Supplementary Materials and Methods

Study group

We followed the STREGA guidelines (21) to describe the study group selection and association analysis. Case-control study groups were selected from the Latvian Genome Database (LGDB), a government funded biobank. All participants of LGDB must be over 18 years old, and information about their health status was affirmed by physicians using ICD-10 codes (International Classification of Diseases). Anthropometric measurements (including weight and stature) were obtained by direct measurement; ethnic, social, environmental information and familial health status were obtained using a questionnaire based interview. Participants of the LGDB were recruited by medical personnel in hospitals or general practices. Recruitment was population based (a specific health condition was not the obligatory requirement for involvement). Signed consent forms were acquired from all participants. The Biobank protocol was approved by the Central Medical Ethics Committee of Latvia (Protocols Nr. A-30, 2005 and A-7, 2007).

Acromegaly patients (n=48) recruited for this study were enrolled on to the LGDB between 2004 and 2008 from two main hospitals, Pauls Stradins Clinical University Hospital (41 patients) and Riga Eastern Clinical University Hospital (7 patients) representing approximately 80% of all the acromegaly patients registered in Latvia as recorded in October 2008 (diagnosed with acromegaly from 1985 to 2007). Additional data were collected based on hospital records and interviews for all the patients selected for the study (ICD-10 code E22.0). Forty-five patients received the somatostatin analogs octreotide (Sandostatin LAR) at dose 10–30 mg every 28 days or lanreotide (Somatuline Autogel) at dose 60–90 mg every 28 days. Tumor size was measured as the maximum diameter obtained from magnetic resonance imaging data and tumors were classified accordingly as microadenomas (<10mm) or macroadenomas (≥10mm). The effect of somatostatin analogs on tumor proliferation and IGF1 level normalization was estimated by comparing the tumor sizes and serum IGF1 (μg/ml) measurements during the course of therapy. In order to estimate the dynamics of adenoma size, the data from last follow-up was compared with the first available measurement with at least a 12 month period in between. We excluded all cases where the therapy was interrupted or adenoma resections were performed during this period. Two groups were defined: “reduced” with observable tumor shrinkage (n=11) and “unchanged” with no observable tumor shrinkage (n=22) together with the cases that showed prolonged expansion (n=2). Only a limited number of cases had IGF1 measurements available before the therapy. For the IGF1 response, only the data at least 6 months after the start of the SA therapy were considered for analysis and only if the therapy was not interrupted. In those cases where several IGF1 measurements were available, the mean IGF1 was calculated, excluding the outliers where possible. Non-responsiveness was defined as the mean IFG1 value above the upper limit of the normal (ULN) value at the corresponding age. Due to the lack of uniformity in the GH measurements, we did not include the GH levels in this analysis.

As controls (Control I), 96 samples were randomly chosen for sequencing from total of 2203 LGDB participants recruited from 2003 to 2005, excluding patients with metabolic and endocrine diseases. In order to minimize the risk of false positivity, we selected an additional control group (Control II) that was sex- and age-matched, consisting of 475 LGDB participants. This control II group was selected from all the 7935 LGDB participants recruited from 2003 to 2008 applying the following selection criteria. First we excluded subjects having any chronic disease including metabolic, endocrine, coronary heart, cancer and renal conditions (detailed list of ICD-10 codes is available upon a request). Similarly we excluded all participants missing the relevant phenotypic data. After this selection, 1225 healthy adults were considered for further use. Matching by age was performed in 10-year age groups, we randomly chose 10 sex-matched controls for each acromegaly patient from the corresponding age group. Five samples were omitted due to practical reasons (96 plate format). Under these criteria, 63 participants from the Control I group were also included in the Control II group. Study protocol was approved by Central Medical Ethics Committee of Latvia (Protocols Nr. A-33, 2005 and A-3, 2008).
DNA analysis

DNA samples were provided by LGDB and aliquoted into 96 well PCR plates or PCR tubes by Tecan Freedom Evo robotic pipette. Final DNA amount was 28ng per well. The SSTR5 gene containing genomic DNA region including 5’ and entire coding region (from –2239 to +1294 respective to start codon) was amplified in six PCR reactions. PCR primers for DNA amplification were designed using Primer3 software (primer sequences can be found in Supplementary Table 1). The following PCR reaction setup was used: 1mM DB buffer, 2.5mM MgCl2, 0.5 units Hot FirePol, 0.2mM dNTP mix (SolisBioDyne, Estonia), 0.3mM primers and 28ng template DNA. PCR temperatures were 95°C – 5min, 40 cycles – 95°C – 30sec, 55°C – 30sec, 72°C – 1min, and 5min at 72°C for final extension and the PCRs were carried out on Veriti96ThermalCycler (Applied Biosystems, USA). The PCR product was confirmed by agarose gel electrophoresis and dephosphorylation of remaining dNTPs was done with shrimp alkaline phosphatase (Fermentas, Lithuania) according to manufacturer’s protocol.

Amplification products had both strands directly sequenced using a set of 14 primers that were designed on Primer3 Software (Supplementary Table 1). Following sequencing reaction setup were used: 1μl BigDye, 2μl 5x BigDye sequencing buffer, 0.5mM corresponding primer, 150–250ng of template DNA (from the previously described PCR reaction) and 5.5μl distillated H2O. Reaction conditions were set: 96°C – 1min, 25 cycles - 96°C – 10sec, 50°C – 5sec, and 60°C – 4min. Sequencing was carried out on Veriti96ThermalCycler (Applied Biosystems, USA). The products were purified using Sephadex G50 (Sigma-Aldrich, USA) and sequenced with ABI Prism 3100 (AME Bioscience, Norway) capillary electrophoresis sequencer. All chromatograms were manually inspected using Contig Express software of Vector NTI Advance 9.0 package. Presence of polymorphisms was confirmed by opposite strand analysis.

Statistical analysis

Statistical analysis was performed using the PLINK 1.07 (23) and SPSS (Standard Version 13;SPSS, Chicago, IL, USA) software. Deviation from Hardy-Weinberg equilibrium was assessed by the exact test described by Wigginton et al (24) which is considered more accurate for rare genotypes. The Cochran-Armitage trend test was used for association analysis in the case control group and Bonferroni correction was applied. Haplotype association was performed as implemented in PLINK. For the quantitative analyses the IGF1 data were transformed as a normalized percentage of upper limit of normal (ULN) of appropriate age according to formula...
\((C_{IGF1} - ULN_{IGF1})/ULN_{IGF1} \times 100\). Normalized IGF and all other continuous variables displayed normal distribution and were further used in linear regression analysis. Two sided Fisher exact test was used to test the allelic distribution in the case of categorical clinical variables, except the analysis of number of adenoma resections where Pearson Chi-Square was calculated from 3x3 table. Permutation tests with 100,000 permutations were performed for each analysis and we used corrected (EMP2) \(P\) values. These values are corrected based on calculation of the proportion of permutations in which any of the test statistics exceeds the particular observed statistic and are more stringent than uncorrected \(P\) values.