Genetic Analyses in Small for Gestational Age Newborns

Susanne E Stalman[1,2], Nita Solanky[2], Miho Ishida[2], Cristina Alemán-Charlet[2], Sayeda Abu-Amero[2], Marielle Alders[3], Lucas Alvizi[1,4], William Baird[2], Charalambos Demetriou[2], Peter Henneman[3], Chela James[5], Lia C Knegt[3], Lydia J Leon[2], Marcel MAM Mannens[3], Adi N Mul[3], Nicole A Nibbering[3], Emma Peskett[2], Faisal I Rezwan[6], Carrie Ris-Stalpers[7], Joris AM van der Post[7], Gerdine A Kamp[8], Frans B Plötz[8], Jan M Wit[9], Philip Stanier[2], Gudrun E Moore[2], Raoul C Hennekam[1].

[1] Department of Pediatrics, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands. [2] Department of Genetics and Genomic Medicine, Institute of Child Health, University College London, London, United Kingdom. [3] Department of Clinical Genetics, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands. [4] Centro de Pesquisas Sobre o Genoma Humano e Células-Tronco, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brasil. [5] UCL Cancer Institute, University College London, London, United Kingdom. [6] Department of Human Development and Health, Southampton General Hospital, University of Southampton, Southampton, United Kingdom. [7] Department of Gynecology and Obstetrics, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands. [8] Department of Pediatrics, Tergooi Hospitals, Blaricum, The Netherlands. [9] Department of Pediatrics, Leiden University Medical Centre, Leiden, The Netherlands.

Abbreviated Title: Genetic analysis in small newborns

Key terms: small for gestational age; fetal growth; copy number variations; methylation disturbances; sequence variants

Word count: 4206

Number of figures and tables: 3 | ***Supplemental figures and tables:*** 9

Disclosure Summary: The authors have nothing to disclose.

Corresponding author and person to whom reprint requests should be addressed:

Susanne E. Stalman, MD

Department of Pediatrics H7-235

Academic Medical Center, University of Amsterdam

Meibergdreef 9, 1105AZ Amsterdam, The Netherlands

Email: s.e.stalman@amc.uva.nl, Telephone: +31 6 12402669

Abstract

Context: Small for gestational age (SGA) can be a result of fetal growth restriction, associated with perinatal morbidity and mortality. Mechanisms that control prenatal growth are poorly understood.

Objective: The aim of the present study was to gain more insight into prenatal growth failure and determine an effective diagnostic approach in SGA newborns. We hypothesized that one or more CNVs and disturbed methylation and sequence variants may be present in genes known to be associated with fetal growth.

Design: A prospective cohort study of subjects with a low birthweight for gestational age.

Setting: The study was conducted at an academic pediatric research institute.

Patients: A total of 21 SGA newborns with a mean birthweight below the 1st centile and a control cohort of 24 appropriate for gestational age newborns were studied.

Intervention: Array comparative genomic hybridization, genome-wide methylation studies and exome sequencing were performed.

Main Outcome Measures: The numbers of copy number variations, methylation disturbances and sequence variants.

Results: The genetic analyses demonstrated three CNVs, one systematically disturbed methylation pattern and one sequence variant explaining the SGA. Additional methylation disturbances and sequence variants were present 20 patients. In 19 patients, multiple abnormalities were found.

Conclusion: Our results confirm the influence of a large number of mechanisms explaining dysregulation of fetal growth. We conclude that copy number variations, methylation disturbances and sequence variants all contribute to prenatal growth failure. Such genetic workup can be an effective diagnostic approach in SGA newborns.

Introduction

The process of human fetal growth is steered by fetal and maternal genetic factors that affect the intrauterine environment to ensure effective nutrient exchange between mother and fetus via the placenta. SGA has been defined either as being <10th centile for weight at a given gestational age or as having a birth length or weight standard deviation score (SDS) of <-2.0 (<2.3rd centile) (1). SGA can be a result of fetal growth restriction (FGR), which is defined as a fetus being unable to reach its individual growth potential (2). FGR is associated with significant perinatal morbidity and mortality (3) and babies with FGR can be predisposed to metabolic diseases later in life (4).

Thirty to fifty percent of the variation in weight at birth can be explained by genetic or epigenetic causes (5), which include chromosome imbalances, sequence variants and epigenetic disturbances. The London Dysmorphology Database contains over 400 entities associated with prenatal growth failure (6). Numerous studies on epigenetic influences, especially DNA methylation disturbances, have also been performed (7). Despite this research, mechanisms behind prenatal growth failure are only poorly understood, at least in part due to the heterogeneous nature of growth disturbances. Consequently, an appropriate diagnostic workup for SGA newborns is not well established, and questions remain to what extent which genetic factors contribute, what the optimal care pathway is for the child, and how we provide adequate counselling to parents.

Aim of the present study is to gain further insight into prenatal growth failure and determine whether a combination of genomic analyses is an effective diagnostic approach for SGA newborns. We used array comparative genomic hybridization (array-CGH) to detect copy number variations (CNVs), genome-wide methylation studies to uncover methylation disturbances, and "whole" exome sequencing (WES) to detect sequence variants in a cohort of SGA newborns. We hypothesized that CNVs explaining the SGA may be found, disturbed methylation may be present in genes known to be aberrantly methylated in low birthweight newborns, and that sequence variants may be present in genes targeted because of their known association with SGA.

Methods

Patients

We selected 21 SGA newborns and their parents from the Baby Bio Bank (BBB) and Moore Cohort. The BBB contains biological samples and clinical data from 2,515 pregnancies, collected between 2000 and 2014. The Moore cohort consists of 319 trio samples collected from newborns and their parents between 1991 and 1994, including a small FGR cohort.

