



(51) International Patent Classification:

A61K 38/19 (2006.01) C12Q 1/6883 (2018.01)
C07K 16/24 (2006.01) C12Q 1/6886 (2018.01)

(21) International Application Number:

PCT/US2018/026588

(22) International Filing Date:

06 April 2018 (06.04.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/482,362 06 April 2017 (06.04.2017) US
62/596,704 08 December 2017 (08.12.2017) US

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 1111 Franklin Street, Twelfth Floor, Oakland, California 94607-5200 (US).

(72) Inventor: FEJZO, Marlena S.; c/o UCLA Technology Development Group, 10889 Wilshire Boulevard, Suite 920, Los Angeles, California 90024 (US).

(74) Agent: CANADY, Karen S. et al.; Canady + Lortz LLP, 3701 Wilshire Blvd, Suite 508, Los Angeles, California 90010 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(54) Title: PREDICTING, DIAGNOSING, AND TREATING NAUSEA AND VOMITING OF PREGNANCY

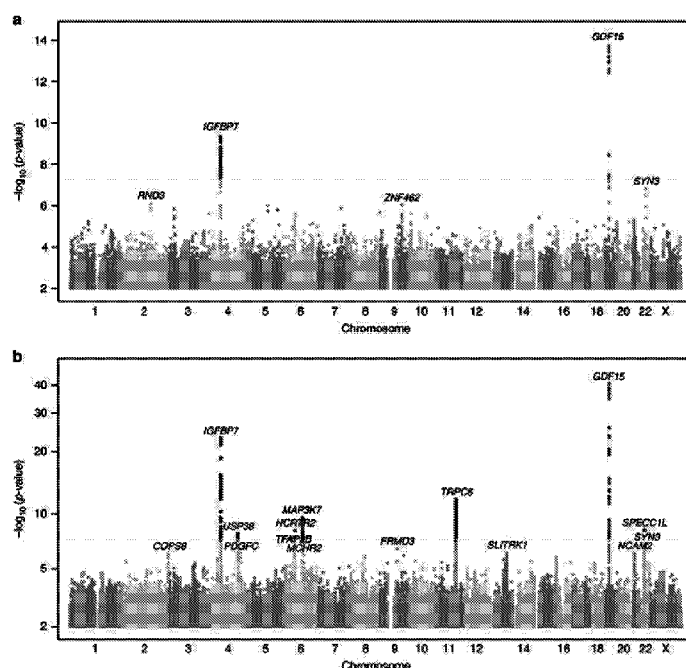


FIGURE 1

(57) Abstract: Assays for identifying patients at risk for nausea and vomiting in pregnancy (NVP) and hyperemesis gravidarum (HG), as well as methods for diagnosis and/or treatment of HG and NVP. Described herein is a method of analyzing genetic expression in a sample obtained from a subject. The method comprises obtaining a sample from the subject; and measuring the level of expression product of at least one of the HG outlook genes. Expression products to be detected include proteins and nucleic acid sequences. Antibodies and/or probes can be used to detect HG outlook gene expression products, such as GDF15 and IGFBP7, in samples, including serum, urine, and blood, obtained from patients. Treatment for NVP and/or HG can be selected on the basis of the assay.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

PREDICTING, DIAGNOSING, AND TREATING NAUSEA AND VOMITING OF PREGNANCY

[0001] This application claims the benefit of United States provisional patent applications number 62/482,362, filed April 6, 2017, and number 62/596,704, filed December 8, 2017, the entire contents of each of which are incorporated herein by reference.

5

BACKGROUND OF THE INVENTION

[0002] Nausea and vomiting of pregnancy (NVP) affects 50-90% of pregnant women. As many as 18% of pregnant women take medication to treat NVP. Hyperemesis Gravidarum (HG) is the most severe form and occurs in 0.3-2% of pregnancies. Its clinical presentation includes severe intractable vomiting, often associated with dehydration, weight loss (> 5% pre-pregnancy weight), ketonuria, metabolic alkalosis and electrolyte disturbances. There is surprisingly little research on this disease that may have killed famous author Charlotte Brontë, was the leading cause of maternal death until intravenous hydration was introduced in the 1950s, and to this day is the second leading cause of hospitalization of pregnancy after preterm labor.

[0003] Despite the prevalence of nausea and vomiting in pregnancy (NVP) and the gravity of the most severe form, hyperemesis gravidarum (HG), decades of research have failed to identify the cause, and a safe and effective treatment has yet to be found. While having no nausea and vomiting has been linked to miscarriage, having severe nausea and vomiting (HG) is also associated with poor outcomes including preterm birth, neurodevelopmental delay, and vitamin K deficient embryopathy. A variety of potential causative factors have been investigated, but the etiology remains unknown. Evidence for a genetic predisposition is provided by classic twin studies of Norwegian, Spanish, and Finnish cohorts that all support a genetic component to NVP. Heritability estimates for presence of NVP are as high as 73%. Siblings and mothers of patients affected with NVP or HG are more likely to be affected than siblings and mothers of unaffected individuals. Sisters of women with HG have a 17-fold increased risk of having HG themselves, and there is a >27-fold increased risk of mother-daughter recurrence when a mother has HG with two daughters.

[0004] There are no diagnostic tests that can predict individual risk of NVP and HG. There remains a need for a practical and accessible diagnostic tool for identifying patients at risk of NVP and HG, facilitating timely therapeutic intervention.

30

SUMMARY OF THE INVENTION

[0005] The assay reagents and methods described herein meet these needs and others by providing biomarkers for nausea and vomiting in pregnancy (NVP) and hyperemesis gravidarum

(HG). An assay as described herein is extremely useful to practicing clinicians to identify patients at risk for NVP and HG. The invention thus provides methods for diagnosis and/or treatment of HG and NVP.

[0006] Described herein is a method of analyzing genetic expression in a sample obtained from a subject. In one embodiment, the method comprises obtaining a sample from the subject; and measuring the level of expression product of at least one of the hyperemesis gravidarum (HG) outlook genes identified in Table 1 in the sample. As used herein, "expression product" includes both the genotypic expression (e.g., presence of a variant) and the phenotypic expression (e.g., RNA or protein expression indicative of a genetic variant). Thus, in some embodiments, measuring an expression product comprises detection of a genetic variant. The method can further comprise assigning a HG outlook score to the measured amount of expression for each of the HG outlook genes, wherein the score is between 0 and 1; and calculating a HG status score, wherein the status score is equal to a combination of the outlook scores assigned. In one embodiment, the combination is an average outlook score. In other embodiments, the combination is a total of HG status scores for each outlook gene measured or detected. The combination can be calculated or determined in other ways that result in an HG status score that reflects the number and/or quality of genetic variants detected. In one embodiment, the presence of a genetic variation associated with HG is assigned a score of 1, and the absence of the genetic variation is assigned a score of 0. Those skilled in the art will appreciate other means of assigning and combining HG outlook scores.

[0007] Also provided are methods and reagents useful for improving treatment efficacy for HG. For example, patients can be tested for the presence of a risk allele that is predictive of ability or failure to respond to medications, such as ondansetron, promethazine, and metoclopramide. As shown herein, failure to respond to ondansetron is linked to the *IGFBP7* HG-risk allele; ineffectiveness of promethazine is also linked to the *IGFBP7* HG-risk allele; and response to metoclopramide is linked to the *GDF15* HG-risk allele. In one embodiment, the invention provides a method of treating HG, the method comprising treating a patient having HG with ondansetron or promethazine if the patient is lacking the *IGFBP7* HG-risk allele, and treating a patient with metoclopramide if the patient is positive for the *GDF15* HG-risk allele.

[0008] Due to the increased efficacy of ondansetron relative to promethazine, which, in turn, is more effective than metoclopramide, a treating physician may opt to initiate treatment with ondansetron regardless of phenotype. Other factors, such as price differences, insurance coverage, and safety data may also influence which drug is used first. Based on effectiveness, in most instances, a physician would not use metoclopramide before any of the other drugs, unless the patient is homozygous for the *GDF15* HG-risk allele. In the latter case,

metoclopramide is predicted to be almost as effective as promethazine in patients with at least one of the *IGFBP7* risk alleles.

[0009] In one embodiment, the method comprises measuring the expression product of an HG outlook gene, wherein the HG outlook gene is selected from: GDF15, IGFBP7, PGR, and RYR2. In a typical embodiment, at least two of the HG outlook genes of Table 1 are measured. In some embodiments, three, four, five, six, seven, eight, nine, ten, or more HG outlook genes are measured, including those listed in Table 1, and other HG outlook genes, such as RYR2.

[0010] In one embodiment, the method comprises measuring the level of GDF15 in a sample obtained from a subject. If the sample shows levels of GDF15 greater than about 10 ng/ml, the subject is likely to have or develop HG. Optionally, the method further comprises measuring the level of IGFBP7 in a sample obtained from the subject, using either the same or another sample. If the sample shows levels of IGFBP7 greater than about 60 ng/ml, the subject is likely to have or develop HG. Subjects exhibiting greater than 10 ng/ml GDF15 and greater than 60 ng/ml IGFBP7 have a 36-fold greater risk of being hospitalized for HG, and are candidates for treatment with medication, such as ondansetron, promethazine, or metoclopramide. In some embodiments, a subject showing greater than about 8 ng/ml GDF15 and/or greater than about 50 ng/ml IGFBP7 are candidates for treatment with medication. Those skilled in the art will know to adjust these cut off levels based on the assays employed and comparison to control or normalization values. For example, the above values are based on data collected from subjects at 12 weeks gestation, and use of a different assay, or collection of samples at a different point in gestation, may involve adjustment to these values. The forementioned immunoassay is performed, in some embodiments, using detection of one or both of GDF15 and IGFBP7, optionally, in combination with an additional marker. In some embodiments, the assay is part of a multi-marker assay that includes reagents for detecting a panel of up to 20 markers. In some embodiments, the panel consists of no more than 15 markers, or up to 10 markers, and in some embodiments, up to 5 markers are used in combination.

[0011] The sample is a biological patient sample, representative examples of which include, a urine sample, a blood sample (including serum), or a saliva sample. In a typical embodiment, the sample is tested for genotype and/or associated phenotype (RNA, protein expression levels) for a minimum of 2 alleles in Table 1 (or for an allele with linkage disequilibrium to the allele in Table 1). Treatment for NVP or HG can be administered to the subject based on the assigned score for each test and/or combination of tests. In some embodiments, treatment is administered to the subject based on detection of protein levels, such as greater than 10 ng/ml GDF15 and/or greater than 60 ng/ml IGFBP7. Those skilled in the art will appreciate that the

cut-off values can be adjusted in accordance with specific adaptations of the methods described herein.

[0012] Thus, in some embodiments, the measuring comprises contacting the sample obtained from the subject with reagents for amplifying polynucleotides. In some embodiments, the level of expression is measured via hybridization of probes that specifically bind polynucleotides encoding the expression product. In yet other embodiments, the level of expression is measured via binding of antibodies that specifically bind the expression product. Standard immunoassay reagents can be used to detect proteins. Representative immunoassays include enzyme immunoassays (including ELISA), microarray assays, nanosurface assays, western blotting, alphaLISA, radioimmunoassay, competitive binding assays, and automated instrumentation-based assays.

[0013] In one embodiment, the HG outlook gene is GDF15, IGFBP7, PGR, SYN3, MMADHC, TMEM38B, HCRTR2, GFRAL, RYR2 and/or PKHD1. The expression can, in a representative embodiment, be measured via detection of the G allele at rs16982345 (or any alleles in a shared haplotype (in linkage disequilibrium) with the G allele at rs16982345) of GDF15. In another embodiment, the HG outlook gene is a combination of two, three, or four of GDF15, IGFBP7, PGR, GFRAL, and RYR2. Examples of such combinations include, but are not limited to, GDF15 and IGFBP7; GDF15 and PGR; GDF15 and RYR2; GDF15 and GFRAL; IGFBP7 and PGR; IGFBP7 and RYR2; IGFBP7 and GFRAL; PGR and RYR2; PGR and GFRAL; GFRAL and RYR2; GDF15, IGFBP7, and GFRAL; GDF15, PGR, and GFRAL; IGFBP7, PGR, and GFRAL; GDF15, IGFBP7, PGR, and GFRAL; GDF15, PGR, and GFRAL; IGFBP7, PGR, and GFRAL; as well as GDF15, IGFBP7, and GFRAL.

[0014] In some embodiments, the method further comprises treating the subject for nausea/vomiting of pregnancy (NVP) or hyperemesis gravidum (HG) if the HG status score is greater than zero. In one embodiment, the gene is GDF15 and the treatment comprises administering an anti-GDF15 agent to the subject. In another embodiment the gene is PGR and the treatment comprises administering an anti-PGR and/or anti-progesterone agent. In another embodiment, the gene is MMADHC and the treatment comprises administering vitamin B12 to the subject. The method can also be used to treat related diseases, including cachexia and low appetite, while the reverse treatment would be applied for detection of the other allele for treating fetal loss and/or overactive appetite. In such a reverse treatment, the treatment comprises administering GDF 15 if the genotype is A/A at rs16982345 in GDF15, and/or progesterone depending on the genotype at the PGR locus.

[0015] Also described is a method of treating a subject for NVP or HG. In some embodiments, the method comprises performing the analyzing described herein; and administering a treatment for NVP or HG to the subject if the analyzing results in a status score that is greater than zero. In some embodiments, the treatment comprises an inhibitor of GDF15, GFRAL, or PGR. In one
5 embodiment, the invention provides a method of treating a subject for HG, wherein the method comprises analyzing a sample obtained from the subject for presence of one or more of the genotype combinations identified in Examples 3 and 5 hereinbelow as associated with HG at a significance of $P \leq 0.05$; and administering an inhibitor of a corresponding gene if the patient is positive for the significant genotype combination. Also provided is a method of treating a subject
10 with an inhibitor of GDF15 or GFRAL, wherein the method comprises assaying a sample obtained from the subject for the presence of a variant of GDF15 or GFRAL as described herein, and administering the inhibitor to the subject whose sample contains the corresponding variant.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] Figure 1 Genome-wide association scans for nausea and vomiting of pregnancy. The Manhattan plot shows distribution of association test statistics vs. genomic position for (a) SCAN1 (binary phenotype), and (b) SCAN2 (ordinal phenotype). The Manhattan plot shows distribution of association test statistics versus genomic position. Chromosomes are arranged along the X-axis. Log10-scaled p-values are shown on the Y-axis. The loci with positions with
15 $p < 5 \times 10^{-8}$ are shown in red and the loci with $p < 10^{-6}$ are labeled with names of nearest genes.
20

[0017] Figure 2 is a scatterplot of serum levels of GDF15 (ng/ml), at 12 weeks gestation in patients hospitalized for HG (HG), patients with NVP (NVP), and controls with no NVP before 24 weeks gestation (NO NVP).

