \frac{1 - HE_{GIP_{Basal}}}{HE_{GIP_{Basal}}}$$

$$m_{4GIP} = 0.4 * \left(\frac{S_{GIP_{Basal}}}{[GIP]_{Plasma_{basal}}} \right) * 1 - HE_{GIP_{Basal}}$$

OXM Degradation and Plasma Level:

$$\frac{d[OXM]_{liver}}{dt} = -(m_{1OXM} + m_{3OXM}(t)) * [OXM]_{liver}(t) + m_{2OXM} * [OXM]_{plasma}(t) * S_{OXM_{Basal}}$$

Which determines the temporal changes in the plasma level of OXM:

$$\frac{d[OXM]_{plasma}}{dt} = -(m_{2OXM} + m_{4OXM}) * [OXM]_{plasma}(t) + m_{1OXM} * [OXM]_{liver}(t)$$

m_{1OXM}, m_{2OXM}, m_{3OXM} (min^{-1}) are rate parameters. m_{3OXM} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of OXM is described by:

$$HE_{OXM}(t) = -m_{5OXM} * S_{OXM_{Basal}} + m_{6OXM}$$

$$m_{3OXM}(t) = \frac{HE_{OXM} * m_{1OXM}}{1 - HE_{OXM}(t)}$$

At basal steady state:

$$m_{6OXM} = m_{5OXM} * S_{OXM_{Basal}} * HE_{OXM_{Basal}}$$

$$m_{3OXM}(0) = \frac{HE_{OXM_{Basal}} * m_{1OXM}}{1 - HE_{OXM_{Basal}}}$$

$$S_{OXM_{Basal}} = m_{3OXM}(0) * [OXM]_{Liver_{basal}} + m_{4OXM} * [OXM]_{Plasma_{basal}} = D_{OXM_{Basal}}$$

where $D_{OXM_{Basal}}$ is the basal total degradation.

33

$$m_{2_{OXM}} = \left(\frac{S_{OXM_{Basal}}}{[OXM]_{Plasma_{basal}}} - \frac{m_{4OXM}}{1 - HE_{OXM_{Basal}}} \right) * \frac{1 - HE_{OXM_{Basal}}}{HE_{OXM_{Basal}}}$$

$$m_{4OXM} = 0.4 * \left(\frac{S_{OXM_{Basal}}}{[OXM]_{Plasma_{basal}}} \right) * 1 - HE_{OXM_{Basal}}$$

CCK Degradation and Plasma Level:

$$\frac{d[CCK]_{liver}}{dt} = -(m_{1CCK} + m_{3CCK}(t)) * [CCK]_{liver}(t) + m_{2CCK} * [CCK]_{plasma}(t) * S_{CCK_{Basal}}$$

Which determines the temporal changes of the plasma level of CCK:

$$\frac{d[CCK]_{plasma}}{dt} = -(m_{2CCK} + m_{4CCK}) * [CCK]_{plasma}(t) + m_{1CCK} * [CCK]_{liver}(t)$$

m_{1CCK}, m_{2CCK}, m_{3CCK} (min^{-1}) are rate parameters. m_{4CCK} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of CCK is described by:

$$HE_{CCK}(t) = -m_{5CCK} * S_{CCK_{Basal}} + m_{6CCK}$$

$$m_{3CCK}(t) = \frac{HE_{CCK} * m_{1CCK}}{1 - HE_{CCK}(t)}$$

At basal steady state:

$$m_{6CCK} = m_{5CCK} * S_{CCK_{Basal}} * HE_{CCK_{Basal}}$$

$$m_{3CCK}(0) = \frac{HE_{CCK_{Basal}} * m_{1CCK}}{1 - HE_{CCK_{Basal}}}$$

$$S_{CCK_{Basal}} = m_{3CCK}(0) * [CCK]_{Liver_{basal}} + m_{4CCK} * [CCK]_{Plasma_{basal}} = D_{CCK_{Basal}}$$

where $D_{CCK_{Basal}}$ is the basal total degradation.

$$m_{2CCK} = \left(\frac{S_{CKK_{Basal}}}{[CCK]_{Plasma_{basal}}} - \frac{m_{4CCK}}{1 - HE_{CKK_{Basal}}}\right) * \left(\frac{1 - HE_{CKK_{Basal}}}{HE_{CKK_{Basal}}}\right)$$

$$m_{4CCK} = 0.4 * \left(\frac{S_{CCK_{Basal}}}{[CCK]_{Plasma_{basal}}}\right) * (1 - HE_{CCK_{Basal}})$$

Insulin Degradation and Plasma Level:

$$\frac{d[INS]_{liver}}{dt} = -(m_{1INS} + m_{3INS}(t)) * [INS]_{liver}(t) + m_{2INS} * [INS]_{plasma}(t) * S_{INS_{Basal}}$$

Which determines a plasma level of:

$$\frac{d[INS]_{plasma}}{dt} = -(m_{2INS} + m_{4INS}) * [INS]_{plasma}(t) + m_{1INS} * [INS]_{liver}(t)$$

m_{1INS}, $m_{2INS \; and} \; m_{3INS}$ (min^{-1}) are rate parameters. m_{4INS} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of insulin is described by:

$$HE_{INS}(t) = -m_{5INS} * S_{INS_{Basal}} + m_{6INS}$$

$$m_{3INS}(t) = \frac{HE_{INS} * m_{1INS}}{1 - HE_{INS}(t)}$$

At basal steady state:

$$m_{6INS} = m_{5INS} * S_{INS_{Basal}} * HE_{INS_{Basal}}$$

35

$$m_{3INS}(0) = \frac{HE_{INS_{Basal}} * m_{1INS}}{1 - HE_{INS_{Basal}}}$$

$$S_{INS_{Basal}} = m_{3INS}(0) * [INS]_{Liver_{basal}} + m_{4INS} * [INS]_{Plasma_{basal}} = D_{INS_{Basal}}$$

where $D_{INS_{Basal}}$ is the basal total degradation.

$$m_{2INS} = \left(\frac{S_{INS_{Basal}}}{[INS]_{Plasma_{basal}}} - \frac{m_{4INS}}{1 - HE_{INS_{Basal}}} \right) * \left(\frac{1 - HE_{INS_{Basal}}}{HE_{INS_{Basal}}} \right)$$

$$m_{4INS} = 0.4 * \left(\frac{S_{INS_{Basal}}}{[INS]_{Plasma_{basal}}} \right) * (1 - HE_{INS_{Basal}})$$

Amylin Degradation and Plasma Level:

$$\frac{d[AMY]_{liver}}{dt} = -(m_{1AMY} + m_{3AMY}(t)) * [AMY]_{liver}(t) + m_{2AMY} * [AMY]_{plasma}(t) * S_{AMY_{Basal}}$$

Which determines the temporal changes of the plasma level of amylin:

$$\frac{d[AMY]_{plasma}}{dt} = -(m_{2AMY} + m_{4AMY}) * [AMY]_{plasma}(t) + m_{1AMY} * [AMY]_{liver}(t)$$

m_{1AMY}, m_{2AMY} and m_{3AMY} (min^{-1}) are rate parameters. m_{4AMY} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of amylin is described by:

$$HE_{AMY}(t) = -m_{5AMY} * S_{AMY_{Basal}} + m_{6AMY}$$

$$m_{3AMY}(t) = \frac{HE_{AMY} * m_{1AMY}}{1 - HE_{AMY}(t)}$$

At basal steady state:

$$m_{6AMY} = m_{5AMY} * S_{AMY_{Basal}} * HE_{AMY_{Basal}}$$

$$m_{3AMY}(0) = \frac{HE_{AMY_{Basal}} * m_{1AMY}}{1 - HE_{AMY_{Basal}}}$$

$$S_{AMY_{Basal}} = m_{3AMY}(0) * [AMY]_{Liver_{basal}} + m_{4AMY} * [AMY]_{Plasma_{basal}} = D_{AMY_{Basal}}$$

where $D_{AMY_{Basal}}$ is the basal total degradation.

$$m_{2AMY} = \left(\frac{S_{AMY_{Basal}}}{[AMY]_{Plasma_{basal}}} - \frac{m_{4AMY}}{1 - HE_{AMY_{Basal}}} \right) * \left(\frac{1 - HE_{AMY_{Basal}}}{HE_{AMY_{Basal}}} \right)$$

$$m_{4AMY} = 0.4 * \left(\frac{S_{AMY_{Basal}}}{[AMY]_{Plasma_{basal}}} \right) * (1 - HE_{AMY_{Basal}})$$

Glucagon Degradation and Plasma Level:

$$\frac{d[GCG]_{liver}}{dt} = -(m_{1GCG} + m_{3GCG}(t)) * [GCG]_{liver}(t) + m_{2GCG} * [GCG]_{plasma}(t) * S_{GCG_{Basal}}$$

Which determines the temporal changes of the plasma level of glucagon:

$$\frac{d[GCG]_{plasma}}{dt} = -(m_{2GCG} + m_{4GCG}) * [GCG]_{plasma}(t) + m_{1GCG} * [GCG]_{liver}(t)$$

m_{1GCG}, m_{2GCG} and m_{3GCG}(min^{-1}) are rate parameters. m_{4GCG} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of glucagon is described by:

$$HE_{GCG}(t) = -m_{5GCG} * S_{GCG\_Basal} + m_{6GCG}$$

$$m_{3GCG}(t) = \frac{HE_{GCG} * m_{1GCG}}{1 - HE_{GCG}(t)}$$

37

At basal steady state:

$$m_{6GCG} = m_{5GCG} * S_{GCG_{Basal}} * HE_{GCG_{Basal}}$$

$$m_{3GCG}(0) = \frac{HE_{GCG_{Basal}} * m_{1GCG}}{1 - HE_{GCG_{Basal}}}$$

$$S_{GCG_{Basal}} = m_{3GCG}(0) * [GCG]_{Liver_{basal}} + m_{4GCG} * [GCG]_{Plasma_{basal}} = D_{GCG_{Basal}}$$

where $D_{GCG_{Basal}}$ is the basal total degradation.