Inclusion criteria for this study included: weight at birth at or below 10^{thth} centile, availability of parental samples and absence of major structural malformations, to be in accordance to our study aims to study severe IUGR newborns without clues for a specific diagnosis. No pre-eclampsia/HELLP syndrome, maternal systemic disease, medication use during pregnancy or maternal smoking was present, except one mother (SGA4) who was a moderate smoker during pregnancy and one other mother (SGA3) who had pre-existing essential hypertension for which she received treatment. SGA17 was a pregnancy termination at 22 weeks of gestation (reason unknown to us), and was included because of the markedly low weight for gestational age without malformations or other clues for a specific diagnosis.

A control cohort of appropriate for gestational age (AGA) newborns (n = 24) was selected from the Preeclampsia And Non-preeclampsia Database (PANDA) Biobank based on birthweight for GA closest to the 50th centile and an equal distribution of GA and mode of delivery in relation to the SGA cases. The PANDA Biobank collected placental biopsies, umbilical cord blood samples and maternal blood samples of 400 women with either preeclampsia or normotensive pregnancies, between 2006 and 2010.

The SDS of weight at birth were calculated using the 1990 British growth references (8) for the British cases and the 1991 reference data for the Dutch controls (9). Descriptive statistics for analyzing demographic data was performed in IBM SPSS Statistics, version 22.

Ethical approval for all studies was obtained (BBB Research Ethics Committee references: 09/H0405/30 and 09/h0405/30+5, Moore cohort reference: 2001/6029, PANDA Biobank AMC2005_133).

Targeted genes

We performed literature searches on (1) genes known to be aberrantly methylated in SGA (**Supplemental Table 1**) (2) genes known to be involved in regulation of DNA methylation (**Supplemental Table 2**) and (3) genes in which sequence variants are associated with disorders with SGA as part of the phenotype (**Supplemental Table 3**). These genes will be referred to as 'targeted genes'.

DNA isolation

DNA was obtained from biopsies from the fetal side of the placenta near the umbilical cord insertion. DNA from parental blood samples and the cases were extracted using DNEasy Blood and Tissue Kit (Qiagen, CA, USA). DNA from the control samples was biopsied from the maternal site of the placenta and extracted according to the Gentra protocol (Qiagen, CA, USA). To minimize the risk of maternal blood contamination, placenta biopsies were washed in phosphate-buffered saline and stored in RNAlater. To verify whether no maternal DNA contamination has occurred, clustering of male samples and female samples was investigated by principal component analysis.

Array CGH

The array comparative genomic hybridization (array-CGH) analysis was performed using Agilent 180K oligo-array (Amadid 023363) (Agilent Technologies, Inc., Palo Alto, CA), with 13kb overall median probe spacing and GRCh37/hg19 browser. Standard methods were used for labelling and hybridization. Samples were hybridized against a pool of 40 healthy sex-matched human reference samples. Data were analyzed with Genomic Workbench 6.5 (Agilent, Santa Clara, CA and Cartagenia (BENCHlab CNV v5.0 (r6643)).

Genome-wide methylation array

Bisulfite conversion of genomic DNA was performed using EZ DNA Methylation Kit (Zymo Research, CA, USA). Converted DNA samples were randomized across one batch and hybridized on Infinium Human Methylation 450K BeadChip array (Illumina, Inc., CA, USA), carried out by a certified Illumina service provider (ServiceXS, Leiden, the Netherlands). The 450K BeadChip applies

both Infinium I and II assays and examines >450.000 CpG sites across the genome. Due to the bisulfite conversion, the array recognized methylated and unmethylated loci and expresses the degree of methylation in β-values, ranging from 0 (fully unmethylated) to 1 (fully methylated).

Quality control of the Illumina 450k assay was performed using MethylAid (10). Raw data were provided by ServiceXS and used for statistical analysis. A file containing the β-value methylation data including annotation was produced by GenomeStudio. Methylation data from GenomeStudio and sample phenotype data were exported to R statistical analysis environment (R version 2.15.2) (http://www.r-project.org), where a single sample analysis (11) was performed. This allows analysis of genome-wide methylation data in small sample sizes, where each case is individually compared to a control cohort. The method combines Illumina Methylation Analyzer (IMA) package (version 3.2.1) and Crawford-Howell t-test (11). The IMA package performs a basic quality control and pre-processes methylation data. Any CpG sites with missing values and samples with >75% CpG sites having a p-value >0.05, CpG sites where >75% samples have detection p-value $>1e^{-5}$, probes on the X and Y chromosomes and probes containing SNP(s) were removed. The β-values were converted to M-values by logit transformation (12). Quantile normalization was used to reduce unwanted technical variation across samples. Peak correction (13) was applied to correct differences between Infinium I and Infinium II type assays. As all cases and controls were hybridized on the same batch, no batch correction was required. Differences between pre-processed M-values of all single cases and the controls were determined using Crawford-Howell t-test.

Given the large number of significantly differentially methylated probes in our patients resulting from the single sample analysis, a script in Python (version 2.7) (https://www.python.org/) was used for further filtering of data. Probes with a β value difference of at least 20%, adjusted p-value <0.05 and a minimum of three differentially methylated probes within 2000 base pairs, allowing for reduction of false positive findings, were selected for hypermethylated and hypomethylated probes, respectively. Probes without gene annotation were removed from further analysis. Genes found to be hypermethylated and hypomethylated at the same time in the same patient were removed. First, genome-wide methylation patterns in SGA newborns were analyzed against previously reported literature (**Supplemental Table 1**). Second, other genes that were differentially methylated in >5

patients were selected.