[0018] Figure 3 is a box plot of serum levels of GDF15 (ng/ml), at 12 weeks gestation in
25 patients hospitalized for HG (HG), patients with NVP (NVP), and controls with no NVP before 24 weeks gestation (NO NVP).

[0019] Figure 4 is a box plot of serum levels of IGFBP7 (ng/ml), at 12 weeks gestation in patients hospitalized for HG (HG), patients with NVP (NVP), and controls with no NVP before 24 weeks gestation (NO NVP).

[0020] Figure 5 is a box plot of serum levels of beta HCG (mIU/ml), at 12 weeks gestation in
30 patients hospitalized for HG (HG), patients with NVP (NVP), and controls with no NVP before 24 weeks gestation (NO NVP).

[0021] Figures 6A-6C. Three of five HG families show segregation of alleles at GDF15 locus associated with increased expression of GDF15. Fig. 6A) Family 1 is of

English/Irish/Swedish/Welsh decent. Case (1A) reported iv fluid, total parenteral nutrition, antiemetic medication, home health care, and a 5% weight loss due to HG. Her cousin (1B)

5 reported normal NVP with no weight loss and no medication to treat NVP. Her aunt (1C)

reported antiemetic medication and weight loss due to HG, and her great aunt (1D) reported antiemetic medication and unrelenting nausea that kept her bedridden for 6 months due to HG.

Fig. 6B) Family 2 is of English/German/Scottish/Irish decent. Case (2A) reported iv fluid, hospitalization, >5% weight loss, and antiemetic medication to treat her HG. Her aunt (2B)

10 reported iv fluid, hospitalization, weight loss, and antiemetic medication to treat her HG. Her cousin [affected aunt's daughter, (2C)] reported antiemetic medication and a 24-pound weight loss in the first trimester due to HG. Participant 2D, the unaffected sister of 2C, reported 2 easy pregnancies with no weight loss nor treatment for NVP. Fig. 6C) Family 3 is of English/Irish

descent. Case (3A) reported iv fluid, hospitalization, >10% weight loss, and antiemetic

15 medication to treat HG. One affected sister (3B) reported iv fluid, hospitalization, and antiemetic medication to treat HG. The other affected sister (3C) reported iv fluid, hospitalization, weight loss, and antiemetic medication to treat HG. Their unaffected aunt (3E) reported mild NVP with no medication nor weight loss.

[0022] Figure 7 is Table 1.2, which shows the results of the genome-wide association study of no NVP vs HG (SCAN1) with $p < 10^{-6}$. The directions of odds ratios (OR) correspond to the minor allele, listed second.

[0023] Figure 8 is Table 1.4, which shows the results of GWAS of NVP as an ordinal response (SCAN2) with $p < 5 \times 10^{-8}$. The directions of effect correspond to the minor allele, listed second.

DETAILED DESCRIPTION OF THE INVENTION

25 [0024] Described herein are methods and kits for detecting and treating nausea and vomiting in pregnancy (NVP) and hyperemesis gravidarum (HG). These markers and methods of using these markers of NVP and/or HG address a long-felt need for diagnostic tests that can predict individual risk of NVP and HG, as well as treatment options for severe nausea and vomiting (HG) that is associated with poor outcomes including preterm birth, neurodevelopmental delay, and vitamin K deficient embryopathy. The invention provides genetic analyses as well as serum

30 tests, e.g. for GDF15 and IGFBP7, that can be used to identify patients at risk for HG.

Definitions

[0025] All scientific and technical terms used in this application have meanings commonly used in the art unless otherwise specified. As used in this application, the following words or phrases have the meanings specified.

5 [0026] As used herein, a "control" or "reference" sample means a sample that is representative of normal measures of the respective marker, such as would be obtained from normal, healthy control subjects, or a baseline amount of marker to be used for comparison. Typically, a baseline will be a measurement taken from the same subject or patient. The sample can be an actual sample used for testing, or a reference level or range, based on known normal
10 measurements of the corresponding marker.

[0027] As used herein, a "significant difference" means a difference that can be detected in a manner that is considered reliable by one skilled in the art, such as a statistically significant difference, or a difference that is of sufficient magnitude that, under the circumstances, can be detected with a reasonable level of reliability. In one example, an increase or decrease of 10%
15 relative to a reference sample is a significant difference. In other examples, an increase or decrease of 20%, 30%, 40%, or 50% relative to the reference sample is considered a significant difference. In yet another example, an increase of two-fold relative to a reference sample is considered significant.

[0028] "Nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides,
20 ribonucleotides, or peptide-nucleic acid sequences that may be assembled from smaller fragments, isolated from larger fragments, or chemically synthesized de novo or partially synthesized by combining shorter oligonucleotide linkers, or from a series of oligonucleotides, to provide a sequence which is capable of expressing the encoded protein.

[0029] As used herein, "pharmaceutically acceptable carrier" or "excipient" includes any
25 material which, when combined with an active ingredient, allows the ingredient to retain biological activity and is non-reactive with the subject's immune system. Examples include, but are not limited to, any of the standard pharmaceutical carriers such as a phosphate buffered saline solution, water, emulsions such as oil/water emulsion, and various types of wetting agents. Preferred diluents for aerosol or parenteral administration are phosphate buffered
30 saline or normal (0.9%) saline.

[0030] Compositions comprising such carriers are formulated by well-known conventional methods (see, for example, *Remington's Pharmaceutical Sciences*, 18th edition, A. Gennaro, ed., Mack Publishing Co., Easton, PA, 1990).

[0031] As used herein, the term "subject" includes any human or non-human animal. The term "non-human animal" includes all vertebrates, e.g., mammals and non-mammals, such as non-human primates, horses, sheep, dogs, cows, pigs, chickens, and other veterinary subjects. In a typical embodiment, the subject is a human.

- 5 [0032] As used herein, "a" or "an" means at least one, unless clearly indicated otherwise.

Methods of the Invention

[0033] The invention provides methods for analyzing genetic expression, or assaying gene expression status in a sample obtained from a subject, for diagnosing NVP and/or HG in a subject, as well as methods for predicting, treating, and/or monitoring NVP and/or HG.

- 10 [0034] In one embodiment, the method comprises obtaining a sample from the subject; and measuring the level of expression product of at least one of the hyperemesis gravidarum (HG) outlook genes identified in Table 1 in the sample. The method can further comprise assigning a HG outlook score to the measured amount of expression for each of the HG outlook genes, wherein the score is between 0 and 1; and calculating a HG status score, wherein the status
15 score is equal to the average outlook score assigned. In other embodiments, the combination is a total of HG status scores for each outlook gene measured or detected. In one embodiment, the presence of a genetic variation associated with HG is assigned a score of 1, and the absence of the genetic variation is assigned a score of 0. Those skilled in the art will appreciate other means of assigning and combining HG outlook scores. The combination can be calculated
20 or determined in other ways that result in an HG status score that reflects the number and/or quality of genetic variants detected.

- [0035] In one embodiment, the method comprises measuring the expression product of an HG outlook gene, wherein the HG outlook gene is selected from: GDF15, IGFBP7, PGR, and RYR2. In a typical embodiment, at least two of the HG outlook genes of Table 1 are measured.
25 In some embodiments, three, four, five, six, seven, eight, nine, ten, or more HG outlook genes are measured, including those listed in Table 1, and other HG outlook genes, such as RYR2. In one embodiment, the method comprises measuring the expression product of GDF15 and IGFBP7. In another embodiment, the method comprises measuring the expression product of at least 3 of the HG outlook genes selected from GDF15, IGFBP7, PGR, and RYR2. In yet another
30 embodiment, the method comprises measuring the expression product of at least 3 of the HG outlook genes selected from GDF15, IGFBP7, PGR, SYN3, MMADHC, TMEM38B, and RYR2. In one embodiment, the method comprises measuring the expression product of each of GDF15, IGFBP7, PGR, and RYR2. In one embodiment, the method comprises measuring GDF15 and/or IGFBP7.

[0036] In one embodiment, the method comprises detecting the presence of a genetic variant as described in Table 1. The detection of a genetic variant as described herein can be used to predict, detect and/or diagnose NVP, HG, cachexia, fetal loss, or an appetite disorder.

Described herein are a number of SNPs, some of which are on separate haplotypes for the

5 same gene (e.g., for GDF15 and PGR). In one embodiment, the method comprises detection of at least one SNP per haplotype.

[0037] The sample is a biological patient sample, representative examples of which include, a urine sample, a blood or serum sample, or a saliva sample. In a typical embodiment, the sample is tested for genotype and/or associated phenotype (RNA, protein expression levels) for a

10 minimum of 2 alleles in Table 1 (or for an allele with linkage disequilibrium to the allele in Table 1). Treatment for NVP or HG can be administered to the subject based on the assigned score for each test and/or combination of tests.

[0038] Thus, in some embodiments, the measuring comprises contacting the sample obtained from the subject with reagents for amplifying polynucleotides. In some embodiments, the level of
15 expression is measured via hybridization of probes that specifically bind polynucleotides encoding the expression product. In yet other embodiments, the level of expression is measured via binding of antibodies that specifically bind the expression product. In one embodiment, the levels of GDF15 and/or IGFBP7 are assayed. In one embodiment, the method comprises an immunoassay for GDF15 and IGFBP7, wherein levels of these proteins are detected in a patient
20 sample, such as, for example, a serum sample.

[0039] In one embodiment, the HG outlook gene is GDF15, IGFBP7, PGR, SYN3, MMADHC, TMEM38B, HCRTR2, GFRAL, RYR2 and/or PKHD1. The expression can, in a representative embodiment, be measured via detection of the G allele at rs16982345 (or any alleles in a shared haplotype (in linkage disequilibrium) with the G allele) at rs16982345 in GDF15.

25 [0040] In one embodiment, the method comprises:

(a) contacting a sample obtained from the subject with reagents that specifically bind to and/or amplify and/or detect at least one HG outlook gene selected from Table 1 or Table 2;

(b) measuring the level or presence of the HG outlook gene;

30 (c) assigning a status score that reflects the measured amount or presence of HG outlook gene relative to a normal control;

(d) referring the subject for treatment of NVP and/or HG if the status score is significantly greater than 1.

[0041] In one embodiment, step (a) is performed with one HG outlook gene from Table 1 and one HG outlook gene from Table 2. In another embodiment, step (a) is performed with at least two HG outlook genes from Table 1 and at least one HG outlook gene from Table 2. In one embodiment, step (a) is performed with one of the combinations of HG outlook genes listed in Table 3 or Example 5.

[0042] In some embodiments, the method further comprises treating the subject for nausea/vomiting of pregnancy (NVP) or hyperemesis gravidum (HG) if the HG status score is greater than zero. In one embodiment, the gene is GDF15 and the treatment comprises administering an anti-GDF15 agent to the subject. In another embodiment the gene is PGR and the treatment comprises administering an anti-PGR and/or anti-progesterone agent. In another embodiment, the gene is MMADHC and the treatment comprises administering vitamin B12 to the subject. The method can be adapted for related diseases, including cachexia and low appetite, while the reverse treatment would be applied for detection of the other allele for treating fetal loss and/or overactive appetite. In such a treatment (in reverse), the method comprises administering GDF15 if the genotype is A/A at rs16982345 in GDF15, and/or progesterone depending on the genotype at the PGR locus.

[0043] Other examples of treatment include, but are not limited to, administering to the subject one or more antiemetic agents (e.g., ondansetron, diclegis, propranolol, or a related drug that inhibits the pathway that includes RYR2 and other calcium channel signaling proteins).

[0044] Also provided is a method of monitoring NVP and/or HG, and/or cachexia, an appetite disorder, and/or fetal loss, in a subject. In one embodiment, the method comprises:

- (a) contacting a sample obtained from the subject with reagents that specifically bind to an expression product of at least one HG outlook gene selected from Table 1;
- (b) measuring the level of binding to the HG outlook gene expression product;
- (c) assigning a status score that reflects the measured amount of HG outlook gene expression product relative to a normal control;
- (d) treating the subject for NVP and/or HG (or related disease) if the status score is significantly greater than 1;
- (e) repeating steps (a) to (c); and
- (f) adjusting the treatment when the status score is not trending toward 1.

[0045] Also provided is a method of treating a subject for NVP or HG. In some embodiments, the method comprises performing the analyzing described herein; and administering a treatment for NVP or HG to the subject if the analyzing results in a status score that is greater than zero.

In some embodiments, the treatment comprises an inhibitor of GDF15, GFRAL, or PGR. In one embodiment, the invention provides a method of treating a subject for HG, wherein the method comprises analyzing a sample obtained from the subject for presence of one or more of the genotype combinations identified in Table 3 or Example 5 hereinbelow as associated with HG at a significance of $P \leq 0.05$; and administering an inhibitor of a corresponding gene if the patient is positive for the significant genotype combination. Also provided is a method of treating a subject with an inhibitor of GDF15 or GFRAL, wherein the method comprises assaying a sample obtained from the subject for the presence of a variant of GDF15 or GFRAL as described herein, and administering the inhibitor to the subject whose sample contains the corresponding variant.

