

Working title:

Genetic aetiology of the persistent lean phenotype, exposed to an unhealthy environment.

Student: JERNEJA MISLEJ

University of Lund

bif15jmi@student.lu.se

Project supervisor: ALAITZ POVEDA

University of Lund

Alaitz.Poveda@med.lu.se

Assistant project supervisors: PAUL W. FRANKS, AZRA KURBASIC

University of Lund

Paul.Franks@med.lu.se

Azra.Kurbasic@med.lu.se

Abstract

Although during the last decades, obesity has reached epidemic proportions in high-income countries, mostly due to changes in lifestyle, there is still a part of the population remaining lean, despite of living in an unhealthy environment.

The master thesis project objective is to design, implement and perform the identification and analysis of a lean phenotype, which is persistent despite an unhealthy environment and potentially present in a large, northern Swedish cohort. An extensive collection of environmental, genealogical, genetic, and clinical longitudinal data is available within this population-based prospective cohort study, enabling exploration of genetic and lifestyle risk factors associated to a persistent lean phenotype.

Further research will be enabled by the successful identification of subjects that are able to maintain a lean phenotype, although exposed to an unhealthy lifestyle on a longitudinal basis. By switching the focus from subjects who are diagnosed, predisposed or at risk for obesity, to subjects with a persistent lean phenotype, protective characteristic against obesity can be explored and established. Proposed analyses could lead to the identification of genetic factors associated with persistent leanness or leastwise confirm the implication of previously established genetic variants.

I Basic information

Within the masters program of Bioinformatics at Lund University, Department of Biology, I will carry out a 30 credit master thesis project under the code BINP30.

The project will take place at the department of Genetic and Molecular Epidemiology in Lund University Diabetes Center at the Clinical Research Center in Malmö, under the supervision of Alaitz Poveda and co-supervision of Paul Franks and Azra Kurbasic.

The project will start in week 4 on 23th of January 2017 and will finish in week 26, on the 26th of June 2017.

II Project

I Introduction

Obesity is a highly prevalent disease and one of the greatest public health concerns in high-income countries. The speed of the progression of obesity epidemic, over the last decades, indicates that at a population level, the global rising in obesity worldwide is due to a change in environmental factors, such as physical activity engagement and diet, rather than to an alteration of the genetic make-up of the population. However, a part of the population remains lean, although exposed to an unhealthy lifestyle. Identification of subjects with a persistent lean phenotype in an unhealthy environment and analysis of the genetic variants contributing to this phenotype could help with the understanding of the epidemiology of obesity[3][4][5].

The project will be conducted with an extensive collection of lifestyle and clinical longitudinal data, available within a population-based prospective cohort study[1]. The data consists of approximately 600 variables, with partially obtained values for approximately 25000 subjects from the northern Swedish population of the province Västerbotten collected within the framework of the Västerbottens Health Survey (also called the Västerbottens Intervention Project) initiated in 1985[1]. In the Västerbottens Health Survey, residents within the county are invited to attend an extensive health examination in the years of their 40th, 50th and 60th birthdays. For the current analysis, health examinations were performed between 1985 and 2013 with a follow-up time of around 10 years between baseline and follow-up measurements. Lifestyle variables cover environmental exposures such as diet, physical activity, smoking, alcohol and education. Clinical variables cover a large range of medical and anthropometric characteristics, including plasma glucose, serum total cholesterol and triglycerides concentrations. A subset of the sample, 5000 participants, also contains information about genetic data as they have been GWAS genotyped.

II Project aims

The overarching aim of the project goal is to identify and analyze the genetic etiology of subjects having a persistent lean phenotype, while being exposed to an unhealthy envi-

ronment. The first step will include familiarization with the data and topic, followed by data exploration and preparation. In the second step, script(s) for identification of certain specific subjects will be designed, implemented and tested. Subjects in the identified subset will have to be persistently exposed to lifestyle factors established to lead to obesity. At the same time, the subjects will have to possess characteristics of a lean and healthy person, with no confounding medical condition or environmental factor that could explain the persistent leanness despite the persistent unhealthy lifestyle. Once participants with persistent leanness have been identified, genetic analyses (heritability and genetic association analyses) will be conducted in order to further research the genetic epidemiology of the lean phenotype.

III Methods

Several lifestyle factors, such as diet, physical activity, smoking and education have been established to lead to obesity, with a high long-range predictive accuracy[2]. Using lifestyle information from the dataset of northern Sweden, one or several lifestyle scores will be constructed, in order to identify participants exposed to an unhealthy lifestyle both at baseline and follow-up. Among those participants, methods serving for the lean phenotype subset identification will be applied in order to identify the subset of participants that could be considered persistent lean although exposed to an unhealthy lifestyle. Additional methods are to be adapted and developed to ensure the final subset, presenting the lean phenotype, will not include subjects who are lean due to confounding medical conditions or environmental factors. Quantitative genetic analyses based on genealogical information will be conducted on the SOLAR program[6], to estimate the heritability of the persistent lean phenotype being exposed to an unhealthy environment. The subset of participants containing GWAS information will be employed to conduct genetic association studies, in order to associate certain gene variants to the persistent lean phenotype.

IV Timeline

- 23.1.2017 - 10.2.2017: Project starts, the student gets acquainted with the topic and available data and prepares the data for processing.
- 13.2.2017 - 24.3.2017: Design, implementation and testing of script(s) for unhealthy lean phenotype subset identification.
- 27.3.2017 - 10.4.2017: Analysis of the lean phenotype subset, focusing on heritability estimation.
- 13.4.2017 - 26.5.2017: Analysis of the lean phenotype subset, focusing on identification and analysis of associated gene variants.
- 29.5.2017 - 9.6.2017: Completion and closure.

- 12.5.2017 - 26.6.2017: Master thesis report and presentation preparation.

References

- [1] A. Kurbasic *et al.*, Gene-Lifestyle Interactions in Complex Diseases: Design and Description of the GLACIER and VIKING Studies. *Curr Nutr Rep*, (2014)
- [2] A. Poveda *et al.*, Innate biology versus lifestyle behaviour in the aetiology of obesity and type 2 diabetes: the GLACIER Study. *Diabetologia*, (2016)
- [3] I. Harosh , Rare Genetic Diseases with Human Lean and/or Starvation Phenotype Open New Avenues for Obesity and Type II Diabetes Treatment. *Current Pharmaceutical Biotechnology*, (2013)
- [4] I. Harosh *et al.*, Enteropeptidase: A Gene Associated with a Starvation Human Phenotype and a Novel Target for Obesity Treatment. *PLoS ONE*7(11): e49612. doi:10.1371/journal.pone.0049612 (2012)
- [5] Jolivet G, Braud S, DaSilva B, Passet B, Harscoet E. *et al.*, Induction of Body Weight Loss through RNAi-Knockdown of APOBEC1 Gene Expression in Transgenic Rabbits. *PLoS ONE*9(9): e106655. doi:10.1371/journal.pone.0106655 (2014)
- [6] Almasy L, Blangero J, Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 62:1198-1211 (1998)