In order to investigate the significance of the present methylation findings, we analysed the cohort of controls as if they were cases: the results in a single control was analysed against the remaining controls, and this was performed for each control. We carried out this analysis for the candidate genes as well as for the untargeted genes differentially methylated in >5 controls.

For the permutation analysis, the fraction of probes showing significant differential methylation (p-values below the threshold of 1.0e-2) were compared between the 50 candidate genes and randomly selected 50 genes within the same sample. This random selection was carried out 1000 times per sample and the mean value was generated for the comparison. The fraction of the probes having significant differential methylation is expected to be higher in the 50 candidate genes compared to the 50 randomly selected samples. If it is significantly higher, the permutation p-value would be less than 0.05, making them more likely candidates.

Exome sequencing

"Whole" Exome Sequencing (WES) was performed by BGI (Hong Kong). In total 41 samples were analyzed using Agilent SureSelect Human All Exon V5 (50M) kit and high throughput sequencing technology of Complete Genomics, at 100x coverage. The samples consisted of 10 trios from newborns with the lowest birthweights and their parents (SGA1, SGA3, SGA6, SGA15 – SGA21) and 11 singletons of the remaining newborns. For each sample BGI analyzed and provided reads, results of mappings, and basic bioinformatics analysis (including alignment and assessment, SNP and InDel calling, basic annotation and statistics, SNP validation). At our institution data were further annotated, including pathogenicity prediction data, allowing for subsequent filtering of variants. Variants were kept for further examination if mutation types (SO terms) with "high" and "moderate" impact (Ensembl Variation - Predicted data, ensemble.org), 1K genome minor allele frequency (MAF) <0.05, ExAC allele frequency <0.05, read depth ≥30, quality score ≥30. Variants with known non-pathogenic significance and a combined SIFT and PolyPhen prediction of "tolerated" and "benign" were discarded.

Subsequently, we checked variants in targeted genes known to cause a low birth weight

(**Supplemental Table 3)**, and determined the likelihood of pathogenicity. Ethnicity-specific MAF were obtained from 1000 Genome, ExAC and GO-ESP databases. Second, potential *de novo* variants were selected and verified in IGV (Integrative Genomics Viewer, Broad Institute) in the 10 patients of whom sequencing results from both newborn and parents were available. Lastly, homozygous and compound heterozygous mutations were analyzed. All variants in genes discussed in **Results** and **Discussion** have been validated by Sanger sequencing.

Results

Patients

All 21 SGA cases (SGA1 – SGA21) had a birth weight (BW) for gestational age (GA) below the 3.4th centile, 19 were below the 2.3nd centile and 14 patients below the 1st centile. **Table 1** shows other demographics of the study group and the control samples. Separate clustering of male cases and control samples from female samples was confirmed, indicating that no maternal DNA contamination was measured (**Supplemental Figure 1**).

Array CGH

The array-CGH yielded abnormalities in three patients. Patient SGA1 showed a mosaic trisomy of chromosome 16 in 70% of the cells (arr[19] 16p13.3q24(64,381-90,163,114)x2~3). Another mosaic imbalance, mosaic Turner syndrome, was seen in patient SGA11 (arr Xp22.33q28(61,091-155,009,479)x1~2). Patient SGA17 had a deletion of 11p13-p14.1 (arr[h19] 11p14.1p13(29,663,942-33,400,789)x1) causing WAGR syndrome (Wilms tumor, Aniridia, Genital anomalies and mental Retardation).

Genome-wide methylation

Quality control of the Illumina 450k assay showed no failed samples for bisulphite conversion, hybridization and overall methylation threshold. **Table 2** shows methylation changes in genes known to be aberrantly methylated in low birthweight newborns which we targeted first (see **Methods** and **Supplemental Table 1**). Differential methylation was seen in 12 patients of which nine had differential methylation in more than one gene. As patient SGA3 shows an extensively aberrant

methylation profile (**Figure 1**), the results for this patient are presented separately. Subsequently, all genes found in an untargeted study to be differentially methylated in five or more patients, were analysed (**Supplemental Table 4**), showing 28 hypermethylated genes and 6 hypomethylated genes. Analysis of our control cohort showed 45 differentially methylated genes present in >5 individuals. Out of the 34 genes resulting from the case analysis, in total 12 genes (35%) appeared to be the same genes. SGA3 showed differential methylation in 26 targeted genes known to be aberrantly methylated in low birthweight newborns (**Supplemental Table 5**). A possible explanation for this extensively disturbed methylation pattern in SGA3 would be an alteration in a gene known to be involved in (the regulation of) DNA-methylation (see **Methods** and **Supplemental Table 2**). Of these, four were hypermethylated and 11 hypomethylated (**Supplemental Table 6**). Additionally, WES data were checked for sequence variants in genes involved in regulating DNA-methylation (**Supplemental Table 6**), showing a heterozygous missense mutation in *MPHOSPH8* (p.Asp460Tyr). The same variant, with a known minor allele frequency (MAF) of 2-3% (rs75390100), was found in two other patients (SGA2 and SGA15). This high MAF excludes this variant to be the sole (Mendelian) cause of the IUGR but we cannot exclude that it contributes to a polygenic or multifactorial origin of the IUGR.

Due to an administrative error, three samples (SGA5, SGA9 and SGA12) could not be included in the genome-wide methylation analysis.