$$m_{2GCG} = \left(\frac{S_{GCG_{Basal}}}{[GCG]_{Plasma_{basal}}} - \frac{m_{4GCG}}{1 - HE_{GCG_{Basal}}} \right) * \left(\frac{1 - HE_{GCG_{Basal}}}{HE_{GCG_{Basal}}} \right)$$

$$m_{4GCG} = 0.4 * \left(\frac{S_{GCG_{Basal}}}{[GCG]_{Plasma_{basal}}} \right) * (1 - HE_{GCG_{Basal}})$$

Leptin Degradation and Plasma Level:

$$\frac{d[LEP]_{liver}}{dt} = -(m_{1LEP} + m_{3LEP}(t)) * [LEP]_{liver}(t) + m_{2LEP} * [LEP]_{plasma}(t) * S_{LEP_{Basal}}$$

Which determines the temporal changes of the plasma level of leptin:

$$\frac{d[LEP]_{plasma}}{dt} = -(m_{2LEP} + m_{4LEP}) * [LEP]_{plasma}(t) + m_{1LEP} * [LEP]_{liver}(t)$$

m_{1LEP}, m_{2LEP} and m_{3LEP} (min^{-1}) are rate parameters. m_{4LEP} is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of leptin is described by:

$$HE_{LEP}(t) = -m_{5LEP} * S_{LEP_{Basal}} + m_{6LEP}$$

$$m_{3LEP}(t) = \frac{HE_{LEP} * m_{1LEP}}{1 - HE_{LEP}(t)}$$

At basal steady state:

$$m_{6LEP} = m_{5LEP} * S_{LEP_{Basal}} * HE_{LEP_{Basal}}$$

$$m_{3LEP}(0) = \frac{HE_{LEP_{Basal}} * m_{1LEP}}{1 - HE_{LEP_{Basal}}}$$

$$S_{LEP_{Basal}} = m_{3LEP}(0) * [LEP]_{Liver_{basal}} + m_{4LEP} * [LEP]_{Plasma_{basal}} = D_{LEP_{Basal}}$$

where $D_{LEP_{Basal}}$ is the basal total degradation.

$$m_{2LEP} = \left(\frac{S_{LEP_{Basal}}}{[LEP]_{Plasma_{basal}}} - \frac{m_{4LEP}}{1 - HE_{LEP_{Basal}}} \right) * \left(\frac{1 - HE_{LEP_{Basal}}}{HE_{LEP_{Basal}}} \right)$$

$$m_{4LEP} = 0.4 * \left(\frac{S_{LEP_{Basal}}}{[LEP]_{Plasma_{basal}}} \right) * (1 - HE_{LEP_{Basal}})$$

Adiponectin Degradation and Plasma Level:

$$\frac{d[ADIPOQ]_{liver}}{dt} =$$

$$-(m_{1ADIPOQ} + m_{3ADIPOQ}(t)) * [ADIPOQ]_{liver}(t) + m_{2ADIPOQ} * [ADIPOQ]_{plasma}(t) * S_{1ADIPOQ_{Basal}}$$

Which determines the temporal changes of the plasma level of adiponectin:

$$\frac{d[ADIPOQ]_{plasma}}{dt} = -(m_{2ADIPOQ} + m_{4ADIPOQ}) * [ADIPOQ]_{plasma}(t) + m_{1ADIPOQ} * [ADIPOQ]_{liver}(t)$$

$m_{1ADIPOQ}$, $m_{2ADIPOQ}$ and $m_{3ADIPOQ}$ (min^{-1}) are rate parameters. $m_{4ADIPOQ}$ is peripheral degradation. Hepatic extraction (HE) that determines liver degradation of adiponectin is described by:

$$HE_{ADIPOQ}(t) = -m_{5ADIPOQ} * S_{ADIPOQ_{Basal}} + m_{6ADIPOQ}$$

$$m_{3ADIPOQ}(t) = \frac{HE_{ADIPOQ} * m_{1ADIPOQ}}{1 - HE_{ADIPOQ}(t)}$$

At basal steady state:

$$m_{6ADIPOQ} = m_{5ADIPOQ} * S_{ADIPOQ_{Basal}} * HE_{1ADIPOQ_{Basal}}$$

$$m_{3ADIPOQ}(0) = \frac{HE_{ADIPOQ_{Basal}} * m_{1ADIPOQ}}{1 - HE_{1ADIPOQ_{Basal}}}$$

$$S_{ADIPOQ_{Basal}} = m_{3ADIPOQ}(0) * [ADIPOQ]_{Liver_{basal}} + m_{4ADIPOQ} * [ADIPOQ]_{Plasma_{basal}} = D_{ADIPOQ_{Basal}}$$

where $D_{ADIPOQ_{Basal}}$ is the basal total degradation.

$$m_{2ADIPOQ} = \left(\frac{S_{ADIPOQ_{Basal}}}{[ADIPOQ]_{Plasma_{basal}}} - \frac{m_{4ADIPOQ}}{1 - HE_{ADIPOQ_{Basal}}} \right) * \frac{1 - HE_{ADIPOQ_{Basal}}}{HE_{ADIPOQ_{Basal}}}$$

$$m_{4ADIPOQ} = 0.4 * \left(\frac{S_{ADIPOQ_{Basal}}}{[ADIPOQ]_{Plasma_{basal}}} \right) * 1 - HE_{ADIPOQ_{Basal}}$$

Hormone Concentration in Muscle and Interstitial Fluid:

$$\frac{d[GHRL]_{tissue}}{dt} = -p_{GHRL} * [GHRL]_{tissue} + p_{GHRL} * ([GHRL]_{tissue}(t) - [GHRL]_{basal})$$

$$\frac{d[GLP1]_{tissue}}{dt} = -p_{GLP1} * [GLP1]_{tissue} + p_{GLP1} * ([GLP1]_{tissue}(t) - [GLP1]_{basal})$$

$$\frac{d[PYY]_{tissue}}{dt} = -p_{PYY} * [PYY]_{tissue} + p_{PYY} * ([PYY]_{tissue}(t) - [PYY]_{basal})$$

$$\frac{d[GIP]_{tissue}}{dt} = -p_{GIP} * [GIP]_{tissue} + p_{GIP} * ([GIP]_{tissue}(t) - [GIP]_{basal})$$

$$\frac{d[OXM]_{tissue}}{dt} = -p_{OXM} * [OXM]_{tissue} + p_{OXM} * ([OXM]_{tissue}(t) - [OXM]_{basal})$$

$$\frac{d[CCK]_{tissue}}{dt} = -p_{CCK} * [CCK]_{tissue} + p_{CCK} * ([CCK]_{tissue}(t) - [CCK]_{basal})$$

$$\frac{d[INS]_{tissue}}{dt} = -p_{INS} * [INS]_{tissue} + p_{INS} * ([INS]_{tissue}(t) - [INS]_{basal})$$

$$\frac{d[AMY]_{tissue}}{dt} = -p_{AMY} * [AMY]_{tissue} + p_{AMY} * ([AMY]_{tissue}(t) - [AMY]_{basal})$$

$$\frac{d[GCG]_{tissue}}{dt} = -p_{GCG} * [GCG]_{tissue} + p_{GCG} * ([GCG]_{tissue}(t) - [GCG]_{basal})$$

$$\frac{d[LEP]_{tissue}}{dt} = -p_{LEP} * [LEP]_{tissue} + p_{LEP} * ([LEP]_{tissue}(t) - [LEP]_{basal})$$

$$\frac{d[ADIPOQ]_{tissue}}{dt} = -p_{ADIPOQ} * [ADIPOQ]_{tissue} + p_{ADIPOQ} * ([ADIPOQ]_{tissue}(t)$$

$$-[ADIPOQ]_{basal})$$

$p_{GHRL}, p_{GLP1}1, p_{PYY}, p_{GIP}, p_{OXM}, p_{CCK}, p_{INS}, p_{AMY}, p_{GCG}, p_{LEP}, p_{ADIPOQ}$ are the rate constants of

the action on peripheral glucose utilization by ghrelin, GLP-1, PYY, GIP, oxyntomodulin,

cholecystokinin, insulin, amylin, glucagon, leptin, adiponectin, respectively.

Quantitative output of the model

The ratio of the Area Under the Curve (AUC) between postprandial state and fasting state was

calculated for each hormone in the lean, obese and post-surgery state using MATLAB2018a

trapezoidal function. (Fig 3B). These quantities utilized as outputs in the following parameter sensitivity analysis.

Parameter sensitivity analysis (PSA) utilizing multivariable regression (MRA)

In order to assess the sensitivity of the kinetic parameters that regulate endocrinology in lean, obese and post-surgery subjects, we performed multivariable regression analysis [18].

The logic of the analysis is to randomly vary a set of inputs n times (trials), utilize the new set of inputs to generate an output according to the function in study and then perform a regression analysis to assess how much the variance in the parameter affects the output.

In our case, the kinetic parameters of the dynamical model were utilized as inputs. We varied the 141 parameters utilizing random scale factors from a log-normal distribution with a median value of 1. The distribution parameter σ, specifying the standard deviation of the distribution of log-transformed variables, controlled the extent to which parameters varied. We generated a total of 300 parameter sets that were stored in the input matrix, I(*300x141*). Each new set of parameters was utilized to simulate the dynamical model and the output (AUC ratios for each hormone) were stored in the output matrix, O (*300x11*).

Values in input and output matrices were mean-centered and normalized by standard deviations for each column. Since random-scale factors obeyed a log-normal distribution, most values in the input matrices were log-transformed before computing the standard deviation.