Markers

[0046] Table 1: Gravidarum Outlook Genes:

[0047] Listed below are genetic variants (listed second in the "alleles" column) indicative of HG/NVP that can be detected via genotypic and/or phenotypic measures. A variant (associated with HG/NVP) can be detected, for example, by sequencing DNA obtained from a sample (or running an allelic discrimination assay), or, in some cases, by detecting an expression product that reveals the presence of the variant.

cytoband	SNP	alleles	gene context
19p13.11	rs45543339	C/T	<i>GDF15-LRRC25</i>
4q12	rs143409503	D/I	<i>IGFBP7-</i>
22q12.3	rs5998706	C/T	<i>SYN3-</i>
2q23.3	rs114571265	C/G	<i>MMADHC-RND3</i>
9q31.2	rs56108151	A/G	<i>TMEM38B-ZNF462</i>
19p13.11	rs16982345	A/G	<i>GDF15-LRRC25</i>
11q22.1	rs2508362	A/G	<i>PGR-TRPC6</i>
6q15	rs4707680	C/T	<i>MAP3K7-</i>
22q11.23	rs201838815	D/I	<i>SPECC1L</i>
6p12.1	rs7761177	C/T	<i>HCRTR2-GFRAL</i>
4q31.21	rs4690766	C/G	<i>INPP4B-USP38</i>
6p12.3	rs2076308	C/G	<i>TFAP2B-PKHD1</i>
9q21.32	rs11793821	A/G	<i>FRMD3-IDNK</i>
22q12.3	rs35970726	D/I	<i>SYN3-LARGE</i>
13q31.1	rs17077506	C/G	<i>SLITRK1-</i>
4q32.1	rs148644662	G/T	<i>CTSO-PDGFC</i>

21q21.1	rs11398016	D/I	<i>NCAM2</i>
2q37.3	rs12992546	C/T	<i>CXR7-COPS8</i>
6q16.2	rs75481205	A/G	<i>PRDM13-MCHR2</i>
	rs34345957	D/I	<i>LRRC25</i>
	rs4808787	A/G	<i>PGPEP1</i>
	rs2508353	G/T	<i>PGR-TRPC6</i>
	rs78571520	A/G	<i>PGR</i>
4q12	rs4865234	A/G	<i>IGFBP7-</i>
11q22.1	rs7948518	C/T	<i>PGR-TRPC6</i>
19p13.11	rs1058587	C/G	<i>GDF15</i>

[0048] In one embodiment, the HG outlook gene is at least one HG outlook gene selected from Table 1 or at least one combination of genes listed in Table 3 or Example 5. In one embodiment, at least two or more HG outlook genes of Table 1 is used in combination. Examples of combinations of HG outlook genes include: one or more of the combinations listed in Table 3 or Example 5; GDF15 and an additional HG outlook gene selected from Table 1; PGR and an additional HG outlook gene selected from Table 1; IGFBP7 and an additional HG outlook gene selected from Table 1; RYR2 and an additional HG outlook gene selected from Table 1; SYN3 and an additional HG outlook gene selected from Table 1; MMADHC and an additional HG outlook gene selected from Table 1; TMEM38B and an additional HG outlook gene selected from Table 1; MAP3K7 and an additional HG outlook gene selected from Table 1; SPECC1L and an additional HG outlook gene selected from Table 1; SPECC1L and an additional HG outlook gene selected from Table 1; GDF15 and IGFBP7 and an additional HG outlook gene selected from Table 1; any combination of 3 or more HG outlook genes selected from Table 1; any combination of 4, 5, 6, 7, 8, 9, 10, or 11 HG outlook genes of Table 1. In one embodiment, the combination of HG outlook genes is 2, 3, 4, or 5 HG outlook genes selected from Table 1. In another embodiment, the combination of HG outlook genes is at least one HG outlook gene from Table 1, and at least one HG outlook gene from Table 2. In another embodiment, the combination of HG outlook genes is at least one HG outlook gene from Table 1, and at least one combination of HG outlook genes from Table 3. In another embodiment, the combination of HG outlook genes is at least one HG outlook gene from Table 2, and at least one combination of HG outlook genes from Table 3. In another embodiment, the combination of HG outlook genes is at least one HG outlook gene from each of Tables 1, 2, and 3.

Kits and Assay Standards

[0049] The invention provides kits comprising a set of reagents as described herein, such as antibodies that specifically bind one or more HG outlook genes of the invention (or their

expression products), and optionally, one or more suitable containers containing reagents of the invention. Reagents include molecules that specifically bind and/or amplify and/or detect one or more HG outlook genes or related expression products of the invention. Such molecules can be provided in the form of a microarray or other article of manufacture for use in an assay

5 described herein. One example of a reagent is an antibody or nucleic acid probe that is specific for the HG outlook gene (or its expression product, e.g. GDF15 and/or IGFBP7). Another example includes probes (or primers) that selectively identify one or more genotypes described herein. Reagents can optionally include a detectable label. Labels can be fluorescent, luminescent, enzymatic, chromogenic, or radioactive.

10 [0050] Kits of the invention optionally comprise an assay standard or a set of assay standards, either separately or together with other reagents. An assay standard can serve as a normal control by providing a reference level of normal expression for a given HG outlook gene that is representative of a healthy individual.

[0051] Kits can include probes for detection of alternative gene expression products in addition
15 to antibodies for protein detection. The kit can optionally include a buffer. Reagents and standards can be provided in combinations reflecting the combinations of HG outlook genes described herein as useful for the detection of HG and/or NVP.

EXAMPLES

20 [0052] The following examples are presented to illustrate the present invention and to assist one of ordinary skill in making and using the same. The examples are not intended in any way to otherwise limit the scope of the invention.

Example 1: Placenta and appetite genes GDF15 and IGFBP7 are associated with HG

[0053] This Example describes a genome-wide association study (GWAS) for binary (HG)
25 and ordinal (severity of nausea and vomiting) phenotypes of pregnancy complications. Two loci, chr19p13.11 and chr4q12, are genome-wide significant ($p < 5 \times 10^{-8}$) in both association scans and are replicated in an independent cohort. The genes implicated at these two loci are GDF15 and IGFBP7 respectively, both known to be involved in placentation, appetite, and cachexia. This GWAS provides insights into the genetic risk factors contributing to the disease.
30 Supplementary materials referenced in this Example, including Supplementary Notes, Data, and Tables, can be found online in connection with the published version of this work at Fejzo et al., Nature Communications 9, Article Number 1178 (2018) doi:10.1038/s41467-018-03258-0, and

can also be found in related provisional patent application number 62/596,704, filed December 8, 2017.

[0054] Genome-wide association scan of the binary phenotype

[0055] For SCAN1, we compared the ends of the clinical spectrum of NVP, HG versus absence of NVP. Participants were female customers of 23andMe, who consented to participate in research, are of European ancestry, and have completed the relevant on-line surveys. A total of 1306 research participants reporting that they received intravenous fluid (IV) therapy for NVP were classified as HG cases, and 15,756 participants who reported no NVP served as controls. Cohort statistics are shown in Table 1.1. Manhattan and quantile–quantile plots are shown in Fig. 1a and Supplementary Fig. 1a, respectively, and SNP-level QC information is shown in Supplementary Table 1b.

[0056] Table 1.1 Demographic characteristics of unrelated female individuals of European descent included in SCAN1.

Phenotype	Group	Total	Female	Age < 30	30–44	45–59	≥60
HG	Case	1306	1306	45 (3%)	431 (33%)	555 (42%)	275 (21%)
	Control	15,756	15,756	392 (2%)	3487 (22%)	5357 (34%)	6520 (41%)

[0057] Most significantly associated single-nucleotide polymorphisms (SNPs) or insertion or deletion polymorphisms (indels) in each locus are shown in Table 1.2. Two loci, chr19p13.11 (rs45543339, OR = 0.67, 95% CI [0.60, 0.74], $p = 1.9 \times 10^{-14}$) and chr4q12 (rs143409503, OR = 0.75 [0.69, 0.82], $p = 4.5 \times 10^{-10}$), reached genome-wide significance ($p < 5 \times 10^{-8}$). The 99% credible set for the chr19p13.11 association signal overlapped two genes, GDF15 and LRRC25 (see Supplementary Figure 2), and contained a common missense variant in GDF15 (rs1058587, p.H202D, MAF = 0.25, OR = 0.68 [0.62, 0.75], $p = 3.4 \times 10^{-14}$) in high linkage disequilibrium (LD) with the lead SNP rs45543339 ($r^2 = 0.98$, $D' = 0.99$). The chr4q12 association signal fell within an intergenic region with the closest genes being an uncharacterized lncRNA LOC101928851 (~18.7 kbp away from rs143409503) and protein-coding gene IGFBP7 (~380 kbp away from rs143409503). In addition, there were three more association signals with $p < 10^{-6}$ (Table 1.2).

[0058] Table 1.2 shows the results of the genome-wide association study of no NVP vs HG (SCAN1) with $p < 10^{-6}$. The directions of odds ratios (OR) correspond to the minor allele, listed second.

[0059] Genome-wide association scan of the ordinal phenotype

[0060] For SCAN2, we next conducted a genome-wide association scan of the 53,731 unrelated female 23andMe research participants of European descent, using NVP severity as an ordinal response: none (N = 14,988), slight (N = 14,292), moderate (N = 17,786), severe (N = 5445), and very severe NVP (N = 1220), coded on a scale of 0 to 4. Details of the definition of each response based on the survey data can be found in the Methods section. Cohort statistics are shown in Table 1.3. Manhattan and quantile–quantile plots are shown in Fig. 1b, and SNP-level QC information is shown in Supplementary Table 1c.

[0061] Table 1.3 Demographic characteristics of unrelated female individuals of European descent included in SCAN2

Phenotype	Group	Total	Age < 30	30–44	45–59	≥60
NVP	None	14,988	373 (2.5%)	3332 (22.2%)	5129 (34.2%)	6154 (41.1%)
	Slight	14,292	233 (1.6%)	2635 (18.4%)	4831 (33.8%)	6593 (46.1%)
	Moderate	17,786	391 (2.2%)	3692 (20.8%)	6233 (35.0%)	7470 (42.0%)
	Severe	5445	140 (2.6%)	1298 (23.8%)	1580 (29.0%)	2427 (44.6%)
	Very severe	1220	42 (3.4%)	409 (33.5%)	519 (42.5%)	250 (20.5%)

[0062] The two most strongly associated loci in SCAN2 were also genome-wide significant in SCAN1: chr19p13.11 (rs16982345, $\beta = -0.104$ [$-0.12, -0.089$], $p = 2.4 \times 10^{-41}$) and chr4q12 (rs143409503, $\beta = -0.072$ [$-0.086, -0.058$], $p = 9.2 \times 10^{-24}$). Five additional regions reached genome-wide significance in SCAN2 (Table 1.4; Supplementary Figure 3) and there were eight additional association signals in SCAN2 with $p < 10^{-6}$ (Supplementary Table 2).

[0063] Table 1.4 shows the results of GWAS of NVP as an ordinal response (SCAN2) with $p < 5 \times 10^{-8}$. The directions of effect correspond to the minor allele, listed second.

[0064] Replication

[0065] Two loci were genome-wide significant in both association scans and were selected for replication genotyping. Given that the replication cohorts contain information on HG as a binary trait and that HG is the more clinically relevant phenotype, three additional loci with $p < 10^{-6}$ in the binary trait scan (SCAN1) were also selected for replication. The first replication cohort (HG-

IV) included 789 women with HG requiring IV fluid treatment and 606 controls reporting normal NVP. The second replication cohort (HG-TPN) included only women at the extreme ends of the clinical spectrum, 110 women requiring total parenteral nutrition (TPN) and 143 women reporting no NVP. Basic demographics of the replication cohorts are shown in Table 1.5. Over 90% of participants self-reported that they were of European descent.

[0066] Table 1.5 Demographic characteristics of the replication cohorts HG-IV consisting of HG patients treated with intravenous (IV) fluids and controls with untreated NVP, and HG-TPN consisting of HG patients treated with total parenteral nutrition (TPN) and controls with no NVP.

	Total	Year born (average)	Ethnicity (European descent)	Attended college	Advanced degree	First child year born
<i>HG-IV</i>						
HG	789	1977	90%	61%	19%	2003
Control	606	1975	92%	62%	18%	2002
<i>HG-TPN</i>						
HG	110	1976	90%	61%	26%	2002
Control	143	1974	93%	51%	25%	2001

[0067] We selected rs16982345 in the chr19p13.11 locus as a proxy for the lead SNP rs45543339 ($r^2 = 0.98$, $D' = 0.99$) for replication in HG-IV (because rs16982345 had a commercially available high quality SNP assay using the Taqman genotyping system). rs16982345 was successfully genotyped in 789 individuals with clinically defined HG and 581 controls reporting normal NVP using the TaqMan genotyping platform (Table 1.6). The call rate was >95%. The replication results for rs16982345 were supportive of the GWAS result ($p = 2.8 \times 10^{-7}$, OR = 1.63 [1.35, 1.98]).

[0068] Table 1.6 Replication results based on Fisher's exact test for the two most significantly associated loci in SCAN1 and SCAN2: rs16982345 (chr19p13.11), rs4865234 (chr4q12), and the three association signals in SCAN1 with $p > 5 \times 10^{-8}$ and $p < 1 \times 10^{-6}$, rs5754397 (SYN3) rs78353059 (MMADHC/RND3), and rs56108151 (TMEM38B/ZNF462).

	N	Genotype 1	Genotype 2	Genotype 3	p-value	OR	95% CI
<i>rs16982345</i>		GG	AG	AA			
HG-IV	789	540	229	20			

CONTROL	581	330	210	41	2.80×10^{-7}	1.63	1.35-1.98
HG-TPN	103	68	33	2			
NO NVP	136	75	51	10	0.04	1.61	1.01-2.60
<i>rs4865234</i>		AA	AG	GG			
HG-IV	778	404	312	62			
CONTROL	603	259	273	71	3.50×10^{-4}	1.35	1.14-1.59
HG-TPN	110	64	38	8			
NO NVP	143	57	66	20	2.81×10^{-3}	1.81	1.21-2.73
<i>rs5754397</i>		CC	CT	TT			
HG-IV	774	218	353	203			
CONTROL	606	180	273	153	0.51	0.95	0.82-1.11
HG-TPN	106	37	44	25			
NO NVP	141	44	64	33	0.72	1.07	0.74-1.56
<i>rs78353059</i>		AA	AG	GG			
HG-IV	779	735	42	2			
CONTROL	596	561	34	1	0.82	1.07	0.66-1.71
HG-TPN	105	99	6	0			
NO NVP	136	125	11	0	0.62	1.43	0.48-4.80
<i>rs56108151</i>		AA	AG	GG			
HG-IV	780	625	150	5			
CONTROL	594	495	97	2	0.06	0.78	0.59-1.02
HG-TPN	108	86	22	0			
NO NVP	141	117	23	1	0.64	0.86	0.45-1.65

[0069] The number (N) in Table 1.6 for each group is the total number of samples that were successfully genotyped. The call rate was >95%. OR estimates, 95% confidence intervals, and p-values were computed using 2×2 contingency tables in R. The effect allele is assumed to be the allele in the left-most (i.e., first) homozygous cell.

- 5 [0070] We also successfully genotyped the variant in 103 women with HG who required TPN and 136 women who reported no NVP in at least two pregnancies (HG-TPN). The genotyping results had comparable effect size, but were not significant after multiple testing correction ($p = 0.04$, OR = 1.61 [1.01, 2.60]) (Table 1.6).

[0071] We also performed replication of the second genome-wide significant locus. rs4865234 was used as a proxy for rs143409503 ($r^2 = 0.95$, $D' = 1.0$) (because rs143409503 is an indel, which cannot be assayed using the TaqMan genotyping system). rs4865234 replicated in HG-IV ($p = 3.5 \times 10^{-4}$, OR = 1.35 [1.14, 1.59]) (Table 1.6). The genotyping results of DNA from HG-TPN were also supportive ($p = 2.8 \times 10^{-3}$, OR = 1.81 [1.21, 2.73]).

[0072] The three additional SNPs with $p < 10^{-6}$ in the binary trait scan (rs5754397, rs78353059, rs56108151) did not replicate in HG-IV nor in HG-TPN ($p > 0.05$, see Table 1.6).

[0073] Combined analysis

[0074] For the replicated loci, we combined the replication results with SCAN1 using a fixed-effect meta-analysis (Table 1.7). Briefly, this test computes a weighted average of odds-ratio estimates across studies while accounting for standard errors. For rs16982345 on chr19p13.11, we estimated a meta-OR of 1.50 [1.38, 1.65] ($p = 2.12 \times 10^{-19}$). Heterogeneity of individual effect sizes did not appear to influence the combined estimate ($p_{het} = 0.66$). For rs4865234 on chr4q12, the meta-OR was 1.33 [1.23, 1.62] ($p = 9.29 \times 10^{-12}$) with little evidence of effect size heterogeneity ($p_{het} = 0.727$).