Analysis of the control cohort yielded a total of eight of the targeted genes (*CDKN1C, NPR3, NR3C1, FOXP1, H19, TBX15, WNT2 and ZIC1*) that were differentially methylated. Seven of those genes were also found in our case analysis with a comparable number of individuals and methylation in the same direction as the cases (or both hypermethylation and hypomethylation). The permutation analysis showed the probes within the 50 candidate genes resulting from our study (**Table 2** and **Supplemental Table 5)** had consistently higher fraction of probes below the threshold value compared to the randomly selected genes, except for four patients (**Supplemental Table 7**). In addition, 7/18 samples had P-values <0.05.

Exome sequencing

Exome sequencing without filtering yielded over 70,000 single nucleotide variants and ~5.000 InDel

variants in the 21 patients studied. After filtering (see Methods) we first evaluated sequence variants in genes that if mutated are known to be associated with disorders in which a low birthweight is part of the phenotype (**Supplemental Table 3**). This targeted analysis yielded potentially pathogenic heterozygous variants in 32 genes, one homozygous variant and two compound heterozygous variants (**Supplemental Table 8**). In this targeted gene panel, no *de novo* variants were identified in newborns of whom sequencing data of the parents were available. In patients in which no WES was performed in their parents, variants were sequenced by Sanger in parents and showed inheritance of all variants from one or both parents.

Second, *de novo* variants in untargeted genes were analyzed *in silico* (see Methods). Two *de* novo single nucleotide variants were predicted to be potentially pathogenic (**Supplemental Table 8**). Third, we analyzed all WES data for homozygous variants in untargeted genes, and found three homozygous missense mutations of potential interest (**Supplemental Table 8**). Lastly, we evaluated data for compound heterozygous mutations in untargeted genes and found one compound heterozygous variant (**Supplemental Table 8**). All variants described have been validated by Sanger sequencing. The recommendation of the American College of Medical Genetics and Genomics was followed in interpreting variants.

Discussion

In the present study we investigated 21 SGA newborns using a combination of array-CGH, genome wide methylation array and exome sequencing. In four patients (19%), we found a genetic abnormality that likely contributes to their low birthweight.

Three CNVs (14%) were detected in the present cohort, a relatively higher number compared to previous reports in patients with SGA or short stature (14). Mosaic trisomy 16 is known to lead to a high risk of prenatal abnormalities (15), frequently including SGA, and can thus be considered a valid explanation for SGA in patient SGA1. Patient SGA11 showed mosaicism for monosomy X. About 50% of individuals with Turner syndrome have a mosaic karyotype (16), and it typically includes a low birth weight. In patient SGA17 an 11p14.1-p13 deletion, as seen in WAGR syndrome,

was identified. This syndrome features reduced intrauterine growth as a known phenotype (17).

As hypothesized in advance, methylation disturbances in several genes known to be aberrantly methylated in low birthweight newborns, were found. In general, more hypermethylation than hypomethylation was found in the present cohort. A methylation abnormality potentially involved in SGA was detected in 13 patients, of which five showed differential methylation in several imprinted genes from the 11p15.5 imprinted region associated with fetal growth restriction, including *CDKN1C, KCNQ1, IGF2AS, INS* and *IGF2* (18). However, when each control was analyzed versus the remaining controls, similar to the case analyses, similar differential methylation of the majority of these genes in control individuals were found. This suggests that these findings are not significant as a Mendelian cause for the disturbed intrauterine growth. On the other, hand, the control versus controls analyses also showed that the differential methylation of *KCNQ1, IGF2AS, INS* and *IGF2* were solely differently methylated in the case series, which increases the likelihood (but does not prove) these contribute to fetal growth restriction. In addition, probes within the candidate genes had consistently higher fraction of probes below the threshold value compared to the randomly selected genes in the majority of cases, as shown in the permutation analysis. Seven out of 19 samples showed a permutation p-value <0.05, when only one would be expected by chance. Therefore, these results were consistent with the overall findings of this study: the 50 candidate genes showed differences in a small proportion of the patients, some of whom also carried potential genetic variants as well.

Methylation disturbances in untargeted genes were considered if the disturbance was present in at least 5 patients. Such changes were detected in 34 genes, six of these (*PIK3R1, DIXDC1, ESRRG, TBX15, GGT1* and *FGF8*) appeared to be of specific interest, in view of their known functions. However, given thatanalysis of the control cohort yielded a similar amount of genes including overlap in 12 genes between the cases and controls, these results are unlikely to be significant, at least when considering them as Mendelian causes for the FGR.

Although promoter methylation is generally associated with reduced gene expression and methylation of a gene itself typically with increased gene expression (19), DNA methylation differs

pre- and postnatally both quantitatively and in terms of CpG versus non-CpG methylation in utero. Therefore, hypermethylation does not necessarily lead to decreased transcription and hypomethylation to increased transcription Furthermore, the location of the hypermethylated or hypomethylated region within the gene can be crucial in relation to the consequences for expression: for example, methylation of a promoter can cause a reduced gene expression while methylation of exons to increased gene expression. Therefore, understanding the consequences of the direction of differential methylation currently remains uncertain. RNA expression studies should provide insight in the consequences of the hyper- and hypomethylation detected in the present candidate genes.