The PLS regression, performed on matrices I and O using the NIPALS algorithm [65, 66], produced a matrix of regression coefficients, B_{PLS}(*141x11*). For each set of inputs, this matrix can be used to predict the resulting outputs. The regression determines B_{PLS} such that $O_{predicted}$ = $I x B_{PLS}$ is close to the original output matrix O.

To test the susceptibility of our results to the number of trials, we performed the analysis for an increasing number of trials (500, 1000, 3000, 10000). PLS regression was performed using routines and methods from [67]. All computations were performed in MATLAB 2018a.

Ranking of the kinetic parameters from PSA

The kinetic parameters were ranked according to their respective PLS coefficients [0-1]. The top ten most sensitive kinetic parameters for lean, obese and post-surgery subjects are shown in **Figure 4**.

Association of kinetic parameters to their respective genes (MBCO)

To scale the previous results to a molecular level, the Molecular Biology of the Cell Ontology (MCBO) algorithm was used. Each parameter was used as a query in PubMed to retrieve the genes associated with it. To be able to identify these terms in our PubMed article sets, a dictionary that associates biological entities with key terms that might be used by authors when writing about the biological entity was generated. The MCBO algorithm screens the PubMed titles and abstracts of each SCP-specific article set and counts how many articles mentioned a certain gene at least once [20]. A minimum number of abstracts within an SCP specific abstract that is needed to mention a gene (level-1 SCPs: at least 4 abstracts, level-2 SCPs at least 3 abstracts, level-3 SCPs at least 2 abstracts, level-4 SCPs at least 1 abstract) is defined. Genes that were not mentioned in the specified number of abstracts were removed. These article counts for each gene-SCP association were subjected to statistical enrichment analysis to calculate the selectivity of each gene-SCP association, i.e. how selective is that particular gene for that particular SCP. Fisher's exact test was used to calculate two different p-values, one with the genes of all SCPs of the same level (same level p-values), and one with the genes of all SCPs

that have the same parent and therefore belong to the same children set (same children set p-values) as a background set.

Average ranking of the genes utilizing PLS and MCBO score

The MCBO score assigned to each identified gene and the sensitivity rank of each kinetic parameter were utilized to calculate the average ranking for each identified gene.

$$AvgRank_{Gene_i} = \frac{(MCBO_{score\,Gene_i} + PLS_{rank\,Parameter_j})}{2}$$

The MCBO score describes the significance of the association of a gene *i* to a specific parameter *j*. The PLS rank assesses the degree of sensitivity of that parameter and it is used to assess sensitivity variability between lean, obese and post-surgery subjects. The average ranking allows one to rank each gene according to its relevance in both the endocrinological control system and cellular processes.

Weighting the unfiltered PPI network by nearest neighbor distance to an average ranked gene

The unfiltered PPI network created in the first step of the study was weighted according to the distance of each of its genes to the average ranked genes in lean, obese and post-surgery subjects. A nearest-neighbor algorithm was utilized to calculate the distances. The distance values (*0, 1,...n*) were utilized to weight the edges of the PPI network. This generated 3 physiologically specific PPI networks for lean, obese and post-surgery subjects, respectively.

Peripheral tissue specific mRNA expression levels

44

The tissue expression of mRNAs of the genes in the lean, obese, and post-surgery subnetworks was retrieved from the GTEx Portal [19]. The GTEx_Analysis_v6p_RNA-seq_RNA-SeQCv1.1.8_gene_median_rpkm database (May 2017) was downloaded. The 75th percentile was used as a cut-off for high expression. Data for subcutaneous adipose tissue, visceral adipose tissue, stomach, pancreas, liver, intestine, and muscle were selected. The expression values used for adipose tissue were the means for subcutaneous and visceral tissues. Heatmaps of the expression levels for each physiologically- specific subnetwork are represented in **Figure 4B**.

Tissue-specific PPI subnetworks for lean, obese and post-surgery subjects.

Tissue-specific genes were utilized as seed genes to generate tissue-specific subnetworks for lean, obese and post-surgery subjects. We utilized the X2K software [5] to create a protein-protein interaction (PPI) network. A path length of 2 was used, the other settings of the software are screenshot in (**Supplementary Figure S12 and S13**).

Validation of tissue selective networks using drug targets from the FAERS database

We identified FDA-approved drugs that were reported in the FAERS database (June 2017) for inducing weight gain and weight loss [20]. A p-value $<10^{-5}$ and number of reports>*1000* were used as a cut-off. The top 100 drugs causing weight gain and the top 100 drugs causing weight loss were so identified. Their respective targets were identified combining multiple databases: Drugbank [19], The Druggable Genome [68, 69], and Therapeutic Targets Database [70]. Targets that are exclusively expressed in the central nervous system were excluded. Targets which are expressed in adipose tissue, intestine, liver, muscle, pancreas, stomach but drugs acting on them do not cross the brain blood barrier were selected for the validation step.

The number of tissue specific targets that were identifiable in the tissue-specific PPI subnetworks was used as validation of the reliability of the previous steps of the study.

Catabolic and anabolic subcellular processes

From the previous analysis, we calculated the number of significant catabolic and anabolic processes and the ratio of significant catabolic and anabolic processes in lean, obese, and post-surgery subjects (**Figure 6A**). The tissue specific ratio of significant catabolic and anabolic processes in lean, obese, and post-surgery is shown in (**Figure 6B**).

Acknowledgments

Author contributions

RI, DKR, V.B. and CVH. conceived and planned the study. IT and VB designed the computational framework and analyzed the genomic, endocrine and network data. CVH performed the FAERS data analysis. RI, PS and DKR supervised the project. EAS supervised the dynamical modeling studies. I.T, VB, EAS, JG, DKR and RI contributed to the interpretation of the results. IT, VB, JG and RI took the lead in writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Figure Legends:

Figure 1. Schematic of the multi-model approach for mapping of genes and subcellular processes to different tissues and organs in the peripheral tissue atlas of metabolism.

The atlas is generated through multiple types of data that are assembled in steps. **Step 1**: Integration of unweighted data sets. The genetic dataset (NCBI gene database) contains genes that are associated with leanness, obesity, overweight, body mass index, and fat mass. The genomic dataset (EBI GWAS) contains SNPs associated with obesity or with obesity-related traits with a *p*-value $\leq 5*10^{-8}$ according to published data. **Step 2**: Creation of unfiltered PPI network: The two datasets were statistically integrated by generating a genetic & genomic interaction network of the encoded proteins. **Step 3**: Generation of endocrinology data from modeling. The endocrinology dataset contains the genes associated with kinetic parameters for hormones that control satiety and hunger and that ranked differently in statistical significance in lean, obese, and post-bariatric surgery human subjects. These parameters were determined from an ODE-based multicompartmental dynamical model of the plasma concentrations of three orexigenic and eight anorexigenic hormones in lean, obese, and post-bariatric surgery human subjects. **Step 4**: Application of the physiology filter. The statistical significance ranking of genes in step 3 was used to weight each protein in the genetic & genomic network of step 2. Since the same gene can rank differently in lean, obese and post-surgery subjects, three subnetworks, one each for the lean, obese, and post-surgery phenotypes, were obtained. **Step 5**: Application of the tissue expression filter. The genes encoding each protein in each subnetwork were localized according to transcript abundance (GTEx database) in 6 organs/tissues: liver, pancreas, intestine, stomach, adipose tissue, and muscle. **Step 6:** Identification of known drug targets in the organ/tissue-specific subnetworks: A dataset extracted from FAERS database of

48

molecular targets of drugs affecting body weight was used to support the analysis from the previous steps. **Step 7**: Application of the subcellular processes filter. We conducted enrichment analysis using Molecular Biology of the Cell Ontology (MBCO) to identify the subcellular processes that are significantly associated with a lean, obese, and post-surgery phenotype in each organ/tissue. **Step 8**: Generation of the peripheral obesity atlas. Integration of these data enable the application of statistical models to generate graphical representations of differences in genes associated with different organs/tissues in the lean, obese, and post-surgical condition, which could be displayed on a peripheral obesity atlas.

Figure 2. Survey of genes and genomic variants associated with human obesity and weight - related traits.

A) Flowchart of the information process. We identified 2527 genes associated with "obesity" (BMI ≥ 30kg/m^2), "body mass index," "overweight," "free fatty mass," or "leanness" (BMI ≤ 25 kg/m^2), in at least two publications from the NCBI Gene database (https://www.ncbi.nlm.nih.gov/gene). Genomic variability was incorporated as SNPs associated with obesity or obesity-related traits with a *p*-value ≤ 5*10^{-8} from the EBI GWAS catalog (http://www.ebi.ac.uk/gwas/). Mapping the SNPs to obesity and obesity-related traits identified 324 genes, 207 of which were not in the original NCBI set. The two datasets were integrated through construction of a PPI, using the nearest neighbor expansion method with the proteins encoded by the genes as seeds and identifying intermediate proteins (and their encoding genes) from the human interactome up to two path lengths of distance. This process resulted in a network of 7268 proteins (and encoding genes); seed genes that lacked connection to the network were not used for further analysis.

B) Mapping the PPI network onto chromosomes. The outermost layer contains the gene symbols (blue text) of genes associated with obesity and/or obesity-related traits according to the NCBI Gene database. The next layer shows the chromosome bands (colored rectangles) of 22 human chromosomes. Chromosome 23 and the Y chromosome do not contain genes of interest and were omitted. The next layer lists the SNPs associated with obesity (Blue). The final innermost layer lists the SNPs associated with obesity-related traits. SNPs were assigned according to the GWAS catalog.

C) **Unfiltered PPI network.** Representation of the unfiltered PPI network: Seed nodes from the GWAS and NCBI datasets are in yellow, intermediate genes are in light blue.