[0075] Table 1.7 Results of meta-analysis of SCAN1 and replication cohorts (A1 is the allele associated with increased risk of HG).

SNP	CHR	BP	A0	A1	Study	HG A1 Freq	CONTROL A1 Freq	OR	95% CI	<i>p</i> -value	Het <i>p</i>
rs16982345	chr19	18500722	A	G	SCAN1	0.8	0.73	1.47	[1.33, 1.63]	6.81E-15	–
					HG-IV	0.83	0.75	1.63	[1.35, 1.98]	2.80E-07	–
					HG-TPN	0.82	0.74	1.61	[1.01, 2.60]	4.00E-02	–
					META	–	–	1.5	[1.38, 1.65]	2.12E-19	6.62E-01
rs4865234	chr4	58355015	G	A	SCAN1	0.7	0.64	1.32	[1.20, 1.43]	7.95E-10	–
					HG-IV	0.72	0.66	1.43	[1.16, 1.79]	9.00E-04	–
					HG-TPN	0.75	0.63	1.82	[1.20, 2.70]	2.81E-03	–
					META	–	–	1.33	[1.23, 1.62]	9.29E-12	7.27E-01

[0076] Conditional analyses

[0077] In order to identify additional association signals, we performed stepwise conditional analysis at all loci that reached genome-wide significance in either SCAN1 or SCAN2.

Conditional analysis detected secondary associations in the chr19p13.11 locus in SCAN1 and in

the chr19p13.11 and PGR/TRPC6 loci in SCAN2 (Supplementary Table 3 and Supplementary Figures 4 and 5). LD between the primary and secondary signals is low ($r^2 < 0.01$), suggesting that they are independent signals corresponding to multiple risk variants within the locus.

[0078] The secondary signal in the chr19p13.11 locus identified from SCAN1 (rs1054221,

5 OR = 1.38 [1.22, 1.56], $p = 1.7 \times 10^{-7}$) maps within the 3' UTR of GDF15. The secondary signal from SCAN2 in this same locus is a common, single-nucleotide deletion with an effect comparable to the primary association (rs34345957, $\beta = 0.11$ [0.089, 0.13], $p = 1.1 \times 10^{-28}$, MAF = 0.14), located in the intron of LRRC25, and in high LD ($r^2 = 0.98$, $D' = 1.0$) with the secondary signal in SCAN1. The tertiary association signal within the chr19p13.11 locus in
10 SCAN2 is weaker (rs4808787, $\beta = -0.037$ [-0.053, -0.021], $p = 6.0 \times 10^{-6}$) and is located in the 5' UTR of the long isoform of PGPEP1.

[0079] Functional analyses of chr19p13.11 and chr4q12 genomic loci

[0080] The genomic context of the replicated loci chr19p13.11 and chr4q12 was analyzed using HaploReg¹⁷ (Supplementary Data 1). Within those two regions, we analyzed A) the index SNPs
15 for strongest association in SCAN1 (rs45543339 and rs143409503) and the proxy SNPs that were confirmed in B) the independent replication cohort (rs16982345 and rs4865234).

[0081] The lead SNP at chr19p13.11, rs45543339, was located in an intron of LRRC25, and the only notable functional annotation was an eQTL for KCNN1 in mucosa of the esophagus.

rs16982345, which is within the 99% credible set at this locus and is in tight LD ($r^2 = 0.98$) with
20 the lead SNP rs45543339, is located within an intron of GDF15 (Supplementary Data 1 and 2). It overlapped enhancer histone marks in three tissues, including placenta, and DNase hypersensitivity sites in nine tissues, including ovary. A summary of the variants in LD with the lead SNP and associated with altered GDF15 expression is shown in Supplementary Data
217,18,19,20,39,40,41,42,44, with the full list of variants in LD with the lead SNP identified using
25 HaploReg listed in Supplementary Data 1. One of the variants associated with HG in SCAN1, rs17725099 ($p = 1.03 \times 10^{-13}$) and in LD with rs16982345 ($r^2 = 0.77$), was the top association signal ($\beta = 0.16$, $p = 1.47 \times 10^{-107}$) in a conference abstract which identified variants influencing GDF15 levels in patients with cardiovascular disease¹⁸. The SNP rs17725099 was also associated with circulating levels of GDF15 in the first study to report on the heritability of
30 GDF15 plasma concentration¹⁹. In addition, rs1058587, a SNP in perfect LD ($r^2 = 1$) with the lead SNP and a part of the 99% credible set at this locus in SCAN1, has previously been associated with altered GDF15 expression. Expression levels of GDF15 were compared in an in vitro model between the cell line DU145 transfected with wild-type GDF15 (rs1058587, C allele) and GDF15 (rs1058587, G allele)²⁰. The C allele of rs1058587 was associated with increased

GDF15 levels in the in vitro model and was also the risk allele associated with HG, providing evidence that the direction of effect is toward higher levels of GDF15 for HG compared to controls.

[0082] The lead indel at chr4q12, rs143409503, mapped to an intergenic region with the closest protein-coding gene being IGFBP7, ~380 kbp away from the marker. No notable features were identified using HaploReg¹⁷, the GTEx portal, nor ExSNP21 databases. Other phenotypes associated with SNPs in GDF15 and IGFBP7 are reported in Supplementary Note

¶ 41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,56

[0083] Discussion

[0084] This GWAS of HG identified two genome-wide significant signals that were subsequently replicated in the larger of two independent cohorts (HG-IV). The most significantly associated locus on chr19p13.11 contained genes GDF15 and LRRC25. While we cannot be certain which gene or genes are implicated, SNPs in LD with the associated variants at the chr19p13.11 locus were found to be associated with altered expression of GDF15. In addition, the gene encoding the receptor for GDF15, the brainstem-restricted receptor GFRAL22, was associated with NVP in SCAN2 (Table 1.4) adding further evidence supporting a previously unknown biological connection between GDF15 and HG. While the results of this study do not establish a causal link between GDF15 and HG, the association between this gene and HG is of particular importance because it highlights the possibility of a pathway involved in the etiology of the condition. GDF15 encodes a TGF- β superfamily member that is expressed at its highest levels in the trophoblast cells of the placenta²³. The protein is found in maternal serum and increases significantly in the first two trimesters²³. GDF15 is believed to suppress production of proinflammatory cytokines in order to facilitate placentation and maintain pregnancy²⁴. In addition to its role in pregnancy, GDF15 has been shown to be a regulator of physiological body weight and appetite via activation of neurons in the hypothalamus and area postrema (vomiting center) of the brainstem^{25,26}. It is also notable that abnormal overproduction of GDF15 in cancer was recently found to be the key driver of cancer anorexia and cachexia which, like HG, exhibits symptoms of chronic nausea and weight loss^{27,28}. Of particular clinical interest, inhibition of GDF15 restored appetite and weight gain in a mouse model of cancer cachexia²⁸, suggesting a therapeutic strategy that may be applicable to patients with HG, if GDF15 proves to be the implicated gene.

[0085] The other locus that reached genome-wide significance and was confirmed in the independent replication cohort, is chr4q12. Although there were no genes that can be directly related to the association signal, the protein-coding gene closest to the lead variant is IGFBP7.

We have not found eQTLs that link the variants associated in this study to expression of IGFBP7, however, this can be explained at least in part by a lack of eQTL data in relevant tissues during pregnancy. The notably similar and functionally relevant roles of IGFBP7 and GDF15 (in placentation, cachexia, and feeding behavior) make a compelling argument for IGFBP7 despite the lack of evidence linking the association signal to its expression. Insulin-like growth factor binding protein 7 (IGFBP7) is a gene involved in implantation and decidualization of the pregnant uterus²⁹. Like GDF15, IGFBP7 is significantly upregulated after implantation and highly expressed in the developing placenta³⁰. Inhibition of IGFBP7 caused pregnancy loss in a mouse model by shifting uterine cytokines to Th1 type dominance and repressing uterine decidualization³⁰. Thus IGFBP7 may play roles in both miscarriage and in the severity of NVP, providing a genetic mechanism for the protective effect of NVP. In addition, like GDF15, evidence points to IGFBP7 as a promising biomarker for cachexia associated with end-stage disease²⁹. An intriguing finding is that the *Drosophila* homolog of IGFBP7 has been suggested to play a role in neuronal coordination between metabolic status and feeding behavior, potentially, like with HG, causing food aversion to normally attractive food, even when starving, and vice versa³¹.

[0086] The strengths of this study include the well-defined phenotype and large sample sizes in two scans and the HG-IV replication sample. The strong independently replicated association signals, and the fact that both candidate genes have roles in early pregnancy when HG sets in, support these genes as functional as well as positional candidate genes^{30,32}. However, it is important to replicate the findings reported herein, both using larger replication sample sizes (for HG-TPN vs no NVP) and in studies of different ethnicities to determine whether the findings can be validated and generalized to other populations. Another caveat is that while the conditional analysis yielded additional association signals at chr19p13.11 that may act independently of the lead association signal, they were not replicated. Replication is necessary to determine if additional independent variants are associated at this locus. Finally, the lack of eQTL data and considerable distance between the chr4q12 association signal and the closest gene IGFBP7 is a noteworthy limitation.

[0087] Since GDF15 and IGFBP7 levels are upregulated during placentation and cachexia^{23,24,28,29}, and downregulated prior to miscarriage^{30,32}, serum concentrations of GDF15 and IGFBP7 should now be studied in pregnant women with and without HG. Drugs targeting these proteins provide clinical utility in treating HG. Such a study may also lead to new methods for prediction and diagnosis. In addition, the SCAN2 results suggest a great potential for discovering additional loci associated with HG or NVP. Finally, the findings herein suggest an answer to an age-old paradox. HG can lead to prolonged dehydration and undernutrition, which can be detrimental to maternal and fetal health and can decrease reproductive fitness. The dual

roles of GDF15 and IGFBP7 in maintaining pregnancy and in increasing the risk of HG may provide a molecular explanation for why NVP still exists in nature.

[0088] Methods

[0089] Human subjects for GWAS

5 [0090] GWAS participants were customers of 23andMe, Inc. who consented to participate in research, and provided answers to the morning sickness-related questions. Two genome-wide association scans were performed, for (1) a binary HG phenotype and (2) an ordinal phenotype related to HG. Phenotype definitions are described below. Due to the nature of the phenotype definitions, some research participants were included in both association scans. All research
10 participants included in the analysis provided informed consent and answered on-line surveys according to a human subjects protocol approved by Ethical & Independent Review Services, a private institutional review board.

[0091] Phenotype definitions

[0092] For the binary HG phenotype analyzed in SCAN1, we compared two ends of the clinical
15 spectrum of NVP: 1306 research participants who reported via an on-line survey that they received IV therapy for NVP were classified as HG cases and 15,756 participants who reported no NVP served as controls. The phenotype definition was ascertained using the interview questions listed in Supplementary Note 2. For the ordinal phenotype related to HG analyzed in SCAN2, data were pulled from four questions also listed in Supplementary Note 2.

20 [0093] Genotyping and imputation and association testing

[0094] A total of 17,062 females in SCAN1 and 53,731 females in SCAN2 were genotyped on one of four custom Illumina genotyping arrays and additional genotypes were imputed using the September 2013 release of the 1000 Genomes Project Phase 1³³ reference haplotypes as described previously^{34,35}. 16,165 individuals in SCAN1 were also included in SCAN2.
25 Breakdown of the number of participants in specific phenotype categories in SCAN1 and SCAN2 are shown in Supplementary Table 1a. Fields that contain 5 or fewer individuals have been masked to protect the privacy of 23andMe customers. All females were filtered to select for European ancestry, and close relatives were removed. We performed logistic (SCAN1) and linear (SCAN2) regression assuming an additive model for allelic effects on NVP, using age and
30 five principal components of genetic ancestry as covariates. Our previously published analysis of categorical phenotypes using ordinal and linear regression showed high concordance between resulting p-values³⁶. Due to comparative ease of implementation and lower

computational burden, we used linear regression in SCAN2. The SNP-level quality control information is shown in Supplementary Tables 1b and 1c. Test statistics were further adjusted for the genomic control inflation factor of $\lambda = 1.044$ for SCAN1 and $\lambda = 1.087$ for SCAN2. The index SNPs were identified by choosing the SNPs with the strongest association in each associated region. Each region contained SNPs with $p < 10^{-5}$ that were grouped into intervals separated by a gap of a minimum of 250 kbp. The SNP with the smallest p-value within each interval was chosen as the lead SNP.

[0095] Credible sets

[0096] To account for the uncertainty in identifying the true causal variant within an association signal, we computed 99% credible sets at each locus. Under the assumptions that there is a single causal variant at a locus and that this variant was genotyped in the association study, 99% credible sets are defined as smallest sets of variants that are 99% likely, based on posterior probability, to contain the causal variant³⁷. While these assumptions are not always true in practice, credible sets provide a helpful summary of the available evidence that one of the SNPs contained in them is causal, and may help in fine-mapping the association signal.

[0097] Replication cohorts

[0098] The first replication cohort (HG-IV) included 789 HG cases treated with IVs and 606 controls with NVP that did not require treatment. The second replication cohort (HG-TPN) included 110 cases requiring TPN due to HG and 143 controls reporting no NVP in any of at least two pregnancies. All participants gave informed consent. This study was approved by the UCLA Institutional Review Board.

[0099] Phenotype definition in the replication cohort

[0100] Affected individuals were included if they had an HG diagnosis and were treated with IV fluids and/or TPN. Participants affected by HG recruited acquaintances that were non-blood related and had at least two pregnancies lasting more than 27 weeks gestation. Controls were eligible to participate in the study if they reported normal or no NVP, did not lose weight due to nausea/vomiting, and did not receive medical attention due to NVP.

[0101] Recruitment of the replication cohort

[0102] For the replication study, the source population for HG cases included patients residing in the US. The majority of cases were recruited through a posting from 2007 to 2017 on the Hyperemesis Education and Research Foundation website (www.HelpHer.org). Minors (under

18 years) were excluded because the risk to benefit ratio to control minors would be difficult to justify and the study requires controls to have had two pregnancies, which is uncommon in minors. Women with pregnancies with chromosome abnormalities and multiple gestations were excluded due to possible distinct etiologies for NVP in these cases. Medical records and
5 recruitment of an acquaintance with at least two pregnancies to serve as a control, were requested of each case. All participants were required to go over an information sheet by phone and return a signed information sheet with all elements of consent in order to enroll in the study.