In contrast to the generally more frequent hypermethylation profile in the present cohort, patient SGA3 showed a predominant hypomethylation pattern and an extensively disturbed methylation profile. First, an external epigenetic influence, such as tobacco smoke or infectious pathogens, could be the cause of this observation (20). The essential hypertension of SGA3's mother may be of importance in this respect. Second, a mutation in genes regulating DNA methylation, such as the *DNMTs* and *TETs*(21), could theoretically cause widespread DNA methylation disturbances. Also maternal mutations in so called 'maternal-effect genes' such as *NLRP5, NLRP7 I* and *KHDC3L(22)* can cause multi-locus imprinting disturbances in their offspring, usually resulting in hypomethylation at multiple loci and seen primarily in female offspring (23). We were unable to detect any of such mutations in SGA3. Lastly, disturbed methylation of 15 genes known to be involved in (regulation of) DNA methylation was present in SGA3. Especially the abnormal methylation of *DNMT1, DNMT3B, TET1, UHRF1* and *ZFP57* may be of interest as abnormal methylation of one of these may have had a subsequent extensive effect.

The evaluation of exome sequencing, targeted for genes associated with disorders in which a low birthweight is part of the phenotype, uncovered 37 sequence variants in 35 genes. In evaluating results we took into account the variability of pattern of inheritance of variants in a single gene, and the possibility that variants may not act in a Mendelian manner but can also act in a polygenic or multifactorial manner. Each reported finding may be involved in fetal growth restriction since if mutated, these genes can cause malformation syndromes, skeletal dysplasias and endocrine disorders. Our analyses showed several sequence variants that are plausible candidates for causality,

while the majority remain of uncertain significance. One strong candidate is the splice acceptor variant in *SOS1* (c.3347-1G>A). This variant has been described previously in patients with Noonan-syndrome and IUGR (24). Given the earlier reported patients and the low MAF, this variant deserves further investigation to confirm its potential pathogenic nature. When assessing the different sequence variants for potential pathogenicity, it is important to stress that in our opinion it is likely that in most patients the cause of FGR will not be monogenic but rather polygenic. This implies that there will be changes in either one or more gene and/or in the methylation pattern, such that each individually will not cause a major health problem (and will be present in controls as well). In this multifactorial model, combinations of a series of such changes will lead to disturbed growth. The size of the present study is too limited to reveal complex interactions, however, to provide comparison to future studies, sequence and differential methylation data is available in the Supplemental Tables.

We found *de novo* mutations in two untargeted genes. *MTUS1* is a tumor suppressor gene controlling cell proliferation, however no function interfering with fetal growth is known so the meaning of this variant remains uncertain. The protein encoded by *LZTS2* acts as a tumor suppressor and is involved in regulating embryonic development by the Wnt signaling pathway. Homozygous mutations were found in four genes, all of uncertain significance. One patient was homozygous for the p.Val316Ala variant in *MTHFD1*. *MTHFD1* is important for folate metabolism and embryonic development and a mutation in this gene has been associated with fetal hypotrophy. Given the contradictory classifications by the prediction programs, the significance of this variant is uncertain. Compound heterozygous variants of interest were found in two targeted genes and one untargeted gene, all likely benign or of uncertain significance.

Our results do not indicate the presence of a single, unifying theme explaining the dysregulation of fetal growth, and confirm previous findings that growth *in utero* is influenced by a large number of genes. We demonstrate that there is no predominant type of genetic abnormality present in SGA newborns: copy number variations, methylation disturbances and sequence variants may all contribute in part to the phenotype. In 19 patients, combinations of a CNV, (multiple) sequence variants and (multiple) methylation disturbances are present. Each of these will require

detailed and sophisticated investigations to better understand their significance. Our results mirror a similar study in children with postnatal growth failure (25), using a similar approach in evaluating variants in genes known to be associated with short stature as well as studying variants in other, untargeted genes (25). The latter authors highlight the multitude of genetic causes for short stature and the complexity of interpretation of variants and their pathogenicity, which resembles the observations in the present study.

We acknowledge limitations of the present study. The size of our cohort is small and the power to draw general conclusions is limited. The genetic heterogeneity within the present cohort also appears high. We therefore used an individual-based data analysis approach for the methylation study to enable suitable data analysis. Using filtering strategies based on population allele frequencies as in the present study limits the ability to determine the pathogenicity of WES variants, since such data lack individual phenotypic data. Ideally, for future studies our AGA newborns should be sequenced to serve as a control population to allow determining pathogenicity of combinations of variants detected by exome sequencing. Furthermore, we had no access to clinical follow-up data of the presently studied cohort. The large number of variants of uncertain significance inhibits investigating each individual variant extensively; ideally, each variant would require a separate, detailed study.

We conclude that copy number variations, methylation disturbances and sequence variants may all contribute in part to prenatal growth failure. This study shows genetic disturbances in SGA are complex and likely polygenic. The results of these studies in individual patients may have important consequences for patient care and counselling of patients and their families. Further research is still required to determine whether such genetic workup can become an effective diagnostic approach in SGA newborns.

Acknowledgments

We would like to thank Jessica Glebbeek, Patricia Schmidt, Sander de Jong, Femke van Sinderen and Jet Bliek for their help in the laboratory. The present study has been made possible by grants of the

following institutions: Tergooi Grant for Support of Scientific Research, KNAW Ter Meulen Fund, Jo Kolk Study Fund, ZonMW Rare Disease Network Grant, the Baby Bio Bank, Wellbeing of Women and Biomedical Research Centre Grant. All sponsors are gratefully acknowledged. Sponsors had no involvement in any stage of the study design, data collection, data analyses or interpretation of the data or decision to publish study results.