D) Enrichment analysis of the unfiltered PPI network for subcellular processes. The MBCO algorithm was used to identify which subcellular processes the gene products in the unfiltered network belong to. The algorithm screens the PubMed titles and abstracts of each SCP-specific article set and counts articles mentioning a given gene at least once. The number of abstracts needed to associate a gene with a process varies with the level of the SCP. A minimum of 4 abstracts is required for a level-1 SCP, 3 abstracts for a level-2 SCP, 2 for a level-3 SCP and 1 for a level-4 SCP. Fisher's exact test is used to calculate p-values.

Figure 3. Dynamical modeling of temporal patterns of plasma concentrations of hormones associated with food intake in lean, obese and post-bariatric surgery patients.

A) Compartmental model of the plasma levels of selected peripheral hormones in human subjects. Glucose-responsive peripheral hormones that control food intake are modeled with the kinetic parameters of secretion, utilization, and degradation. Ingestion of glucose stimulates ghrelin (GHRL) secretion from the stomach. Absorption by the intestine stimulates the secretion of glucagon-like peptide-1 (GLP-1), cholecystokinin (CCK), oxyntomodulin (OXM), gastric inhibitor peptide (GIP), and peptide YY (PYY). Through mesenteric blood circulation, glucose delivery to the pancreas stimulates the secretion of insulin (INS), glucagon (GCG), and amylin (AMY) and delivery to adipocytes stimulates secretion of leptin (LEP) and adiponectin (ADIPOQ). All the hormones, except for leptin and adiponectin are subjected to first pass through the liver, and then they are delivered by the circulation to the muscles for utilization, or they undergo degradation. (No data was available of measuring ADIPOQ in the post-bariatric surgery subjects)

B) Simulated time courses of the plasma concentrations of the hormones. Simulated time courses for lean (black with gray circles), obese (red) and post-bariatric surgery (blue) subjects are shown. Simulations were conducted using parameters chosen so that the curves closely resembled experimental data (in dotted lines) from different studies (**Supplementary File 3**). Subjects are in a fasting state from -120 to +0 min. At 0 minutes, administration of glucose initiates the postprandial state, which is monitored for 120 min.

C) Quantitative output of the simulation of the dynamical model. Ratio of the area under the curve (AUC) of the postprandial plasma hormone concentration (2 hours) divided by the AUC of the fasting plasma concentration (2 hours for each hormone was chosen as the quantitative output of the simulation for use in the multivariable regression analysis (MRA).

D) Multivariable regression analysis (MRA) of simulated data identifies kinetic parameters whose variation produces the greatest changes in glucose-evoked hormone release in lean, obese, and post-bariatric surgery subjects. To account for the variation in the kinetic regulation of hormone physiology in a population, and to compensate for the lack of extensive clinical data, we generated 3 simulated sets of human subjects (lean, obese, post-surgery) by randomly varying each of the 141 kinetic parameters of the model, with the mean centered on the parameter value and the standard deviation set to $\pm$ 1. We used partial least square regression (PLSR) to identify which kinetic parameters are more likely to induce a significant perturbation of the output of the dynamical model (AUC ratio). MRA was repeated with progressively larger populations (representing 300, 500, 1,000, 3,000, 10,000 individuals). The MRA for the obese population of 10,000 is represented here: We generated a total of 10,000 parameter sets that were stored in the input matrix, $I(10,000x141)$. Each new set of parameters was then utilized to simulate the dynamical model, and the output (AUC ratios for each hormone) was stored in the output matrix, O $(10,000x11)$. The PLSR, performed on matrices I and O, produced a matrix of regression coefficients, B_{PLS} $(141x11)$, that can be used to predict the resulting outputs such that $O_{predicted} = I \times B_{PLS}$ is close to the original output matrix O.

E) MRA ranking in each subject population. The top 10 most sensitive kinetic parameters from simulations of lean (grey), obese (red), and post-bariatric surgery (blue) subjects, according

53

to PLSR coefficients determined through multivariable regression analysis of the ratio between the area under the curve (AUC) of the postprandial plasma hormone concentration and the AUC of the fasting plasma hormone concentration. The resulting PLSR coefficients were then used to rank the parameters.

Figure 4. Integration of the PPI network and the dynamical model data to obtain physiologically weighted tissue-specific PPI networks for lean, obese, and post-bariatric surgery subjects

A) Average ranking of genes associated with the most sensitive kinetic parameters in lean, obese, and post-surgery subjects. Top left. The genes associated with each of the sensitive parameters were obtained utilizing the Molecular Biology of the Cell Ontology (MBCO) algorithm and were ranked according to their relative MBCO score. **Top right.** The most sensitive kinetic parameters in lean, obese, and post-surgery subjects were ranked according to their PLSR rank obtained from the MRA. **Top center.** The average ranking score for each gene was calculated as the mean of the MBCO rank and the MRA rank.

B) Weighting of the encoding genes in the unweighted PPI network using the results from dynamical modeling analysis and MBCO classification. Bottom left: The average rank score for each gene (left) was used to weight each corresponding node in the unfiltered genetic & genomic PPI network (**center**). Since the same gene can rank differently in lean, obese, and post-surgery subjects, the analysis is performed separately for each condition. The remaining nodes encoding proteins within the lean-, obese-, and post-surgery PPI networks were assigned a rank according to their distances from each of the average ranked genes (*edge thickness*). Distances are calculated from 1 to *n* and generate a physiologically weighted PPI network for each physiological condition (lean, obese, post-surgery) (**right**).

Figure 5. Organ/tissue-specific gene expression patterns in the physiologically weighted PPI networks. The physiologically weighted PPIs for lean, obese, and post-surgery conditions were divided according to the expression of each gene in the indicated tissues/organs, resulting in the generation of a PPI subnetwork for each site.

Figure 6: The 10 FDA-approved drugs that have the highest FAERS risk for weight gain and the 10 with the highest risk for weight loss as side effects, and their respective targets. Targets are listed by protein name and were obtained from multiple databases: DrugBank, The Druggable Genome, and Therapeutic Targets Database. See **Supplementary File 5** for the complete list of all the drugs and their respective targets that have weight gain or weight loss as side effects

Organ/tissue distribution of protein targets of weight gain- & weight loss-inducing drugs. Those targets of the 20 FDA-approved drugs listed in table 1 which are expressed in the 6 peripheral organs/tissues of interest are listed according to their expression retrieved from the GTEx database.

Figure 7. An initial human obesity atlas reveals organ- and tissue-specific metabolic differences in lean, obese, and post-bariatric surgery conditions.

A) Statistically significant catabolic and anabolic processes in lean, obese, and post-surgery subjects. (Top) The number of catabolic and anabolic processes within the top 100 biological subcellular processes was identified by the MCBO algorithm for the whole body for each physiological condition. (**B**ottom) Ratio of catabolic to anabolic processes for lean, obese and post-bariatric surgery patients.