[0103] Sample collection

[0104] Saliva samples were collected for DNA analysis from all cases and controls. DNA
10 Genotek saliva kits (Oragene, Ottawa, Canada) were mailed to all cases and controls. The saliva collection kit is self-administered and comes with directions for submitting 2 ml of saliva into a collection vial and returning the sample to the study site via an addressed and postage-paid return envelope provided with the collection kit.

[0105] DNA extraction

15 [0106] DNA was extracted from the saliva samples according to manufacturer's instructions (Oragene, Ottawa Canada). Using the kit, we have successfully isolated, on average, 197 µg of DNA of high quality (260/280 1.84) from 2 ml of saliva. The low end of expected DNA quantity reported by the manufacturer is 30 µg/ml of saliva or 60 µg/sample. After the extraction, the DNA was stored at -20 °C.

20 [0107] Quality control for replication data

[0108] The TaqMan genotyping platform was performed on 384-well plates with a minimum of two blank samples per plate and a minimum of two duplicate samples per plate. Once genotypes were determined from the first 384-well plate, a minimum of three positive controls or one positive control for each genotype was added to the remaining plates. The minimum call
25 rate for each SNP was >95%.

[0109] Genotyping and statistical analysis in the replication study

[0110] rs16982345, the SNP with the strongest association to NVP in the chr19p13.11 locus was selected for genotyping in the replication sample HG-IV consisting of 789 cases and 606 controls. For the non-coding marker rs16982345 in chr19p13.11, genotyping a minimum of 780
30 cases and 580 controls, and assuming a proportion of the GG genotype similar to SCAN1 (0.53 for controls and 0.64 for cases), we estimated a power greater than 97% to detect between-

group genotype differences using a two-sample, two-sided test of proportions at significance level $\alpha = 0.05$.

[0111] In addition, we genotyped an extreme cohort (HG-TPN) consisting of 110 women who reported TPN for HG and 143 controls with no NVP in two or more pregnancies. TaqMan genotyping primers for rs16982345 were available from Thermo Fisher Scientific and Applied Biosystems PRISM 7900HT Sequence Detection System (TaqMan) was used for large-scale screening. The call rate was >95%.

[0112] Genotypes were tested for significant association with HG using Fisher's exact test using 2×2 contingency tables counting the number effect vs non-effect alleles for cases and controls using R.

[0113] The SNP rs4865234 was genotyped as a proxy for the most significant association signal at chr4q12 because the most significant association signal, rs143409503, is an indel, which cannot be assayed using the TaqMan genotyping system. The SNP rs4865234 is in tight LD with rs143409503 ($r^2 = 0.95$, $D' = 1.0$) and genotyped on DNA isolated from HG-IV and HG-TPN cohorts as stated previously for rs16982345.

[0114] The 3 additional loci that reached significance $p < 10^{-6}$ in the binary trait scan (SCAN1) were also genotyped in the two replication cohorts (HG-IV and HG-TPN) using the TaqMan genotyping platform. The SNP rs5754397 was used as a proxy for the lead SNP rs5998706 ($r^2 = 0.55$, $D' = 0.83$), and the SNP rs78353059 was used as a proxy for the lead SNP rs114571265 ($r^2 = 0.87$, $D' = 1.0$) because they were the closest SNPs listed in HaploReg¹⁷ to the lead SNP with an available commercially validated assay from Thermo Fisher Scientific and they also reached significance $p < 10^{-6}$ in SCAN1. The lead SNP rs56108151 was commercially available and used for genotyping the two replication cohorts.

[0115] Meta-analysis of SCAN1 and replication

[0116] We performed a fixed-effect meta-analysis using the R language and software meta to compute the weighted average odds-ratio across the replication group and SCAN1 for sites rs16982345 and rs4865234. The fixed-effect meta-analysis assumes that there exists a true underlying effect in the population and combines individual study estimates (e.g., replication and SCAN1) to obtain a meta estimate. This assumption may be invalid in practice due to differences in linkage disequilibrium patterns across studies, or differences in covariates included. To test for heterogeneity in effects across studies, or significant study-wide variance, we estimated Cochran's Q^{38} for each SNP. This measures the variance across studies using the

meta estimate as the true mean, which can be tested in a null-hypothesis framework under a χ^2_{n-1} distribution. Here, the number of studies is $n=2$, so our heterogeneity p-values reflect a one-tailed test under a χ^2_1 distribution.

[0117] Conditional analysis

5 [0118] The stepwise conditional analysis detects statistically independent signals in a GWAS locus. At each step, we assessed association between variants within 20 kbp around a GWAS locus and the HG or related ordinal phenotype, using top SNPs from preceding steps as additional covariates. This was repeated until no significant association was detected at $p < 10^{-5}$. Finally, we fit a joint model with all of the primary and secondary signals and evaluated
10 its goodness of fit through the likelihood ratio test (LRT) against the model with just the primary signal. The reported effect sizes of the secondary signals are from the joint model, and the p-values are from the likelihood ratio test comparing the full model with the leave-one-out models.

[0119] In silico functional follow-up

[0120] We used the HaploReg¹⁷ v4.1 tool to identify SNPs in the haplotype blocks containing
15 the two significant GWAS association signals, their LD r^2 and D' scores, associated alleles, allele frequencies, genes, functional annotation, conservation, regulatory elements, protein-binding sites, motif changes, and eQTLs. We queried the two most significant loci identified, chr19p13.11 and chr4q12. Within those two regions, we analyzed SNPs with HaploReg, which uses LD information from the 1000 Genomes Project³³. The LD threshold in HaploReg v4.1 was
20 set at $r^2 = 0.2$: the population used for LD calculations was the 1000 Genomes Project Phase 1 European descent; the source for epigenomes was the ChromHMM (Core 15-state model); the mammalian conservation algorithms GERP and SiPhy-omega, and positions are shown relative to GENCODE genes (Supplementary Data 1). HaploReg v4.1 includes GWAS and eQTL (including GTEx) and GRASP eQTL updates which were queried for the SNPs and included in
25 Supplementary Data 1.

[0121] We also queried the GTEx portal and ExSNP integrated eQTL databases²¹ for the top candidate genes (based on loci identified in this study and similar function in early pregnancy reported previously^{30,32}, GDF15 and IGFBP7) to identify eQTLs (17 and 28 studies, respectively) included in the databases. The Genotype-Tissue Expression (GTEx) Project was
30 supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from (gtexportal.org/home/bubbleHeatmapPage/)

the GTEx Portal on 05 May 2017. The SNPs identified were cross checked with the SNPs identified in haplogroups by HaploReg.

[0122] Functional annotation of all SNPs (SNPs $r^2 > 0.2$ with SNPs in the two loci that were found to be in coding regions by HaploReg were queried for all predicted consequences using the 1000 Genomes Project³³.

[0123] Finally, published studies reporting SNPs linked to altered expression of the top candidate genes (GDF15 and IGFBP7) and SNPs in the top candidate genes reported in association studies in the GWAS catalog (<http://www.ebi.ac.uk/gwas/> accessed on May 2017) and/or in published studies were cross checked with the SNPs identified in HaploReg to determine whether they are in LD with the SNPs linked to HG in this study.

[0124] Code availability

[0125] All relevant software, including version details, used for the primary GWAS and conditional analyses are described in the references included in the Methods section.

[0126] Data availability

[0127] Qualified researchers can contact apply.research@23andMe.com to gain access to full GWAS summary statistics following an agreement with 23andMe that protects 23andMe participant privacy.

[0128] References cited

1. Clark S. M., et al. *Obstet. Gynecol. Int.* 2012, 252676 (2012).
2. Piwko, C., et al. *J. Popul Ther. Clin. Pharmacol.* 20, e149–e160 (2013).
3. Kallen, B. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 26, 291–302 (1987).
4. Verberg, M. F., et al. *Hum. Reprod. Update* 11, 527–539 (2005).
5. Drife, J. O. *BMJ* 344, e567 (2012).
6. Gazmararian, J. A. et al. *Obstet. Gynecol.* 100, 94–100 (2002).
7. Matthews, A., et al. *Cochrane Database Syst Rev.* 9, CD000145 (2015).
8. Hinkle, S. N. et al. *JAMA Intern. Med.* 176, 1621–1627 (2016).
9. Fejzo, M. S. et al. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 170, 71–76 (2013).
10. Fejzo, M. S., et al. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 189, 79–84 (2015).
11. Toriello, H. V. et al. *Am. J. Med. Genet.* 161A, 417–429 (2013).
12. Corey, L. A., et al. *Obstet. Gynecol.* 80, 989–994 (1992).
13. Colodro-Conde, L. et al. *Behav. Genet.* 46, 481–491 (2016).

14. Vellacott, I. D., et al. *Int J. Gynecol. Obstet.* 27, 57–62 (1988).
15. Zhang, Y. et al. *Am. J. Obstet. Gynecol.* 204, 230e.1–7 (2011).
16. Vikanes, A. et al. *BMJ* 29, 340 (2010).
17. Ward, L. D. & Kellis, M. *Nucleic Acids Res.* 40, D930–D934 (2012).
- 5 18. Eriksson, N. et al. In *Annual Meeting of the American Society of Human Genetics* (American Society of Human Genetics, 2013).
19. Ho, J. E. et al. *Clin. Chem.* 58, 1582–1591 (2012).
20. Wang, X., et al. *Prostate* 72, 677–689 (2012).
21. Yu, C. H., et al. *OMICS* 20, 400–414 (2016).
- 10 22. Hsu, J.-Y. et al. *Nature* 550, 255–259 (2017).
23. Moore, A. G. et al. *J. Clin. Endocrinol. Metab.* 85, 4781–4788 (2000).
24. Marjono, A. B. et al. *Placenta* 24, 100–106 (2003).
25. Tsai, V. W. et al. *PLoS ONE* 8, e55174 (2013).
26. Tsai, V. W. et al. *PLoS ONE* 9, e100370 (2014).
- 15 27. Lerner, L. et al. *J. Cachexia Sarcopenia Muscle* 6, 317–324 (2015).
28. Lerner, L. et al. *J. Cachexia Sarcopenia Muscle* 7, 467–482 (2016).
29. Loncar, G., et al. *Int. J. Mol. Sci.* 15, 23878–23896 (2014).
30. Liu, Z. K., et al. *PLoS ONE* 7, e45224 (2012).
31. Bader, R. et al. *J. Cell Sci.* 126, 2571–2576 (2013).
- 20 32. Tong, S. et al. *Lancet* 363, 129–130 (2004).
33. The 1000 Genomes Project Consortium et al. *Nature* 526, 68–74 (2015).
34. Hromatka, B. S. et al. *Hum. Mol. Genet.* 24, 2700–2708 (2015).
35. Tung, J. Y., et al. *PLoS ONE* 6, e23473 (2011).
36. Jones, A. V. et al. *Hum. Mol. Genet.* 26, 1391–1406 (2017).
- 25 37. The Wellcome Trust Case Control Consortium, et al. *Nat. Genet.* 44, 1294–1301 (2012).
38. Schwarzer, G., et al. *Meta-analysis with R* (Springer International Publishing, Switzerland, 2015).
39. Lappalainen, T. et al. *Nature* 501, 506–511 (2013).
40. Barreiro, L. B. et al. *Proc. Natl Acad. Sci. USA* 109, 1204–1209 (2012).
- 30 41. Ek, W. E. et al. *Hum. Mol. Genet.* 25, 817–827 (2016).
42. Athiyarath, R., et al. *Ann. Hematol.* 12, 2093–2095 (2014).
43. Wang, X. et al. *Clin. Sci.* 118, 137–145 (2009).
44. Brown, D. A. et al. *Clin. Cancer Res.* 15, 6658–6664 (2009).
45. Locke, A. E. et al. *Nature* 518, 197–206 (2015).
- 35 46. Julià, A. et al. *Hum. Mol. Genet.* 23, 6927–6934 (2014).
47. Arakawa, S. 1 et al. *Nat. Genet.* 43, 1001–1004 (2011).
48. Sun, Y. 1 et al. *Invest. Ophthalmol. Vis. Sci.* 54, 8199–8203 (2013).
49. Lauc, G. et al. *PLoS Genet.* 9, e1003225 (2013).

50. Yang, B. Z., et al. Am. J. Med. Genet. B Neuropsychiatr. Genet. 174, 261–268 (2017).
 51. Patel, Y. M. et al. Cancer Epidemiol. Biomark. Prev. 24, 119–127 (2015).
 52. Huang, Y.-J. et al. Mutat. Res. 702, 32–39 (2010).
 53. Nicoletti, P. et al. Oncologist 17, 279–287 (2012).
 54. Fan, Y. 1 et al. Anim. Genet. 45, 740–744 (2014).
 55. García-Gámez, E. et al. PLoS ONE 6, e21158 (2011).
 56. Luo, W. et al. Int. J. Biol. Sci. 8, 870–881 (2012).

Example 2: IGFBP7 and PGR, in addition to GDF15, associated with nausea and vomiting of pregnancy and hyperemesis gravidarum

- 10 [0129] This Example provides a replication study confirming the results described in Example 1, particularly with respect to IGFBP7 and PGR.

[0130] Table 2. Loci Replication Results

Replicated Loci	GWAS 1 P-Value	GWAS 2 P-Value	Replication P-Value, OR [95% CI]
GDF15-LRRC25	1.8X10 ⁻¹⁴	2.4X10 ⁻⁴¹	0.0001, 2.9 [1.7-5.0]
lncRNA LOC101928851-IGFBP7	4.9X10 ⁻¹⁰	9.2X10 ⁻²⁴	0.0009, 1.4 [1.2-1.8]
TRPC6-PGR	9.5X10 ⁻⁵	7.3X10 ⁻¹²	0.0001, 1.9 [1.5-2.3]

[0131] One SNP/Locus was tested in a replication sample of over 1300 individuals.

- 15 Example 3: Combinations of markers increase risk detection of hyperemesis gravidarum

[0132] This Example demonstrates the increased risk detection achieved when different combinations of 3 or 4 genes are measured.

[0133] Table 3. Combinations of genotypes in genes increase risk for HG.

GDF15, IGFBP7, PGR				
Control (n=134)	TPN (n=106)	Odds ratio	95 % CI:	Significance:
7 (5.2%)	22 (20.8%)	4.8	1.9 to 11.6	P = 0.0006
GDF15, PGR, RYR2				
Control (n=133)	TPN (n=106)	Odds ratio	95 % CI:	Significance:
18 (13.5%)	35 (33.0%)	3.1	1.7 to 6.0	P = 0.0005

GDF15, IGFBP7, RYR2				
Control (n=135)	TPN (n=106)	Odds ratio	95 % CI:	Significance:
28 (20.7%)	35 (33.0%)	1.9	1.1 to 3.4	P = 0.0325
IGFBP7, PGR, RYR2				
Control (n=138)	TPN (n=109)	Odds ratio	95 % CI:	Significance:
13 (9.4%)	30 (27.5%)	3.7	1.8 to 7.4	P = 0.0003
GDF15, IGFBP7, PGR, RYR2				
Control (n=133)	TPN (n=106)	Odds ratio	95 % CI:	Significance:
7 (5.3%)	22 (20.8%)	4.7	1.9 to 11.5	P = 0.0007

[0134] Combinations of genotypes of selected risk genes are compared for HG cases requiring feeding tube (TPN) vs Controls with no nausea/vomiting in at least 2 pregnancies.