References

1. Clayton PE, Cianfarani S, Czernichow P, Johannsson G, Rapaport R, Rogol A. Management of the child born small for gestational age through to adulthood: a consensus statement of the International Societies of Pediatric Endocrinology and the Growth Hormone Research Society. J Clin Endocrinol Metab 2007; 92:804-810

2. ACOG. Intrauterine growth restriction - Clinical management guidelines for obstetrician–gynecologists. Int J Gynaecol Obstet 2001; 72:85-96

3. Alexander GR, Kogan M, Bader D, Carlo W, Allen M, Mor J. US birth weight/gestational age-specific neonatal mortality: 1995-1997 rates for whites, hispanics, and blacks. Pediatrics 2003; 111:e61-66

4. Barker DJ. The developmental origins of chronic adult disease. Acta paediatrica (Oslo, Norway : 1992) Supplement 2004; 93:26-33

5. Lunde A, Melve KK, Gjessing HK, Skjaerven R, Irgens LM. Genetic and environmental influences on birth weight, birth length, head circumference, and gestational age by use of population-based parent-offspring data. American journal of epidemiology 2007; 165:734-741

6. Oxford Medical Databases: London Dysmorphology and Dysmorphology Photo Library Version 3.0. 2001 Oxford University Press.

7. Moore GE, Ishida M, Demetriou C, Al-Olabi L, Leon LJ, Thomas AC, Abu-Amero S, Frost JM, Stafford JL, Chaoqun Y, Duncan AJ, Baigel R, Brimioulle M, Iglesias-Platas I, Apostolidou S, Aggarwal R, Whittaker JC, Syngelaki A, Nicolaides KH, Regan L, Monk D, Stanier P. The role and interaction of imprinted genes in human fetal growth. Philosophical transactions of the Royal Society of London Series B, Biological sciences 2015; 370:20140074

8. Freeman JV, Cole TJ, Chinn S, Jones PR, White EM, Preece MA. Cross sectional stature and weight reference curves for the UK, 1990. Arch Dis Child 1995; 73:17-24

9. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. An update of the Swedish reference standards for weight, length and head circumference at birth for given gestational age (1977-1981). Acta paediatrica Scandinavica 1991; 80:756-762

10. van Iterson M, Tobi EW, Slieker RC, den Hollander W, Luijk R, Slagboom PE, Heijmans BT. MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics 2014; 30:3435-3437

11. Rezwan FI, Docherty LE, Poole RL, Lockett GA, Arshad SH, Holloway JW, Temple IK, Mackay DJ. A statistical method for single sample analysis of HumanMethylation450 array data: genome-wide methylation analysis of patients with imprinting disorders. Clinical epigenetics 2015; 7:48

12. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, Lin SM. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC bioinformatics 2010; 11:587

13. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics 2011; 3:771-784

14. Wit JM, van Duyvenvoorde HA, van Klinken JB, Caliebe J, Bosch CAJ, Lui JC, Gijsbers ACJ, Bakker E, Breuning MH, Oostdijk W, Losekoot M, Baron J, Binder G, Ranke MB, Ruivenkamp CAL. Copy Number Variants in Short Children Born Small for Gestational Age. Horm Res Paediatr 2014; 82:310-318

15. Chareonsirisuthigul T, Worawichawong S, Parinayok R, Promsonthi P, Rerkamnuaychoke B. Intrauterine growth retardation fetus with trisomy 16 mosaicism. Case reports in genetics 2014; 2014:739513

16. Hennekam RCM, Krantz ID, Allanson JE. "Turner syndrome" Gorlin's syndromes of the head and neck. Oxford: Oxford University Press; 2010.

17. Tezcan B, Rich P, Bhide A. Prenatal Diagnosis of WAGR Syndrome. Case reports in obstetrics and gynecology 2015; 2015:928585

18. Cordeiro A, Neto AP, Carvalho F, Ramalho C, Doria S. Relevance of genomic imprinting in intrauterine human growth expression of CDKN1C, H19, IGF2, KCNQ1 and PHLDA2 imprinted genes. Journal of assisted reproduction and genetics 2014; 31:1361-1368

19. Wagner JR, Busche S, Ge B, Kwan T, Pastinen T, Blanchette M. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. Genome biology 2014; 15:R37

20. Ho SM, Johnson A, Tarapore P, Janakiram V, Zhang X, Leung YK. Environmental epigenetics and its implication on disease risk and health outcomes. ILAR journal / National Research Council, Institute of Laboratory Animal Resources 2012; 53:289-305

21. Hamidi T, Singh AK, Chen T. Genetic alterations of DNA methylation machinery in human diseases. Epigenomics 2015; 7:247-265

22. Docherty LE, Rezwan FI, Poole RL, Turner CL, Kivuva E, Maher ER, Smithson SF, Hamilton-Shield JP, Patalan M, Gizewska M, Peregud-Pogorzelski J, Beygo J, Buiting K, Horsthemke B, Soellner L, Begemann M, Eggermann T, Baple E, Mansour S, Temple IK, Mackay DJ. Mutations in NLRP5 are associated with reproductive wastage and multilocus imprinting disorders in humans. Nat Commun 2015; 6:8086

23. Sano S, Matsubara K, Nagasaki K, Kikuchi T, Nakabayashi K, Hata K, Fukami M, Kagami M, Ogata T. Beckwith-Wiedemann syndrome and pseudohypoparathyroidism type Ib in a patient with multilocus imprinting disturbance: a female-dominant phenomenon? J Hum Genet 2016;