B) Mapping the ratio of catabolic to anabolic processes onto individual tissues/organs in lean, obese, and post- bariatric surgery subjects. Using the obesity atlas, differences in overall metabolism can be mapped to individual tissues and organs in persons with different physiological phenotypes.
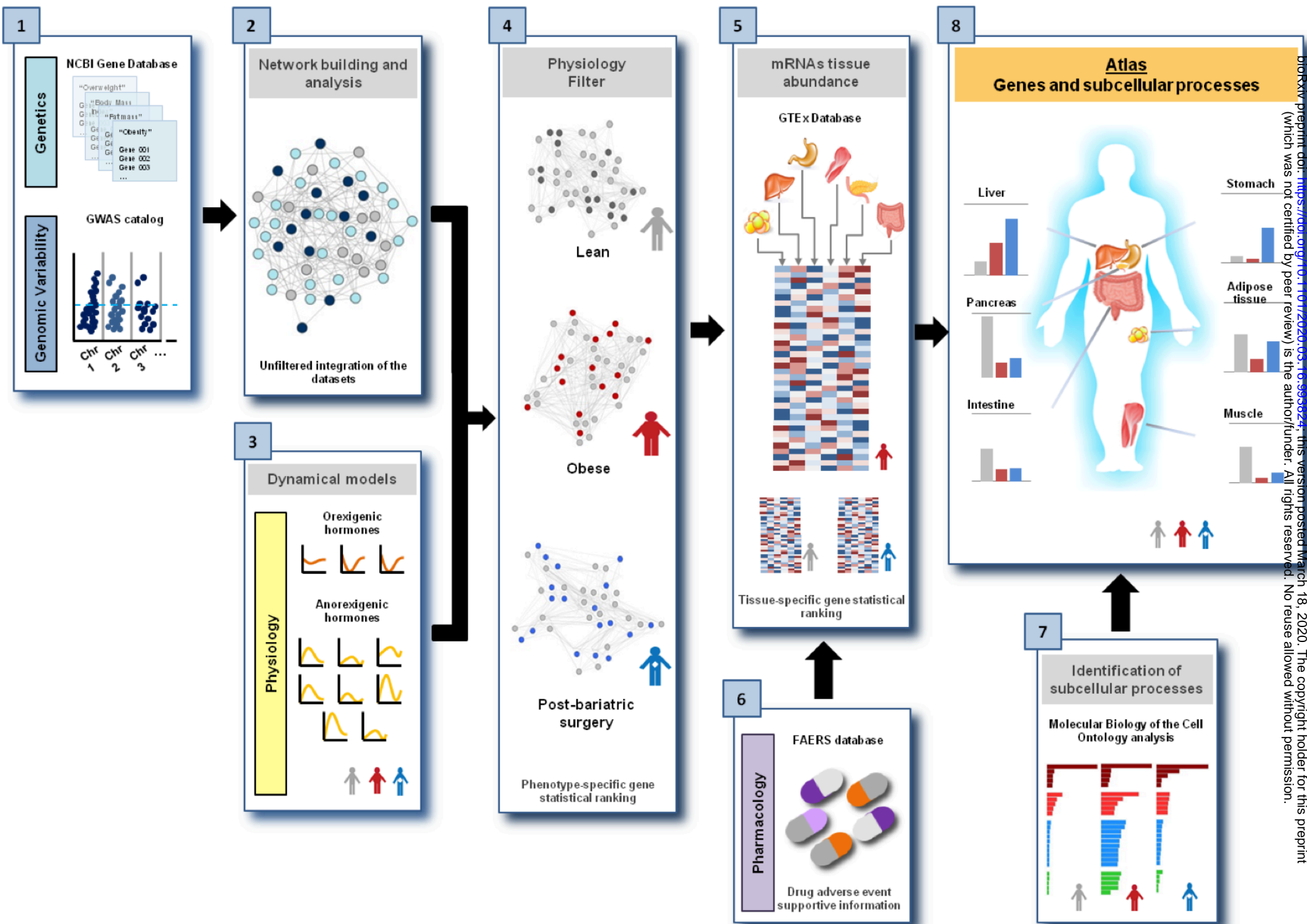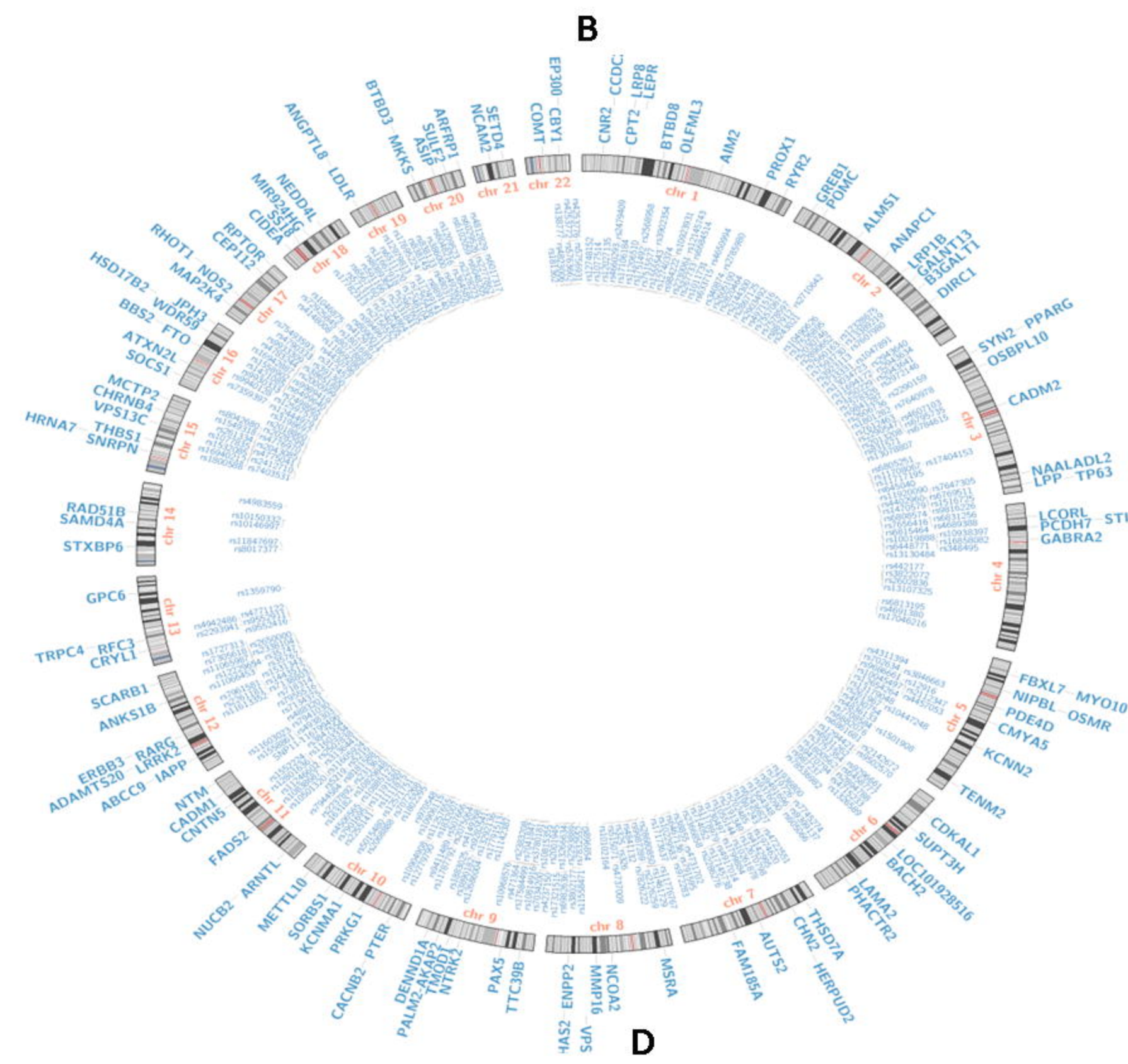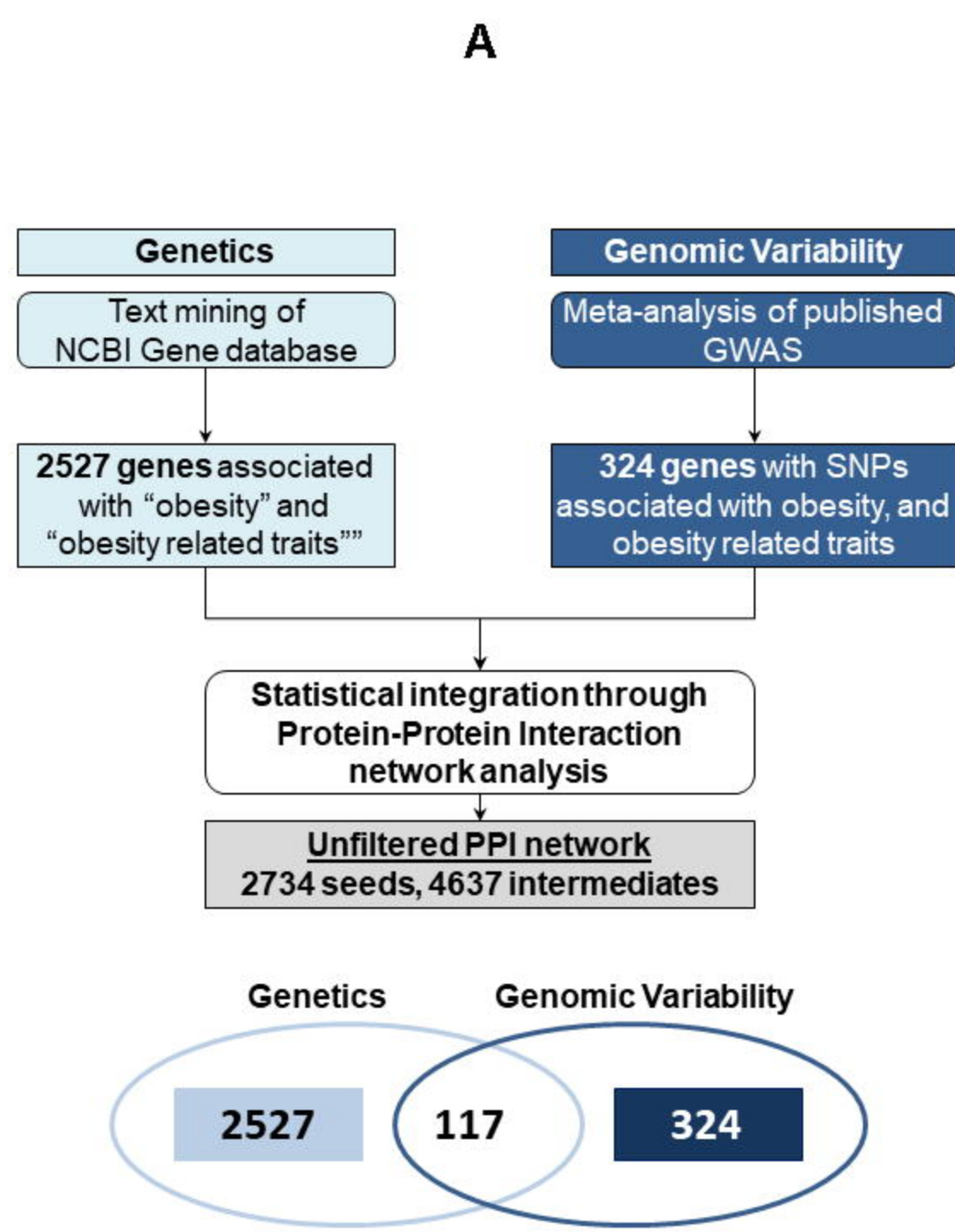
REFERENCES