Example 4: Analysis of GDF15 and IGFBP7 in Hyperemesis Gravidarum lends further support for a role in the etiology of severe nausea and vomiting of pregnancy

[0135] This Example shows that GDF15 and IGFBP7 can be used to identify patients at risk for HG. A test for HG can now be provided, as the data described herein shows that a blood test may be done that predicts over a 36-fold increased risk for hospitalization for HG ($p=0.0014$) based on the protein levels of GDF15 and IGFBP7 combined.

[0136] The GWAS shows placentation, appetite, and cachexia genes *GDF15* and *IGFBP7* are linked to Hyperemesis Gravidarum (HG). This Example shows that GDF15 and IGFBP7 are upregulated in HG patients. The Example also shows that alleles are associated with symptoms, medication effectiveness, and segregation with disease in families. GDF15 serum levels are significantly increased in women with HG compared to controls. GDF15 alleles may be linked to metoclopramide response and segregate with HG in 3 of 5 families. The IGFBP7 risk allele may be associated with ondansetron and promethazine response, and prolonged symptoms. This study supports GDF15 and possibly IGFBP7 in the pathobiology of HG and provides a tool for prediction and diagnosis. In addition, the GDF15-GFRAL brainstem-activated pathway was recently identified, and therapies to treat conditions of abnormal appetite are under intense investigation. HG should be added to the list.

[0137] Mounting evidence for a genetic predisposition to HG suggests understanding the genetic component is essential in discovering an etiology. Using an existing cohort with well-

defined clinical criteria of intravenous fluid treatment for HG, the results presented herein successfully validated two candidate genes from the genome-wide association study (GWAS) that are highly likely to play a role in HG. The top two genes, *GDF15* and *IGFBP7*, are both involved in placentation and feeding behavior, which leads to a fascinating paradigm shift away from HCG. These results support the use of GDF15 and IGFBP7 as biomarkers for prediction, diagnosis, and identification of novel treatments.

[0138] The proteins encoded by the genes identified in the GWAS (*GDF15*, *IGFBP7*) are involved in placentation, are upregulated in early pregnancy, and cause loss of appetite in animal models. To determine whether serum levels are dysregulated in HG patients, the levels of these biomarkers were compared in pregnant women with HG. Serum concentrations of GDF15, IGFBP7, and HCG were analyzed at 12 weeks gestation in 11 women hospitalized for HG, compared to 9 women with normal NVP, and 20 women with no NVP.

[0139] Demographic characteristics are shown in Table 4.1 below. The serum concentrations of GDF15 and IGFBP7 were significantly increased ($p < 0.001$, $p = 0.002$, respectively) in women hospitalized for HG compared to women with NVP. The serum concentrations of GDF15 were significantly higher in patients with HG compared to women with no NVP ($p < 0.001$), but this level of significance was not observed for serum levels of IGFBP7 ($p = 0.123$). There was no difference in serum GDF15 or IGFBP7 levels in patients with NVP compared to NO NVP ($p = 0.945$, $p = 0.274$, respectively). Using the same samples, the serum concentrations of beta HCG at 12 weeks gestation were not significantly different between women with and without HG ($p = 0.142$). In summary, GDF15 is upregulated in pregnancies affected by HG compared to both control groups, but the results for IGFBP7 showed more variability, and beta HCG is not upregulated in HG pregnancies compared to either control group.

[0140] The scatterplot shown in Figure 2 displays the levels (ng/ml) of GDF15 (protein 1) and IGFBP7 (protein 2) from women with HG, NVP, or no NVP. Serum levels at 12 weeks gestation were compared for 11 women hospitalized for HG compared to 29 women who did not have HG (see Figures 3-5). Evaluation of odds ratios shows that patients are 36-fold more likely to be hospitalized for HG with values of GDF15 above 10 ng/ml and IGFBP7 above 60 ng/ml.

Odds ratio	36.2500
95 % CI:	4.0180 to 327.0399
z statistic	3.199

Significance level	P = 0.0014
--------------------	------------

[0141] 10/11 HG greater than 10 ng/ml gdf15, greater than 60 ng/ml igfbp7 blood levels. (1 less); 8/29 NO HG (NVP+NO NVP) greater than 10 ng/ml gdf15, greater than 60 ng/ml igfbp7 blood levels. (21 less). These data show that a blood test may be done that predicts over a 36-
5 fold increased risk for hospitalization for HG (p=0.0014).

[0142] Statistical analysis

[0143] GDF15 > 14 and IGFBP7 > 60

	HG	no HG
serum high	8	4
serum low	3	25

[0144] Classification of serum values:

- serum high: GDF15 > 14 and IGFBP7 > 60
- serum low: GDF15 ≤ 14 and/or IGFBP7 ≤ 60

[0145] The fisher test tests the table for significance. The p-value is 0.00074, thus there is a significant difference for HG vs. no HG in the categories high and low serum values.

[0146] GDF15 > 10 and IGFBP7 > 60

	HG	no HG
serum high	10	7
serum low	1	22

[0147] Classification of serum values:

- serum high: GDF15 > 10 and IGFBP7 > 60
- serum low: GDF15 ≤ 10 and/or IGFBP7 ≤ 60

[0148] The fisher test tests the table for significance. The p-value is 0.000199, thus there is a significant difference for HG vs. no HG in the categories high and low serum values. The significance for GDF15>10ng/ml as a test alone, for IGFBP7 >60ng/ml as a test alone, was compared to using both combined, and the results show that it is better to do both combined (more significant result and greater odds of detecting HG).

[0149] Table 4.1. Demographic characteristics of cohort with serum (HG=pregnant women hospitalized for HG $\leq$ 20 weeks gestation, NV=pregnant women with normal NVP, NO=pregnant women with no NVP $\leq$ 24 weeks gestation)

[0150] A. Continuous variables

Variable	Group	N	Min	Q1	Median	Mean	Q3	Max	Sd	Missing
Maternal age	HG	11	20.0	25.5	29.0	28.8	31.5	38.0	4.8	0
	NV	9	26.0	29.0	31.0	30.2	32.0	33.0	2.2	0
	NO	20	22.0	29.8	34.0	33.4	37.2	42.0	5.2	0
	all	40	20.0	28.0	31.0	31.4	34.2	42.0	4.9	0
BMI pre-pregnancy	HG	11	15.0	21.0	22.4	23.6	26.4	32.9	5.2	0
	NVP	8	22.2	24.3	25.5	26.0	26.6	34.0	3.6	1
	NO	18	18.7	21.6	23.0	23.3	24.7	27.9	2.4	2
	all	37	15.0	21.7	23.7	24.0	25.6	34.0	3.8	3

5 [0151] B. Categorical variables

Variable	Levels	HG %1	NV %2	NO %3	nall %all
Para	0	8 72.7	6 66.7	9 45.0	23 57.5
	1	3 27.3	2 22.2	9 45.0	14 35.0
	2	0 0.0	1 11.1	2 10.0	3 7.5
	all	11 100.0	9 100.0	20 100.0	40 100.0
Fetal sex	male	0 0.0	3 33.3	13 68.4	16 41.0
	female	11 100.0	6 66.7	6 31.6	23 59.0
	all	11 100.0	9 100.0	19 100.0	39 100.0
Education	junior high school	5 45.5	6 75.0	10 50.0	21 53.9
	High school	2 18.2	1 12.5	2 10.0	5 12.8
	College	4 36.4	1 12.5	8 40.0	13 33.3
	all	11 100.0	8 100.0	20 100.0	39 100.0
Marital status	not married/ no relationship	0 0.0	1 11.1	1 5.0	2 5.1
	married/relationship	10 100.0	8 88.9	19 95.0	37 94.9
	all	10 100.0	9 100.0	20 100.0	39 100.0
Smoking	Never	9 90.0	4 44.4	9 45.0	22 56.4
	before	1 10.0	4 44.4	11 55.0	16 41.0

at present	0	0.0	1	11.1	0	0.0	1	2.6
all	10	100.0	9	100.0	20	100.0	39	100.0

[0152] The SNPs linked to HG in the GWAS replication study were rs16982345 (closest gene *GDF15*) and rs4865234 (closest gene *IGFBP7*). The study also explored whether the SNPs were associated with medication effectiveness, HG symptoms, and whether they segregated with disease in HG families.

5 [0153] Table 4.2. Demographic characteristics of genotyped cohort.

Demographic Characteristics	
Ethnicity (% white)	87%
Average maternal birth year	1976
Average number of HG pregnancies	2
Percent of women who voluntarily terminated	15.6%
Average number of living children	2
First child (average year born)	2003
Women who had vaginal deliveries (%)	83.3%
Women who lost one or more pregnancies (%)	34.7%
Women who received partial or complete higher education (%)	83.6%

[0154] In the previous study of 254 women treated with intravenous fluids for HG, the most common antiemetic treatments were ondansetron (79%), promethazine (73%), and metoclopramide (55%). The self-reported effectiveness in relieving symptoms was approximately 50% for ondansetron, with less than 20% reporting promethazine was effective, and less than 10% reporting metoclopramide was effective. According to a recent Cochrane review, there is insufficient data on the efficacy of any of these treatments used to treat HG. It is also unknown whether genetic mechanisms may influence the effectiveness of one medication over another. By genotyping over 780 women with HG for the risk alleles in *GDF15* and *IGFBP7* and collecting data on self-reported medication effectiveness, it is possible to explore whether HG risk alleles in the genes might be used to predict effectiveness of medications used to treat HG. This allows more rapid and effective treatment, and reduces dehydration and malnutrition of pregnancy associated with refractory HG. To that end, risk alleles in HG patients were compared to medication effectiveness for ondansetron, promethazine, and metoclopramide.

[0155] Self-reported failure to respond to ondansetron was significantly linked to the *IGFBP7* HG-risk allele ($p=0.03$, $OR=4.88$, $1.14-20.92$); self-reported ineffectiveness of promethazine was also significantly linked to the *IGFBP7* HG-risk allele ($p=0.04$, $OR=2.47$, $1.04-5.89$); and self-reported response to metoclopramide was significantly linked to the *GDF15* HG-risk allele

($p=0.02$, $OR=0.37$ 0.165-0.833). The results were not corrected for multiple testing and therefore should be viewed as exploratory.

[0156] Among 277 women with HG that were genotyped and provided information on duration of symptoms, women with an AA genotype in rs4865234 were significantly more likely to have prolonged symptoms compared to the women carrying the G allele (AG+GG) ($p=0.0106$) (Table 3). There was no between-group difference for rs16982345.

[0157] Among 547 women with HG genotyped, 41% (223 HG women) reported that they experienced hypersalivation while 59% (324 HG women) reported that they did not experience ptyalism. Genotypes for rs16982345 and rs4865234 were not significantly different between the groups.

[0158] Weight loss due to HG was self-reported by survey in 259 women that were genotyped ranging from 5% to 36% of their pre-pregnancy weight. We divided the groups depending on whether a woman lost 5-10% of her bodyweight versus 11%-36% of her bodyweight during an HG pregnancy. Forty-six percent (119 HG women) lost 5-10% and 54% (140 HG women) lost 11%-36% of their bodyweight. Genotypes for rs16982345 and rs4865234 were not significantly different between the groups. (Table 4.3)

[0159] Table 4.3. rs4865234 and rs16982345 genotypes associated (unadjusted p -value <0.05) with medications and symptoms

rs4865234					
Total N reporting duration	Long AA (%)	Long GG+AG (%)	p-value	OR	95% CI
277	127/138 (92%)	113/139 (81%)	0.0100	2.66	[1.26, 5.62]
Total N reporting Zofran	Effective GG (%)	Effective AA+AG (%)	p-value	OR	95% CI
333	29/31 (94%)	226/302 (75%)	0.0330	4.88	[1.14, 20.92]
Total N reporting Phenergan	Effective GG (%)	Effective AA+AG (%)	p-value	OR	95% CI
314	10/23 (43%)	69/291 (24%)	0.0406	2.47	[1.04, 5.89]
rs16982345					
Total N reporting Reglan	Effective GG (%)	Effective AA+AG (%)	p-value	OR	95% CI
266	39/180 (22%)	8/86 (9%)	0.0163	2.7	[1.20, 6.06]

[0160] HG patients were divided into two groups depending on whether they had a recurrence or not. A recurrence was defined as having HG treated with IV fluids in two pregnancies

(N=331). A nonrecurring pregnancy was defined as no IV fluids required in the second pregnancy even if NVP symptoms were present, where HG was treated with IV fluids in the first pregnancy (N=225). The SNPs tested (rs16982345 and rs4865234) were not predictive of recurrence.

5 [0161] Finally, the study explored whether risk alleles segregated with disease in families. Five families with 3 family members whose pregnancies were affected by HG and one family member with normal NVP were genotyped to determine whether the risk alleles segregate with disease. All three affected individuals in Family 1 carried at least one risk allele (G) for rs16982345, while the unaffected family member had the AA genotype (Figure 6A). Risk alleles
10 in rs16982345 and rs4865234 did not segregate with disease in the remaining 4 families (the unaffected family member did not have a different genotype that was shared by all 3 affected individuals). However, two SNPs associated with increased expression of GDF15 in previous studies, segregated with disease in family 2 (rs1054564), and family 3 (rs1055150) (Figure 6B and 6C). In summary, in three of five families, the affected family members carry genotypes that
15 are linked to altered GDF15 levels, while the unaffected family members do not.

[0162] The first GWAS of HG identified two candidate genes, *GDF15* and *IGFBP7*, but causality may only be hypothesized as functional studies are prohibitory in pregnant women. This Example provides further evidence of a role for these loci by showing that GDF15, and also IGFBP7, are dysregulated in HG pregnancies, while beta HCG, which was the leading
20 hypothesis for the cause of HG, is not. Additional supportive evidence comes from previous studies showing both GDF15 and IGFBP7 are upregulated in early pregnancy when HG occurs, play critical roles in placentation, decrease prior to miscarriage, and are drivers of stress-induced feeding behavior and cachexia, a disease characterized by symptoms similar to HG: nausea, weight loss, and muscle wasting. While a 2.7 ng/μl increase in mean serum GDF15
25 levels (5.4 ng/μl cachexia vs 2.7 ng/μl controls) is suspected to initiate cancer, herein an increase in mean serum levels of 5.4 ng/ul (15.8 ng/μl HG vs 10.4 ng/μl NVP) was detected in patients hospitalized for HG compared to normal NVP. Interestingly, serum concentrations of GDF15 in cancer patients with cachexia are significantly increased, even in the pre-cachectic state, and remain relatively steady thereafter. Therefore, it is now imperative to determine
30 whether the same holds true for HG. Levels of GDF15 (and IGFBP7) must be studied in early pregnancy, to determine whether they can be used to predict an HG pregnancy. In addition, the present findings linking SNPs to medication effectiveness, prolonged symptoms, and familial HG can provide a basis for individualized therapy. Furthermore, the brainstem-restricted receptor for GDF15, GFRAL, was recently identified and the pathway is currently under intense
35 investigation to identify novel therapeutic strategies for weight gain (and weight loss), which may be applicable to HG. Of particular clinical interest, inhibition of GDF15 restored appetite and

weight gain in a mouse model of cancer cachexia. Finally, the findings herein suggest an answer to an age-old paradox. HG can lead to prolonged dehydration and undernutrition, which can be detrimental to maternal and fetal health and can decrease reproductive fitness. The dual roles of GDF15 (and possibly IGFBP7) in maintaining pregnancy, and in increasing the risk of HG, may provide a molecular explanation for why NVP still exists in nature.