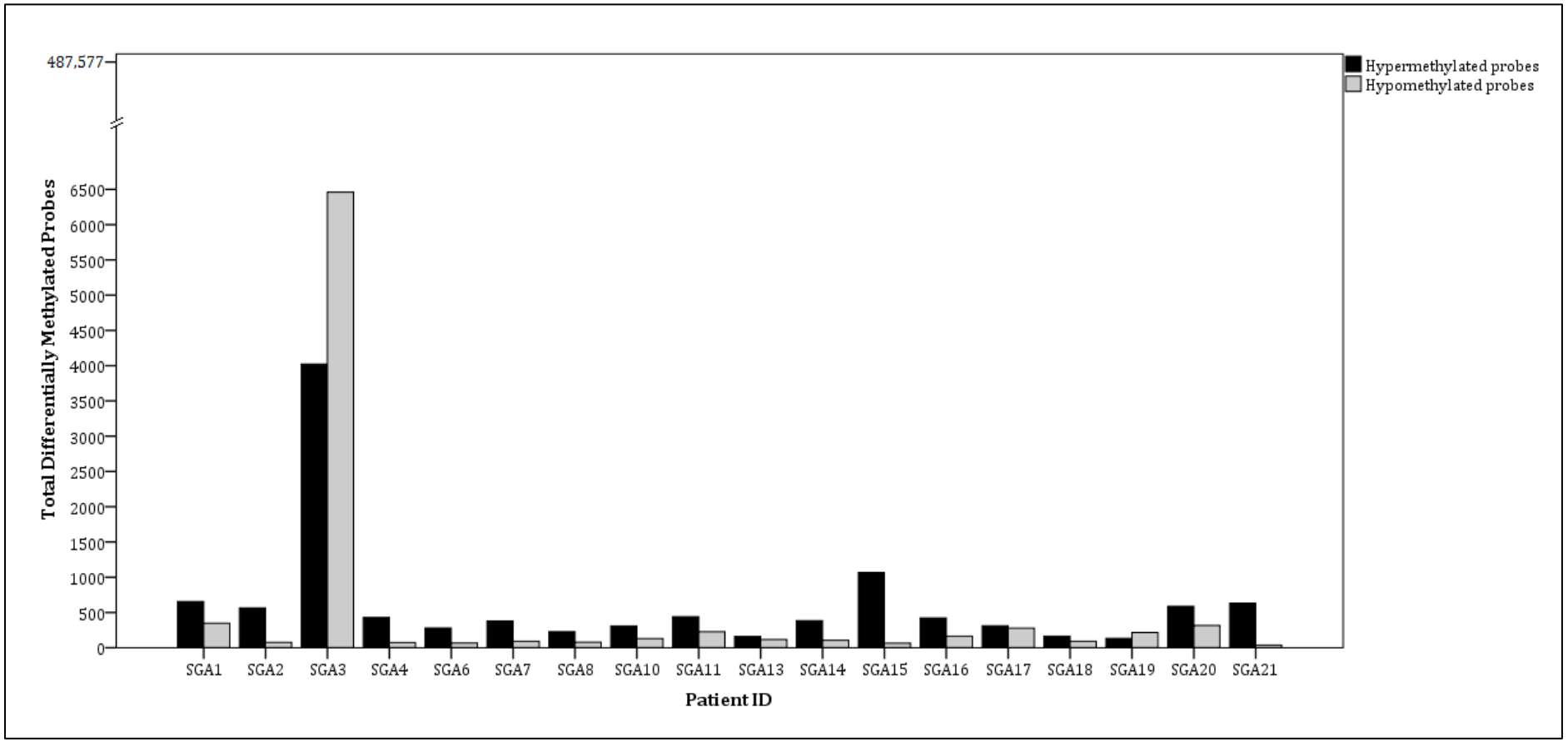
24. Justino A, Dias P, Joao Pina M, Sousa S, Cirnes L, Berta Sousa A, Carlos Machado J, Costa JL. Comprehensive massive parallel DNA sequencing strategy for the genetic diagnosis of the neuro-cardio-facio-cutaneous syndromes. European journal of human genetics : EJHG 2015; 23:347-353

25. Wang SR, Carmichael H, Andrew SF, Miller TC, Moon JE, Derr MA, Hwa V, Hirschhorn JN, Dauber A. Large-scale pooled next-generation sequencing of 1077 genes to identify genetic causes of short stature. J Clin Endocrinol Metab 2013; 98:E1428-1437

Figure 1. Number of Differentially Methylated Probes per Patient

Legend Figure 1.
Total number of differentially methylated probes per patient out of 485,577 interrogated probes, after single case analysis and further probe filtering (see **Methods**). An extensively disturbed methylation profile is evident in patient SGA3, and patient SGA15 has an increased number of hypermethylated probes in comparison to the other patients

Figure 1. Number of Differentially Methylated Probes per Patient



Legend Figure 1. Total number of differentially methylated probes per patient out of 485,577 interrogated probes, after single case analysis and further probe filtering (see **Methods**). An extensively disturbed methylation profile is evident in patient SGA3, and patient SGA15 has an increased number of hypermethylated probes in comparison to the other patients.

Table 1. Demographics of 21 SGA Newborns and 24 Controls Appropriate for Gestational Age