1. Hales, C.M., et al., *Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007-2008 to 2015-2016.* Jama, 2018. **319**(16): p. 1723-1725.
2. James, W.P.T., *WHO recognition of the global obesity epidemic.* International journal of obesity, 2009. **32**(S7): p. S120.
3. Herrera, B.M. and C.M. Lindgren, *The genetics of obesity.* Current diabetes reports, 2010. **10**(6): p. 498-505.
4. O'rahilly, S., *Human genetics illuminates the paths to metabolic disease.* Nature, 2009. **462**(7271): p. 307.
5. McCarthy, M.I., *Genomics, type 2 diabetes, and obesity.* New England Journal of Medicine, 2010. **363**(24): p. 2339-2350.
6. Lu, Y. and R.J. Loos, *Obesity genomics: assessing the transferability of susceptibility loci across diverse populations.* Genome medicine, 2013. **5**(6): p. 55.
7. Sun, X., et al., *From genetics and epigenetics to the future of precision treatment for obesity.* Gastroenterology report, 2017. **5**(4): p. 266-270.
8. Chronaiou, A., et al., *Lower ghrelin levels and exaggerated postprandial peptide-YY, glucagon-like peptide-1, and insulin responses, after gastric fundus resection, in patients undergoing Roux-en-Y gastric bypass: a randomized clinical trial.* Obesity surgery, 2012. **22**(11): p. 1761-1770.
9. Nguyen, N.T. and J.E. Varela, *Bariatric surgery for obesity and metabolic disorders: state of the art.* Nature reviews Gastroenterology & hepatology, 2017. **14**(3): p. 160.
10. Tavassoly, I., J. Goldfarb, and R. Iyengar, *Systems biology primer: the basic methods and approaches.* Essays in biochemistry, 2018. **62**(4): p. 487-500.
11. Ghosh, S. and C. Bouchard, *Convergence between biological, behavioural and genetic determinants of obesity.* Nature Reviews Genetics, 2017. **18**(12): p. 731.
12. Benelam, B., *Satiation, satiety and their effects on eating behaviour.* Nutrition Bulletin, 2009. **34**(2): p. 126-173.
13. Brown, G.R., et al., *Gene: a gene-centered information resource at NCBI.* Nucleic acids research, 2014. **43**(D1): p. D36-D42.
14. Rosenbloom, K.R., et al., *ENCODE data in the UCSC Genome Browser: year 5 update.* Nucleic acids research, 2012. **41**(D1): p. D56-D63.
15. Kent, W.J., et al., *The human genome browser at UCSC.* Genome research, 2002. **12**(6): p. 996-1006.
16. Tyner, C., et al., *The UCSC genome browser database: 2017 update.* Nucleic acids research, 2016. **45**(D1): p. D626-D634.
17. Carithers, L.J. and H.M. Moore, *The genotype-tissue expression (GTEx) project*, 2015, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA.
18. Zhao, S., et al., *Systems pharmacology of adverse event mitigation by drug combinations.* Science translational medicine, 2013. **5**(206): p. 206ra140-206ra140.
19. Wishart, D.S., et al., *DrugBank: a comprehensive resource for in silico drug discovery and exploration.* Nucleic acids research, 2006. **34**(suppl_1): p. D668-D672.
20. Hansen, J., et al., *A flexible ontology for inference of emergent whole cell function from relationships between subcellular processes.* Scientific reports, 2017. **7**(1): p. 17689.

21.   Oliván, B., et al., *Effect of weight loss by diet or gastric bypass surgery on peptide YY3–36 levels.* Annals of surgery, 2009. **249**(6): p. 948.
22.   le Roux, C.W., et al., *Gut hormone profiles following bariatric surgery favor an anorectic state, facilitate weight loss, and improve metabolic parameters.* Annals of surgery, 2006. **243**(1): p. 108.
23.   Druce, M., et al., *Ghrelin increases food intake in obese as well as lean subjects.* International journal of obesity, 2005. **29**(9): p. 1130.
24.   Labayen, I., et al., *Role of baseline leptin and ghrelin levels on body weight and fat mass changes after an energy-restricted diet intervention in obese women: effects on energy metabolism.* The Journal of Clinical Endocrinology & Metabolism, 2011. **96**(6): p. E996-E1000.
25.   Falkén, Y., et al., *Changes in glucose homeostasis after Roux-en-Y gastric bypass surgery for obesity at day three, two months, and one year after surgery: role of gut peptides.* The Journal of Clinical Endocrinology & Metabolism, 2011. **96**(7): p. 2227-2235.
26.   Stępień, M., et al., *Serum concentrations of adiponectin, leptin, resistin, ghrelin and insulin and their association with obesity indices in obese normo-and hypertensive patients–pilot study.* Archives of medical science: AMS, 2012. **8**(3): p. 431.
27.   Martins, C., et al., *Impact of sustained weight loss achieved through Roux-en-Y gastric bypass or a lifestyle intervention on ghrelin, obestatin, and ghrelin/obestatin ratio in morbidly obese patients.* Obesity surgery, 2011. **21**(6): p. 751.
28.   Carlson, J.J., et al., *Pre-and post-prandial appetite hormone levels in normal weight and severely obese women.* Nutrition & metabolism, 2009. **6**(1): p. 32.
29.   Bergman, R.N., L.S. Phillips, and C. Cobelli, *Physiologic evaluation of factors controlling glucose tolerance in man: measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose.* The Journal of clinical investigation, 1981. **68**(6): p. 1456-1467.
30.   Sumithran, P., et al., *Long-term persistence of hormonal adaptations to weight loss.* New England Journal of Medicine, 2011. **365**(17): p. 1597-1604.
31.   Nakazato, M., et al., *Islet amyloid polypeptide, a novel pancreatic peptide, is a circulating hormone secreted under glucose stimulation.* Biochemical and biophysical research communications, 1990. **169**(2): p. 713-718.
32.   Jacobsen, S.H., et al., *Changes in gastrointestinal hormone responses, insulin sensitivity, and beta-cell function within 2 weeks after gastric bypass in non-diabetic subjects.* Obesity surgery, 2012. **22**(7): p. 1084-1096.
33.   Holst, J., et al., *Plasma enteroglucagon after jejunoileal bypass with 3: 1 or 1: 3 jejunoileal ratio.* Scandinavian journal of gastroenterology, 1979. **14**(2): p. 205-207.
34.   Sobie, E.A., *Parameter sensitivity analysis in electrophysiological models using multivariable regression.* Biophysical journal, 2009. **96**(4): p. 1264-1274.
35.   Heppner, K.M. and D. Perez-Tilve, *GLP-1 based therapeutics: simultaneously combating T2DM and obesity.* Frontiers in neuroscience, 2015. **9**: p. 92.
36.   Cotto, K.C., et al., *DGIdb 3.0: a redesign and expansion of the drug–gene interaction database.* Nucleic acids research, 2017. **46**(D1): p. D1068-D1073.
37.   Lodish, H., et al., *Molecular cell biology.* 2008: Macmillan.
38.   Li, X., et al., *Transcriptional atlas of cardiogenesis maps congenital heart disease interactome.* Physiological genomics, 2014. **46**(13): p. 482-495.
39.   Puchalski, R.B., et al., *An anatomic transcriptional atlas of human glioblastoma.* Science, 2018. **360**(6389): p. 660-663.
40.   Rani, J., et al., *T2DiACoD: a gene atlas of type 2 diabetes mellitus associated complex disorders.* Scientific reports, 2017. **7**(1): p. 6892.
41.   Kunej, T., et al., *Obesity gene atlas in mammals.* Journal of Genomics, 2013. **1**: p. 45.
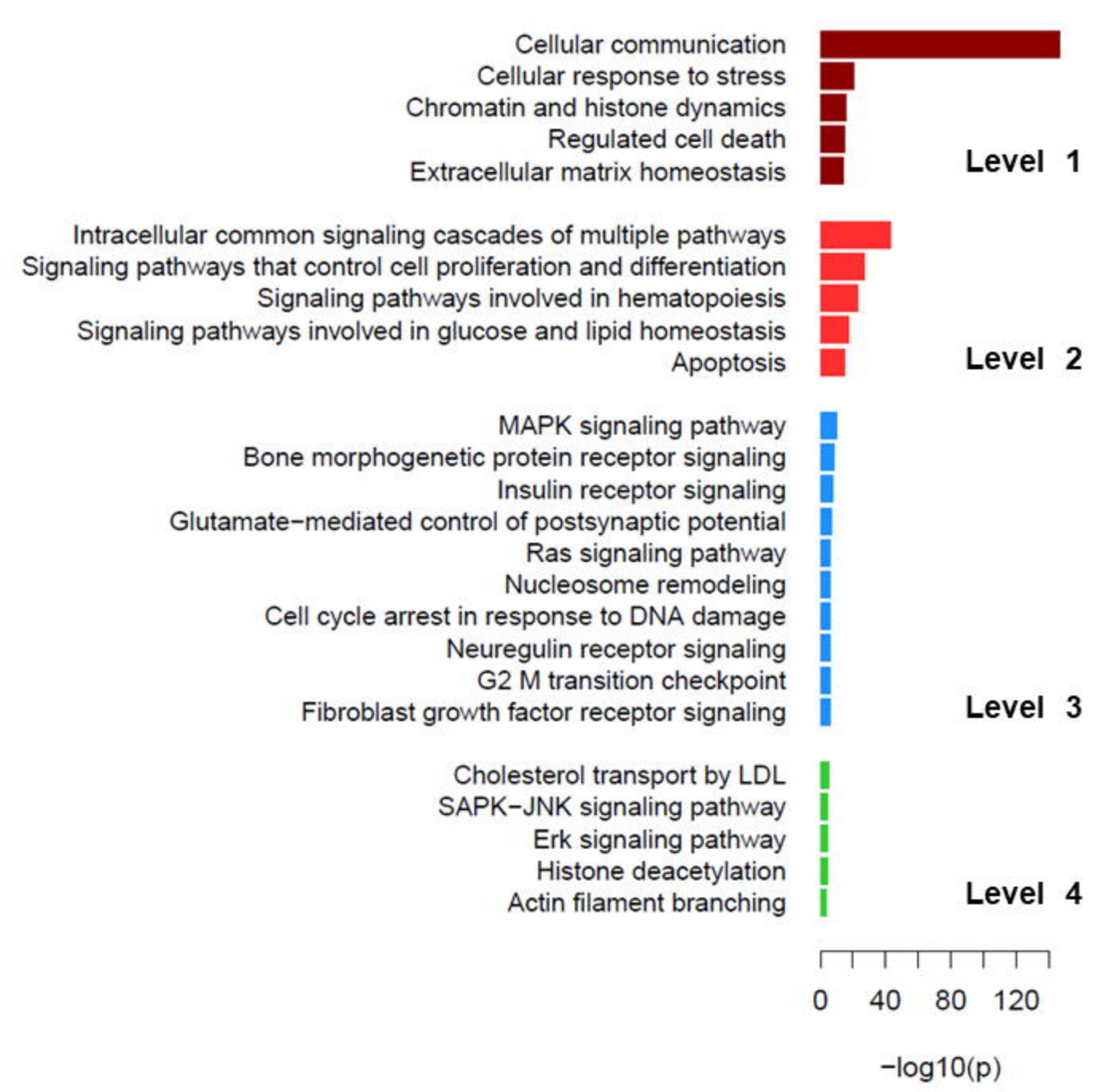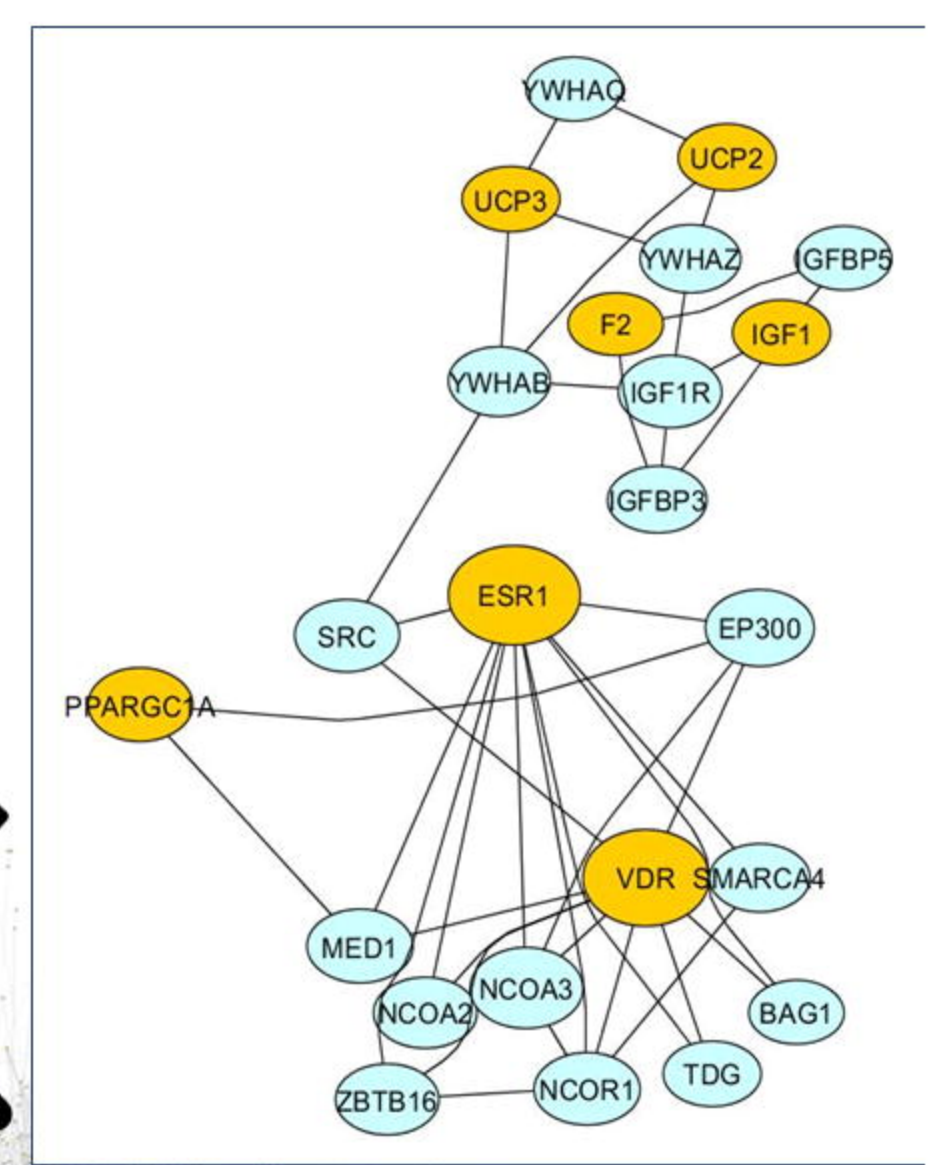
42. Tavassoly, I., et al., *Genomic signatures defining responsiveness to allopurinol and combination therapy for lung cancer identified by systems therapeutics analyses.* Molecular Oncology, 2019. **13**(8): p. 1725-1743.