Example 5: Combinations of Genotypes and Their Frequencies

[0163] This Example analyzes the frequencies of various genotype combinations. The genes analyzed were (1) GDF15, corresponding to SNP rs16982345; (2) IGFBP7, corresponding to SNP rs4865234; (3) PGR, corresponding to SNP rs7948518; (4) RYR2; and (5) GFRAL, corresponding to SNP rs7761177. The data in the table below shows the significantly increased risk of HG for certain allele combinations. Inhibitors of GDF15, GRAL, PGR can be selected as medications to block the corresponding gene product if a patient is positive for those variants.

Frequencies of Genotype Combinations between genes 1 and 2

	Control (No HG) n=560	HG/TPN n=737	Odds ratio	95 % CI	Significance
AA-AA	18 (3.2%)	11 (1.5%)	0.4977	0.2377 to 1.0420	P = 0.0642
AA-AG	19 (3.4%)	7 (0.9%)	0.2727	0.1138 to 0.6532	P = 0.0036
AA-GG	2 (0.4%)	1 (0.1%)	0.3786	0.0342 to 4.1856	P = 0.4282
AG-AA	89 (15.9%)	126 (17.1%)	1.0896	0.8098 to 1.4660	P = 0.5711
AG-AG	87 (15.5%)	84 (11.4%)	0.6983	0.5059 to 0.9639	P = 0.0290
AG-GG	28 (5.0%)	11 (1.5%)	0.2836	0.1399 to 0.5747	P = 0.0005
GG-AA	132 (23.6%)	249 (33.8%)	1.6511	1.2888 to 2.1152	P = 0.0001
GG-AG	147 (26.3%)	202 (27.4%)	1.0588	0.8261 to 1.3571	P = 0.6518
GG-GG	38 (6.8%)	46 (6.2%)	0.9131	0.5855 to 1.4241	P = 0.6886

Frequencies of Genotype Combinations between genes 1 and 3

	Control (No HG) n=566	HG/TPN n=734	Odds ratio	95 % CI	Significance
AA-CT	11 (1.9%)	8 (1.1%)	0.556	0.2221 to 1.3915	P = 0.2098
AA-TT	26 (4.6%)	9 (1.2%)	0.2578	0.1198 to 0.5547	P = 0.0005
AA-CC	1 (0.2%)	1 (0.1%)	0.7708	0.0481 to 12.3507	P = 0.8541
AG-CT	63 (11.1%)	95 (12.9%)	1.187	0.8455 to 1.6664	P = 0.3219
AG-TT	129 (22.8%)	110 (15.0%)	0.5972	0.4503 to 0.7919	P = 0.0003
AG-CC	13 (2.3%)	16 (2.2%)	0.9479	0.4522 to 1.9872	P = 0.8874
GG-CT	90 (15.9%)	216 (29.4%)	2.2054	1.6742 to 2.9052	P < 0.0001
GG-TT	222 (39.2%)	254 (34.6%)	0.8200	0.6533 to 1.0291	P = 0.0868
GG-CC	11 (1.9%)	25 (3.4%)	1.7791	0.8679 to 3.6470	P = 0.1157

Frequencies of Genotype Combinations between genes 1 and 4

	Control (No HG) n=405	HG/TPN n=559	Odds ratio	95 % CI	Significance
--	-----------------------	--------------	------------	---------	--------------

AA-AA 0 (0.0%)	0 (0.0%)	0.7248	0.0143 to 36.6041	P = 0.8722
AA-AG 0 (0.0%)	0 (0.0%)	0.7248	0.0143 to 36.6041	P = 0.8722
AA-GG 29 (7.2%)	15 (2.7%)	0.3575	0.1891 to 0.6760	P = 0.0016
AG-AA 0 (0.0%)	0 (0.0%)	0.7248	0.0143 to 36.6041	P = 0.8722
AG-AG 5 (1.2%)	15 (2.7%)	2.2059	0.7952 to 6.1194	P = 0.1286
AG-GG 145 (35.8%)	144 (25.8%)	0.6222	0.4712 to 0.8215	P = 0.0008
GG-AA 0 (0.0%)	0 (0.0%)	0.7248	0.0143 to 36.6041	P = 0.8722
GG-AG 11 (2.7%)	22 (3.9%)	1.4674	0.7034 to 3.0614	P = 0.3067
GG-GG 215 (53.1%)	363 (64.9%)	1.6367	1.2603 to 2.1254	P = 0.0002

Frequencies of Genotype Combinations between genes 1 and 5

	Control (No HG) n=551	HG/TPN n=715	Odds ratio	95 % CI	Significance
AA-CT	17 (3.1%)	7 (1.0%)	0.3106	0.1279 to 0.7543	P = 0.0098
AA-TT	13 (2.4%)	7 (1.0%)	0.4092	0.1621 to 1.0326	P = 0.0585
AA-CC	6 (1.1%)	5 (0.7%)	0.6397	0.1942 to 2.1070	P = 0.4626
AG-CT	103 (18.7%)	117 (16.4%)	0.8510	0.6356 to 1.1393	P = 0.2784
AG-TT	48 (8.7%)	63 (8.8%)	1.0126	0.6834 to 1.5002	P = 0.9504
AG-CC	51 (9.3%)	31 (4.3%)	0.4443	0.2802 to 0.7046	P = 0.0006
GG-CT	167 (30.3%)	228 (31.9%)	1.0765	0.8466 to 1.3689	P = 0.5476
GG-TT	74 (13.4%)	163 (22.8%)	1.7599	1.3039 to 2.3756	P = 0.0002
GG-CC	72 (13.1%)	94 (13.1%)	1.0070	0.7245 to 1.3997	P = 0.9668

Frequencies of Genotype Combinations between genes 2 and 3

	Control (No HG) n=585	HG/TPN n=760	Odds ratio	95 % CI	Significance
AA-CT	76 (13.0%)	174 (22.9%)	1.9886	1.4813 to 2.6696	P < 0.0001
AA-TT	165 (28.2%)	202 (26.6%)	0.9215	0.7238 to 1.1732	P = 0.5069
AA-CC	12 (2.1%)	22 (2.9%)	1.4234	0.6986 to 2.9005	P = 0.3309
AG-CT	73 (12.5%)	130 (17.1%)	1.4473	1.0621 to 1.9722	P = 0.0192
AG-TT	182 (31.1%)	157 (20.7%)	0.5722	0.4465 to 0.7334	P < 0.0001
AG-CC	7 (1.2%)	15 (2.0%)	1.6625	0.6734 to 4.1044	P = 0.2703
GG-CT	22 (3.8%)	24 (3.2%)	0.8345	0.4631 to 1.5037	P = 0.5470
GG-TT	44 (7.5%)	32 (4.2%)	0.5405	0.3382 to 0.8637	P = 0.0101
GG-CC	4 (0.7%)	4 (0.5%)	0.7685	0.1914 to 3.0859	P = 0.7105

Frequencies of Genotype Combinations between genes 2 and 4

	Control (No HG) n=423	HG/TPN n=579	Odds ratio	95 % CI	Significance
AA-AA	0 (0.0%)	0 (0.0%)	0.7308	0.0145 to 36.9066	P = 0.8755
AA-AG	9 (2.1%)	22 (3.8%)	1.8169	0.8280 to 3.9869	P = 0.1364
AA-GG	181 (42.8%)	291 (50.3%)	1.3509	1.0497 to 1.7386	P = 0.0194
AG-AA	0 (0.0%)	0 (0.0%)	0.7308	0.0145 to 36.9066	P = 0.8755
AG-AG	7 (1.7%)	16 (2.8%)	1.7295	0.7052 to 4.2415	P = 0.2313
AG-GG	173 (40.9%)	207 (35.8%)	0.8041	0.6214 to 1.0406	P = 0.0974
GG-AA	0 (0.0%)	0 (0.0%)	0.7308	0.0145 to 36.9066	P = 0.8755

GG-AG 1 (0.2%)	1 (0.2%)	0.7301	0.0455 to 11.7063	P = 0.8242
GG-GG 52 (12.3%)	42 (7.3%)	0.5580	0.3639 to 0.8558	P = 0.0075

Frequencies of Genotype Combinations between genes 2 and 5

	Control (No HG) n=569	HG/TPN n=738	Odds ratio	95 % CI	Significance
AA-CT	130 (22.8%)	195 (26.4%)	1.2127	0.9397 to 1.5651	P = 0.1384
AA-TT	64 (11.2%)	132 (17.9%)	1.7188	1.2468 to 2.3694	P = 0.0009
AA-CC	52 (9.1%)	61 (8.3%)	0.8958	0.6082 to 1.3195	P = 0.5777
AG-CT	132 (23.2%)	134 (18.2%)	0.7345	0.5607 to 0.9622	P = 0.0251
AG-TT	64 (11.2%)	94 (12.7%)	1.1517	0.8211 to 1.6154	P = 0.4131
AG-CC	56 (9.8%)	65 (8.8%)	0.8848	0.6079 to 1.2878	P = 0.5226
GG-CT	39 (6.9%)	32 (4.3%)	0.6160	0.3808 to 0.9963	P = 0.0483
GG-TT	9 (1.6%)	15 (2.0%)	1.2909	0.5608 to 2.9716	P = 0.5483
GG-CC	23 (4.0%)	10 (1.4%)	0.3261	0.1539 to 0.6908	P = 0.0034

Frequencies of Genotype Combinations between genes 3 and 4

	Control (No HG) n=428	HG/TPN n=580	Odds ratio	95 % CI	Significance
CT-AA	0 (0.0%)	0 (0.0%)	0.7382	0.0146 to 37.2774	P = 0.8794
CT-AG	6 (1.4%)	19 (3.3%)	2.3821	0.9431 to 6.0163	P = 0.0663
CT-GG	122 (28.5%)	237 (40.9%)	1.7331	1.3266 to 2.2641	P = 0.0001
TT-AA	0 (0.0%)	0 (0.0%)	0.7382	0.0146 to 37.2774	P = 0.8794
TT-AG	11 (2.6%)	19 (3.3%)	1.2839	0.6045 to 2.7271	P = 0.5156
TT-GG	273 (63.8%)	280 (48.3%)	0.5299	0.4104 to 0.6843	P < 0.0001
CC-AA	0 (0.0%)	0 (0.0%)	0.7382	0.0146 to 37.2774	P = 0.8794
CC-AG	0 (0.0%)	1 (0.2%)	2.2183	0.0901 to 54.5894	P = 0.6259
CC-GG	16 (3.7%)	24 (4.1%)	1.1115	0.5830 to 2.1191	P = 0.7481

Frequencies of Genotype Combinations between genes 3 and 5

	Control (No HG) n=576	HG/TPN n=734	Odds ratio	95 % CI	Significance
CT-CT	83 (14.4%)	164 (22.3%)	1.709	1.2785 to 2.2845	P = 0.0003
CT-TT	45 (7.8%)	111 (15.1%)	2.1024	1.4592 to 3.0292	P = 0.0001
CT-CC	39 (6.8%)	50 (6.8%)	1.0065	0.6523 to 1.5530	P = 0.9766
TT-CT	207 (35.9%)	175 (23.8%)	0.5581	0.4388 to 0.7097	P < 0.0001
TT-TT	87 (15.1%)	114 (15.5%)	1.0335	0.7632 to 1.3995	P = 0.8314
TT-CC	91 (15.8%)	80 (10.9%)	0.6519	0.4722 to 0.9001	P = 0.0093
CC-CT	12 (2.1%)	19 (2.6%)	1.2490	0.6012 to 2.5945	P = 0.5512
CC-TT	8 (1.4%)	15 (2.0%)	1.4812	0.6236 to 3.5183	P = 0.3734
CC-CC	4 (0.7%)	6 (0.8%)	1.1786	0.3310 to 4.1963	P = 0.7998

Frequencies of Genotype Combinations between genes 4 and 5

	Control (No HG) n=422	HG/TPN n=567	Odds ratio	95 % CI	Significance
AA-CT	0 (0.0%)	0 (0.0%)	0.7445	0.0147 to 37.5988	P = 0.8828

AA-TT 0 (0.0%)	0 (0.0%)	0.7445	0.0147 to 37.5988	P = 0.8828
AA-CC 0 (0.0%)	0 (0.0%)	0.7445	0.0147 to 37.5988	P = 0.8828
AG-CT 4 (0.9%)	20 (3.5%)	3.8208	1.2962 to 11.2631	P = 0.0151
AG-TT 5 (1.2%)	11 (1.9%)	1.6500	0.5689 to 4.7852	P = 0.3566
AG-CC 8 (1.9%)	7 (1.2%)	0.6469	0.2327 to 1.7981	P = 0.4036
GG-CT 217 (51.4%)	268 (47.3%)	0.8468	0.6580 to 1.0897	P = 0.1962
GG-TT 102 (24.2%)	163 (28.7%)	1.2658	0.9494 to 1.6877	P = 0.1083
GG-CC 86 (20.4%)	98 (17.3%)	0.8164	0.5918 to 1.1262	P = 0.2165

G/G (gene 1), A/A (gene 2), C/T (gene 5)

Control (No HG) (n=541)	HG/TPN (n=711)	Odds ratio	95% CI	Significance
68 (12.6%)	124 (17.4%)	1.4694	1.0674 to 2.0227	P = 0.0183

G/G (gene 1), C/T (gene 3), C/T (gene 5)

Control (No HG) (n=547)	HG/TPN (n=708)	Odds ratio	95% CI	Significance
43 (7.9%)	106 (15.0%)	2.0638	1.4204 to 2.9987	P = 0.0001

A/A (gene 2), C/T (gene 3), C/T (gene 5)

Control (No HG) (n=565)	HG/TPN (n=732)	Odds ratio	95% CI	Significance
37 (6.5%)	91 (12.4%)	2.0259	1.3592 to 3.0197	P = 0.0005

G/G (gene 1), A/A (gene 2), C/T (gene 3), C/T (gene 5)