Patient ID	Gender	GA	BW (grams)	BW (centile)	BW (SDS)	Ethnicity	Mode of delivery
Cases							
SGA1	Female	33.00	1220	0.41	-2.64	Caucasian	Caesarean section
SGA2	Female	38.00	1980	0.51	-2.57	African	Vaginal
SGA3	Female	33.71	640	4.7E-5	-4.91	South American	Caesarean section
SGA4	Female	39.00	2435	3.36	-1.83	Caribbean	Caesarean section
SGA5	Female	34.00	1350	0.59	-2.52	Asian	Caesarean section
SGA6	Female	39.57	2120	0.22	-2.85	Caucasian	Vaginal
SGA7	Male	38.00	2080	0.69	-2.46	South American	Vaginal
SGA8	Male	38.00	2140	1.04	-2.31	Caucasian	Vaginal
SGA9	Male	34.43	1543	1.04	-2.31	Caucasian	Caesarean section
SGA10	Male	39.57	2320	0.62	-2.50	Caucasian	Vaginal
SGA11	Female	38.57	2180	1.10	-2.29	African	Caesarean section
SGA12	Female	39.00	2385	2.56	-1.95	African	Caesarean section
SGA13	Male	38.57	2280	1.36	-2.21	Asian	Caesarean section
SGA14	Female	37.14	2017	1.83	-2.09	Caribbean	Vaginal
SGA15	Female	31.71	474	3.14E-4	-4.52	Caucasian	Caesarean section
SGA16	Male	39.00	2090	0.24	-2.82	African	Caesarean section
SGA17	Male	22.00	236	0.13	-3.00	Caucasian	Termination of pregnancy
SGA18	Male	36.00	1600	0.21	-2.86	Caucasian	Caesarean section
SGA19	Male	37.00	1782	0.26	-2.80	Caucasian	Caesarean section
SGA20	Male	40.00	1874	0.01	-3.69	Caucasian	Vaginal
SGA21	Male	40.00	2220	0.20	-2.88	Caucasian	Vaginal
Mean ± SD	-	36.49 ± 4.14	1760 ± 640	0.78 ± 0.88	-2.76 ± 0.77	-	-
Controls							
Mean ± SD	-	37.48 ± 4.10	2953 ± 926	53.83 ± 15.51	0.10 ± 0.42	-	-

GA=gestational age; BW=birth weight; SDS=standard deviation score

Table 2. Differential Methylation in Genes Known to Be Aberrantly Methylated in Low Birthweight Newborns

Patient ID	Gene	Chromosome (MapInfo)	Control β-value (mean)	Case β-value (mean)	No. of probes	Main gene function(s) and influence(s) on fetal growth[a]	Comparison with control cases	
Hypermethylation							**Direction of methyation**	**No. of controls**
SGA11 SGA15	*CDKN1C*	11 (2905931-2907008) 11 (2906667-2907073)	0.14 0.21	0.41 0.48	5 4	Imprinted gene in 11p15.5 region, highly expressed in placenta. Upregulation associated with IUGR placentas, loss of function associated with Beckwith-Wiedemann syndrome, gain of function with Silver-Russel syndrome*	Hypermethylation	1
SGA4	*FGF14*	13 (103052362-103052943)	0.16	0.44	3	Hypomethylation associated with SGA or FGR	-	-
SGA13	*GNAS; GNASAS*	20 (57414162-57414539)	0.62	0.83	3	Hypomethylation of GNASAS associated with SGA. Decreased expression of GNAS observed in IUGR placentas	-	-
SGA1 SGA7	*FOXP1*	3 (71631050-71631744) 3 (71631050-71631744)	0.09 0.09	0.45 0.46	4 4	Increased methylation associated with FGR	Hypermethylation	1
SGA14 SGA20	*NPR3*	5 (32710614-32711429) 5 (32710231-32711517)	0.24 0.12	0.49 0.39	4 6	Hypermethylation associated with FGR	Hypermethylation	3
SGA14	*NR3C1*	5 (142784522-142785258)	0.22	0.47	3	Differential methylation in this glucocorticoid receptor in placenta correlated with birth weight	Hypermethylation	3
SGA11 SGA14 SGA15 SGA17 SGA20 SGA21	*TBX15*	1 (119530600-119530702) 1 (119530600-119531093) 1 (119530600-119530702) 1 (119530048-119530932) 1 (119530600-119530702) 1 (119530600-119531093)	0.28 0.31 0.28 0.35 0.28 0.30	0.58 0.55 0.64 0.58 0.57 0.61	3 3 3 4 3 4	Promotor hypomethylation leads to TBX15 decrease in FGR placentas	Both	3**
SGA7 SGA11 SGA13 SGA16	*WNT2*	7 (116963193-116963502) 7 (116962950-116964012) 7 (116962950-116963502) 7 (116962950-116963502)	0.18 0.19 0.17 0.17	0.52 0.51 0.48 0.48	5 7 6 6	WNT2 promoter methylation in placenta is associated with low birthweight	Both	6**
SGA10 SGA15 SGA21	*ZIC1; ZIC4*	3 (147125714-147127662) 3 (147125712-147126206) 3 (147126763-147127662)	0.30 0.43 0.21	0.58 0.66 0.57	5 6 6	Decreased methylation associated with SGA or FGR	Both	6**
Hypomethylation								
SGA17	*IGF2AS; INS-IGF2; IGF2*	11 (2162406-2162616)	0.44	0.17	5	*IGF2* is imprinted and highly expressed in placenta, hypomethylation of *H19/IGF2* control region is associated with FGR. *INS-IGF2* involved in growth and metabolism. *IGF2AS* isimprinted and expressed in antisense to *IGF2*	-	-
SGA13	*KCNQ1; KCNQ1OT1*	11(2721207-2721383)	0.49	0.24	4	Upregulated *KCNQ1* and loss of *KCNQ1OT1* associated with IUGR; genetic variants of *KCNQ1* associated with Beckwith-Wiedemann syndrome	-	-
SGA1	*TBX15*	1 (119526060-119527377)	0.68	0.42	4	Promotor hypomethylation leads to TBX15 decrease in FGR placentas	Both	1**
SGA17	*WNT2*	7 (116964012-116964802)	0.35	0.11	4	WNT2 promoter methylation in placenta is associated with low birthweight	Both	2**

For references see **Supplemental Table 1**; *Clinical diagnosis uncertain due to unavailable detailed phenotyping, ** controls with methylation in same direction as case.