43. Schwartz, M.W., *Central nervous system regulation of food intake.* Obesity, 2006. **14**(S2): p. 1S-8S.

44. Heisler, L.K. and D.D. Lam, *An appetite for life: brain regulation of hunger and satiety.* Current opinion in pharmacology, 2017. **37**: p. 100-106.

45. MacArthur, J., et al., *The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog).* Nucleic acids research, 2016. **45**(D1): p. D896-D901.

46. Chen, E.Y., et al., *Expression2Kinases: mRNA profiling linked to multiple upstream regulatory layers.* Bioinformatics, 2011. **28**(1): p. 105-111.

47. Alfarano, C., et al., *The biomolecular interaction network database and related tools 2005 update.* Nucleic acids research, 2005. **33**(suppl_1): p. D418-D424.

48. Keshava Prasad, T., et al., *Human protein reference database—2009 update.* Nucleic acids research, 2008. **37**(suppl_1): p. D767-D772.

49. Chatr-Aryamontri, A., et al., *MINT: the Molecular INTeraction database.* Nucleic acids research, 2006. **35**(suppl_1): p. D572-D574.

50. Nishimura, D., *BioCarta.* Biotech Software & Internet Report: The Computer Software Journal for Scient, 2001. **2**(3): p. 117-120.

51. Lynn, D.J., et al., *InnateDB: facilitating systems‐level analyses of the mammalian innate immune response.* Molecular systems biology, 2008. **4**(1): p. 218.

52. Mewes, H.-W., et al., *MIPS: a database for genomes and protein sequences.* Nucleic acids research, 2002. **30**(1): p. 31-34.

53. Ma'ayan, A., et al., *SNAVI: Desktop application for analysis and visualization of large-scale signaling networks.* BMC systems biology, 2009. **3**(1): p. 10.

54. Chatr-Aryamontri, A., et al., *The BioGRID interaction database: 2017 update.* Nucleic acids research, 2017. **45**(D1): p. D369-D379.

55. Kerrien, S., et al., *IntAct—open source resource for molecular interaction data.* Nucleic acids research, 2006. **35**(suppl_1): p. D561-D565.

56. Xenarios, I., et al., *DIP: the database of interacting proteins.* Nucleic acids research, 2000. **28**(1): p. 289-291.

57. Beuming, T., et al., *PDZBase: a protein–protein interaction database for PDZ-domains.* Bioinformatics, 2004. **21**(6): p. 827-828.

58. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic acids research, 2000. **28**(1): p. 27-30.

59. Dalla Man, C., R.A. Rizza, and C. Cobelli, *Meal simulation model of the glucose-insulin system.* IEEE Transactions on biomedical engineering, 2007. **54**(10): p. 1740-1749.

60. Hovorka, R., et al., *Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes.* Physiological measurement, 2004. **25**(4): p. 905.

61. Tavassoly, I., et al., *Dynamic modeling of the interaction between autophagy and apoptosis in mammalian cells.* CPT: pharmacometrics & systems pharmacology, 2015. **4**(4): p. 263-272.

62. Tavassoly, I., *Dynamics of Cell Fate Decision Mediated by the Interplay of Autophagy and Apoptosis in Cancer Cells: Mathematical Modeling and Experimental Observations.* 2015: Springer.

63. Dorvash, M., et al., *Dynamic Modeling of Signal Transduction by mTOR Complexes in Cancer.* Journal of Theoretical Biology, 2019: p. 109992.

64. Dorvash, M., M. Farahmandnia, and I. Tavassoly, *A Systems Biology Roadmap to Decode mTOR Control System in Cancer.* Interdisciplinary Sciences: Computational Life Sciences, 2019.

65. Miyashita, Y., et al., *Comments on the NIPALS algorithm.* Journal of chemometrics, 1990. **4**(1): p. 97-100.

66. Lorber, A., L.E. Wangen, and B.R. Kowalski, *A theoretical foundation for the PLS algorithm.* Journal of Chemometrics, 1987. **1**(1): p. 19-31.

67. Abdi, H., *Partial least squares regression.* Encyclopedia of measurement and statistics, 2007. **2**: p. 740-744.

68. Griffith, M., et al., *DGIdb: mining the druggable genome.* Nature methods, 2013. **10**(12): p. 1209.

69. Hopkins, A.L. and C.R. Groom, *The druggable genome.* Nature reviews Drug discovery, 2002. **1**(9): p. 727.

70. Chen, X., Z.L. Ji, and Y.Z. Chen, *TTD: therapeutic target database.* Nucleic acids research, 2002. **30**(1): p. 412-415.

A

B

C

D

Genetics

Text mining of NCBI Gene database

2527 genes associated with "obesity" and "obesity related traits"

Genomic Variability

Meta-analysis of published GWAS

324 genes with SNPs associated with obesity, and obesity related traits

Statistical integration through Protein-Protein Interaction network analysis

Unfiltered PPI network
2734 seeds, 4637 intermediates

Genetics   Genomic Variability

2527   117   324

YWHAQ
UCP2
UCP3
YWHAZ
IGFBP5
F2
IGF1
YWHAB
IGF1R
IGFBP3
ESR1
SRC
EP300
PPARGC1A
VDR
SMARCA4
MED1
BAG1
NCOA2
NCOA3
ZBTB16
NCOR1
TDG

Cellular communication
Cellular response to stress
Chromatin and histone dynamics
Regulated cell death
Extracellular matrix homeostasis
Level 1

Intracellular common signaling cascades of multiple pathways
Signaling pathways that control cell proliferation and differentiation
Signaling pathways involved in hematopoiesis
Signaling pathways involved in glucose and lipid homeostasis
Apoptosis
Level 2

MAPK signaling pathway
Bone morphogenetic protein receptor signaling
Insulin receptor signaling
Glutamate-mediated control of postsynaptic potential
Ras signaling pathway
Nucleosome remodeling
Cell cycle arrest in response to DNA damage
Neuregulin receptor signaling
G2 M transition checkpoint
Fibroblast growth factor receptor signaling
Level 3

Cholesterol transport by LDL
SAPK-JNK signaling pathway
Erk signaling pathway
Histone deacetylation
Actin filament branching
Level 4

0   40   80   120

−log10(p)

A

Text mining of PubMed using MBCO algorithm

873 unique genes associated with kinetic parameters

MBCO ranking of the genes associated with each sensitive kinetic parameter

Kinetic Parameters

873 genes

MBCO Rank

AVERAGE RANK

$$\frac{\text{MBCO Rank (Gene)} + \text{MRA Rank (Parameter)}}{2}$$

MRA Rank of kinetic parameters

Kinetic Parameters	Lean	Obese	Post-Surgery
amylin secretion	4	11	17
CCK degradation	17	3	25
CCK secretion	10	14	8
CCK utilization	19	26	4
ghrelin degradation	3	13	7
ghrelin secretion	2	2	1
ghrelin utilization	30	8	10
GIP secretion	29	9	16
GLP-1 degradation	6	29	9
GLP-1 secretion	28	6	29
GLP-1 utilization	9	7	21
insulin degradation	7	16	11
insulin utilization	8	4	5
leptin degradation	26	10	13
leptin secretion	15	1	2
leptin utilization	33	21	3
OXM degradation	1	24	24
PYY degradation	5	25	14
PYY utilization	25	20	6

Unfiltered Genetic and Genomic Network

B

Average ranking of the genes associated with the most sensitive kinetic parameters in lean, obese, and post-surgery subjects

Kinetic Parameters

873 genes

Identification of the 873 genes within the unfiltered genetic and genomic network

Ranking of nodes by Average Rank

Distances of genetic & genomic nodes from each average ranked node

LEAN Physiologically weighted PPI network

873 genes

Identification of the 873 genes within the unfiltered genetic and genomic network

Ranking of nodes by Average Rank

Distances of genetic & genomic nodes from each average ranked node

OBESE Physiologically weighted PPI network

873 genes

Identification of the 873 genes within the unfiltered genetic and genomic network

Ranking of nodes by Average Rank

Distances of genetic & genomic nodes from each average ranked node

POST-SURGERY Physiologically weighted PPI network

Top 10 drugs for weight gain by FAERS risk

Drug	Targets												
FLUOXETINE	PTGFR	C1QBP	CD9	SLC6A4	HTR2A	BRCA2							
BREXPIPRAZOLE	HTR1A	DRD2	HTR2A	ADRA2C	ADRA1B	DDX6	HTR7						
PHENELZINE	MAOA	GNA11	MAOB	SLC6A3	SLC6A2	CD9							
OLANZAPINE	GABRA1	ADRB1	ALDH2	LEP	RGS4	CYP2D6	ADRB3	GNB3	KCNH2	SLC1A1	CYP1A2	FAAH	MC4R
ZIPRASIDONE	BRCA1	DRD2	DNMT3A	CD9	HTR1A	HTR2A	CHRNA7	CCND1	ADRA1A	ADRB1	AKT1	HTR1E	ALK	A2MP1
HALOPERIDOL	HTR2A	DRD2	BRCA1	ALK	DRD1	ADAM9	A2MP1	BRCA2	SERPINA3	ALDH2	ALDH3A1	DNMT3A	BAAT	DDX6
TRIFLUOPERAZINE	DRD2	ADRA1A	CALY	BRCA1	DDX6	ALDH2	DNMT3A	CCND1					
CORTICOTROPIN	MC2R	CRH	APAF1	MC1R	ANGPT1	ADRB1	MC5R	NUDT2					
GEMTUZUMAB	BRAF	DDX6	DNMT3A	HTR6	PPARG	CCND1	CDK4	ANKK1	GIPR				
TIOTIXENE	CCND1	ADRA1A	ADRB1	AKT1	HTR1E	ALK	A2MP1						

Top 10 drugs for weight loss by FAERS risk

Drug	Targets						
EXENATIDE	GLP1R						
LOMITAPIDE	MTTP	MT1B	MTNR1A	ANK2	MTNR1B		
PAMIDRONIC ACID	FDPS	CLTC	BCR				
FLUTAMIDE	AR	AHR	NR1I2				
PRAMLINTIDE	CALCR	RAMP1	RAMP2	RAMP3	ACLS	ACHE	
OCTREOTIDE	SSTR2	AQP2	SHROOM2	SSTR1	SSTR5		
ROFLUMILAST	PDE4D	PDE4A	PDE4B	PDE4C	COL11A1		
NATEGLINIDE	ABCC8	PPARG	HCN2	CYP3A5	CYP3A4	KCNJ11	KCNJ1
FULVESTRANT	ESR1	ESR2	EREG	ALDH3B1	CPD		
PIOGLITAZONE	CCND1	ATP4A	PPARG	PPARA			

Adipose	Muscle	Intestine	Stomach	Pancreas	Liver
PPARG	CACNA1S	EPHB2	PPARG	CYP3A5	CPS1
ADRB2	PPARA	CPS1	CYP3A5	GABRE	NR1I2
ANGPT1		PPARG	JAK2		CYP21A2
JAK2		NR1I2	PTGS1		CYP3A5
PTGS1		IL7R	CYP2C9		CYP3A4
AR		CYP3A5	PTGDR2		ADRA1A
GABRE		CYP3A4	ALDH3A1		BAAT
		JAK2			CYP2C19
		PTGS1			CYP2C9
		CYP2C19			APOF
		CYP2C9			CYP1A2
		PTGDR2			DHODH
		SLC6A4			CYP2A6
		PPARA			PPARA
		OTC			OTC
		GABRE			AR

A

B

Table 1

Top 10 drugs for weight gain by risk

Drug	FAERS Risk	Targets												
FLUOXETINE	12.540	PTGFR	C1QBP	CD9	SLC6A4	HTR2A	BRCA2							
BREXPIPRAZOLE	7.980	HTR1A	DRD2	HTR2A	ADRA2C	ADRA1B	DDX6	HTR7						
PHENELZINE	7.536	MAOA	GNA11	MAOB	SLC6A3	SLC6A2	CD9							
OLANZAPINE	7.182	GABRA1	ADRB1	ALDH2	LEP	RGS4	CYP2D6	ADRB3	GNB3	KCNH2	SLC1A1	CYP1A2	FAAH	MC4R
ZIPRASIDONE	6.827	BRCA1	DRD2	DNMT3A	CD9	HTR1A	HTR2A	CHRNA7	CCND1	ADRA1A	ADRB1	AKT1	HTR1E	ALK	A2MP1
HALOPERIDOL	5.923	HTR2A	DRD2	BRCA1	ALK	DRD1	ADAM9	A2MP1	BRCA2	SERPINA3	ALDH2	ALDH3A1	DNMT3A	BAAT	DDX6
TRIFLUOPERAZINE	5.887	DRD2	ADRA1A	CALY	BRCA1	DDX6	ALDH2	DNMT3A	CCND1						
CORTICOTROPIN	5.767	MC2R	CRH	APAF1	MC1R	ANGPT1	ADRB1	MC5R	NUDT2						
GEMTUZUMAB	5.703	BRAF	DDX6	DNMT3A	HTR6	PPARG	CCND1	CDK4	ANKK1	GIPR					
TIOTIXENE	5.614	CCND1	ADRA1A	ADRB1	AKT1	HTR1E	ALK	A2MP1							

Top 10 drugs for weight loss by risk

Drug	FAERS Risk	Targets						
LOMITAPIDE	9.857	MTTP	MT1B	MTNR1A	ANK2	MTNR1B		
EXENATIDE	16.654	GLP1R						
PAMIDRONIC ACID	8.343	FDPS	CLTC	BCR				
FLUTAMIDE	7.816	AR	AHR	NR1I2				
PRAMLINTIDE	6.923	CALCR	RAMP1	RAMP2	RAMP3	ACLS	ACHE	
OCTREOTIDE	6.133	SSTR2	AQP2	SHROOM2	SSTR1	SSTR5		
ROFLUMILAST	5.594	PDE4D	PDE4A	PDE4B	PDE4C	COL11A1		
NATEGLINIDE	5.564	ABCC8	PPARG	HCN2	CYP3A5	CYP3A4	KCNJ11	KCNJ1
FULVESTRANT	5.507	ESR1	ESR2	EREG	ALDH3B1	CPD		
PIOGLITAZONE	5.486	CCND1	ATP4A	PPARG	PPARA			

Table 2

Adipose	Muscle	Intestine	Stomach	Pancreas	Liver
PPARG	CACNA1S	EPHB2	PPARG	CYP3A5	CPS1
ADRB2	PPARA	CPS1	CYP3A5	GABRE	NR1I2
ANGPT1		PPARG	JAK2		CYP21A2
JAK2		NR1I2	PTGS1		CYP3A5
PTGS1		IL7R	CYP2C9		CYP3A4
AR		CYP3A5	PTGDR2		ADRA1A
GABRE		CYP3A4	ALDH3A1		BAAT
		JAK2			CYP2C19
		PTGS1			CYP2C9
		CYP2C19			APOF
		CYP2C9			CYP1A2
		PTGDR2			DHODH
		SLC6A4			CYP2A6
		PPARA			PPARA
		OTC			OTC
		GABRE			AR