Control (No HG) (n=538)	HG/TPN (n=705)	Odds ratio	95% CI	Significance
18 (3.3%)	59 (8.4%)	2.6385	1.5372 to 4.5287	P = 0.0004

G/G (gene 1), A/A (gene 2), T/T (gene 5)

Control (No HG) (n=541)	HG/TPN (n=711)	Odds ratio	95% CI	Significance
33 (6.1%)	80 (11.3%)	1.9517	1.2798 to 2.9763	P = 0.0019

G/G (gene 1), C/T (gene 3), T/T (gene 5)

Control (No HG) (n=547)	HG/TPN (n=708)	Odds ratio	95% CI	Significance
21 (3.8%)	72 (10.2%)	2.8356	1.7205 to 4.6734	P < 0.0001

A/A (gene 2), C/T (gene 3), T/T (gene 5)

Control (No HG) (n=565)	HG/TPN (n=732)	Odds ratio	95% CI	Significance
21 (3.7%)	61 (8.3%)	2.355	1.4161 to 3.9163	P = 0.0010

G/G (gene 1), A/A (gene 2), C/T (gene 3), T/T (gene 5)

Control (No HG) (n=538)	HG/TPN (n=705)	Odds ratio	95% CI	Significance
8 (1.5%)	40 (5.7%)	3.985	1.8495 to 8.5863	P = 0.0004

Gene	gene	SNP
1	GDF15	rs16982345
2	IGFBP7	rs4865234
3	PGR	rs7948518
4	RYR2	See Example 6
5	GFRAL	rs7761177

Example 6: Genetic analysis of hyperemesis gravidarum reveals association with intracellular calcium release channel (RYR2)

This Example identifies a link between HG and RYR2. Whole-exome sequencing of 5 families was performed and followed by analysis of variants in 584 cases/431 controls. Variants in RYR2 segregated with disease in 2 families. The novel variant L3277R was not found in any case/control. The rare variant, G1886S was more common in cases ($p = 0.046$) and extreme cases ($p = 0.023$). Replication of G1886S using Norwegian/Australian data was supportive. Common variants rs790899 and rs1891246 were significantly associated with HG and weight loss. Copy-number analysis revealed a deletion in a patient. RYR2 encodes an intracellular calcium release channel involved in vomiting, cyclic-vomiting syndrome, and is a thyroid hormone target gene. Additionally, RYR2 is a downstream drug target of Inderal, used to treat HG and CVS. These results provide genetic evidence for a pathway and therapy for HG. More details relating to this example can be found in Fejzo et al., Molecular and Cellular Endocrinology, Volume 439, 5 January 2017, 308-316.

[0164] Throughout this application various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to describe more fully the state of the art to which this invention pertains.

[0165] Those skilled in the art will appreciate that the conceptions and specific embodiments disclosed in the foregoing description may be readily utilized as a basis for modifying or designing other embodiments for carrying out the same purposes of the present invention. Those skilled in the art will also appreciate that such equivalent embodiments do not depart from the spirit and scope of the invention as set forth in the appended claims.

What is claimed is:

1. A method of analyzing genetic expression in a sample obtained from a subject, the
5 method comprising:
 - (a) obtaining a sample from the subject;
 - (b) measuring the level of expression product of at least one of the hyperemesis
gravidarum (HG) outlook genes identified in Table 1 in the sample;
 - (c) assigning a HG outlook score to the measured amount of expression for each of the
10 HG outlook genes, wherein the score is between 0 and 1;
 - (d) calculating a HG status score, wherein the status score is equal to or representative
of the combined outlook score(s) assigned in (c).
2. The method of claim 1, wherein at least two of the HG outlook genes of Table 1 are
measured.
- 15 3. The method of claim 1, wherein the measuring comprises contacting the sample
obtained from the subject with reagents for amplifying and/or detecting polynucleotides.
4. The method of claim 1, wherein the level of expression is measured via hybridization of
probes that specifically bind polynucleotides encoding the expression product.
5. The method of claim 1, wherein the level of expression is measured via binding of
20 antibodies that specifically bind the expression product.
6. The method of claim 1, wherein the sample is a urine sample, a blood sample, or a
saliva sample.
7. The method of claim 1, further comprising treating the subject for nausea/vomiting of
pregnancy (NVP) or hyperemesis gravidum (HG) if the HG status score is greater than zero.
- 25 8. The method of claim 1, wherein the HG outlook gene is GDF15, IGFBP7, PGR, SYN3,
MMADHC, TMEM38B, HCRT2, GFRAL, RYR2 and/or PKHD1.
9. The method of claim 8, wherein the expression is measured via detection of the G allele
at rs16982345, or any alleles in a shared haplotype with the G allele at rs16982345, of GDF15.
10. The method of claim 7, wherein the gene is GDF15 and the treatment comprises
30 administering an anti-GDF15 agent to the subject.
11. The method of claim 7, wherein the gene is MMADHC and the treatment comprises
administering vitamin B12 to the subject.
12. The method of claim 1, further comprising measuring the expression product of an
additional HG outlook gene, wherein the additional HG outlook gene is selected from: GDF15,
35 IGFBP7, PGR, and RYR2.

13. A method of treating a subject for NVP or HG, the method comprising:
- (a) performing the analyzing of any one of claims 1 to 12; and
 - (b) administering a treatment for NVP or HG to the subject if the analyzing results in a status score that is greater than zero.
- 5 14. The method of claim 13, wherein the treatment comprises an antiemetic agent.
15. A method of analyzing a biological sample, the method comprising:
- (a) measuring the amount of GDF15 in the sample; and
 - (b) measuring the amount of IGFBP7 in the sample.
- 10 16. The method of claim 15, wherein the analyzing further comprises detecting whether the amount of GDF15 is greater than 10 ng/ml, and the amount of IGFBP7 is greater than 60 ng/ml.
17. The method of claim 15 or 16, wherein the sample is obtained from a pregnant subject.
18. The method of claim 17, wherein the subject is treated for HG or NVP if the amount of GDF15 is greater than about 10 ng/ml, and the amount of IGFBP7 is greater than about 60 ng/ml.
- 15 19. A method of treating a subject for NVP or HG, the method comprising:
- (a) performing the analyzing of claim 17; and
 - (b) administering a treatment for NVP or HG to the subject if the amount of GDF15 is greater than 10 ng/ml, and the amount of IGFBP7 is greater than about 60 ng/ml.
20. A method of treating a subject for NVP or HG, the method comprising:
- (a) analyzing a sample obtained from the subject for presence of the G allele in rs16982345 (positive for the *GDF15* HG-risk allele) and/or the GG genotype at rs4865234 (lacking the *IGFBP7* HG-risk allele); and
 - (b) administering ondansetron or promethazine if the patient is lacking the *IGFBP7* HG-risk allele, and treating a patient with metoclopramide if the patient is positive for the *GDF15* HG-risk allele.
- 25 21. A method of treating a subject for HG, the method comprising:
- (a) analyzing a sample obtained from the subject for presence of one or more of the genotype combinations identified in Example 5 as associated with HG at a significance of $P \leq 0.05$; and
 - (b) administering an inhibitor of a corresponding gene if the patient is positive for the
- 30 significant genotype combination.

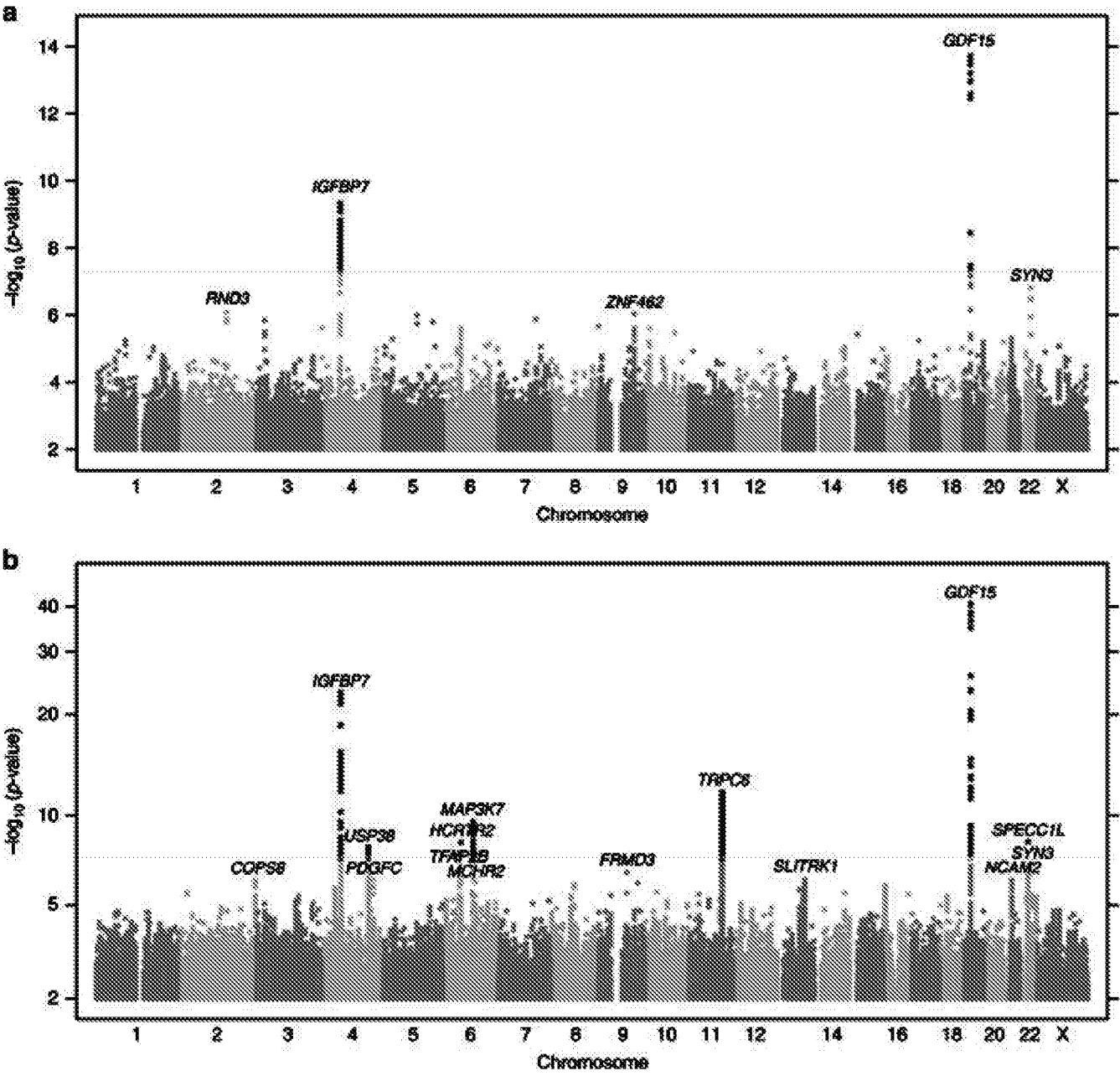


FIGURE 1

Protein 1 vs. Protein 2

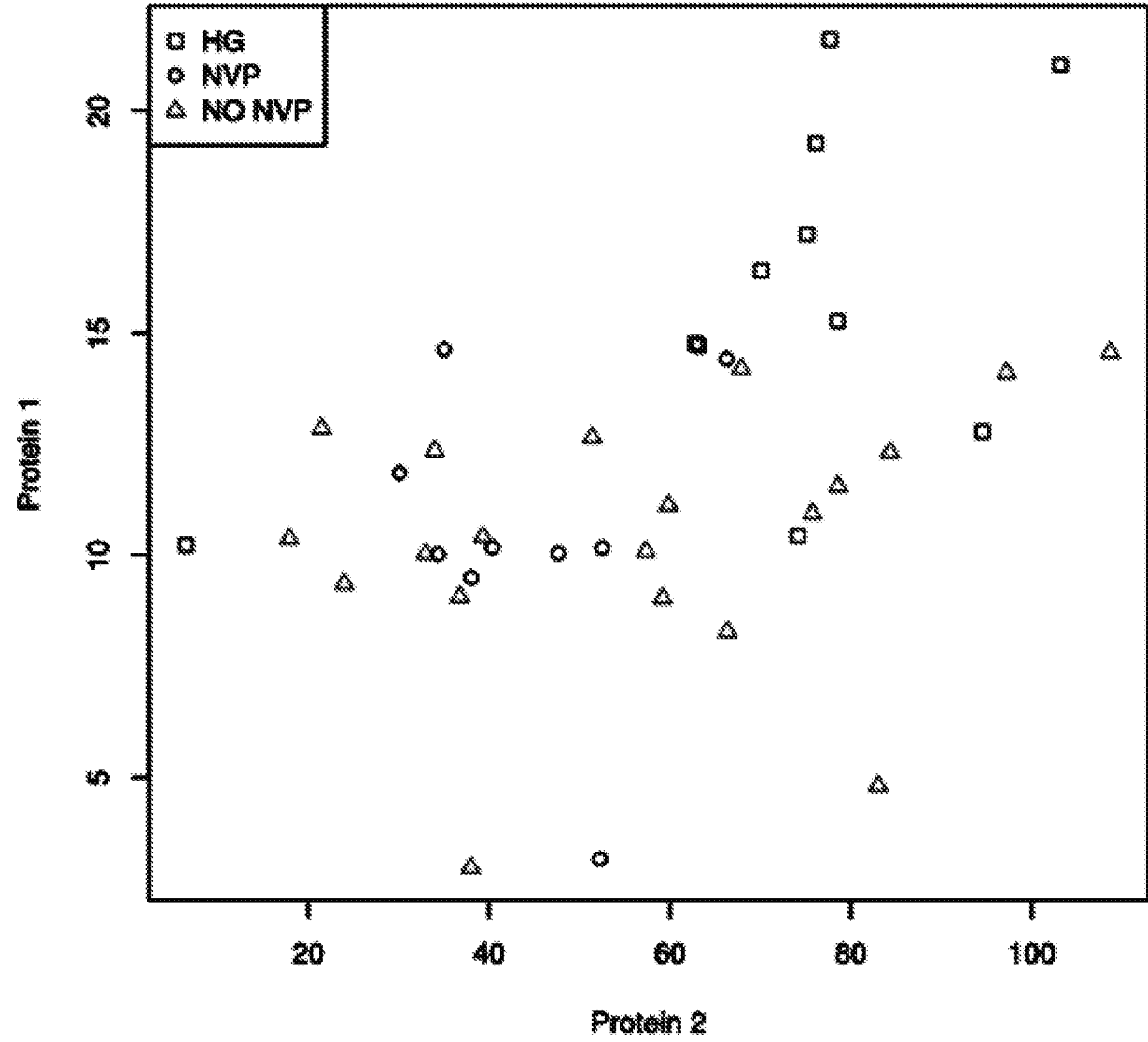


FIGURE 2

3/9

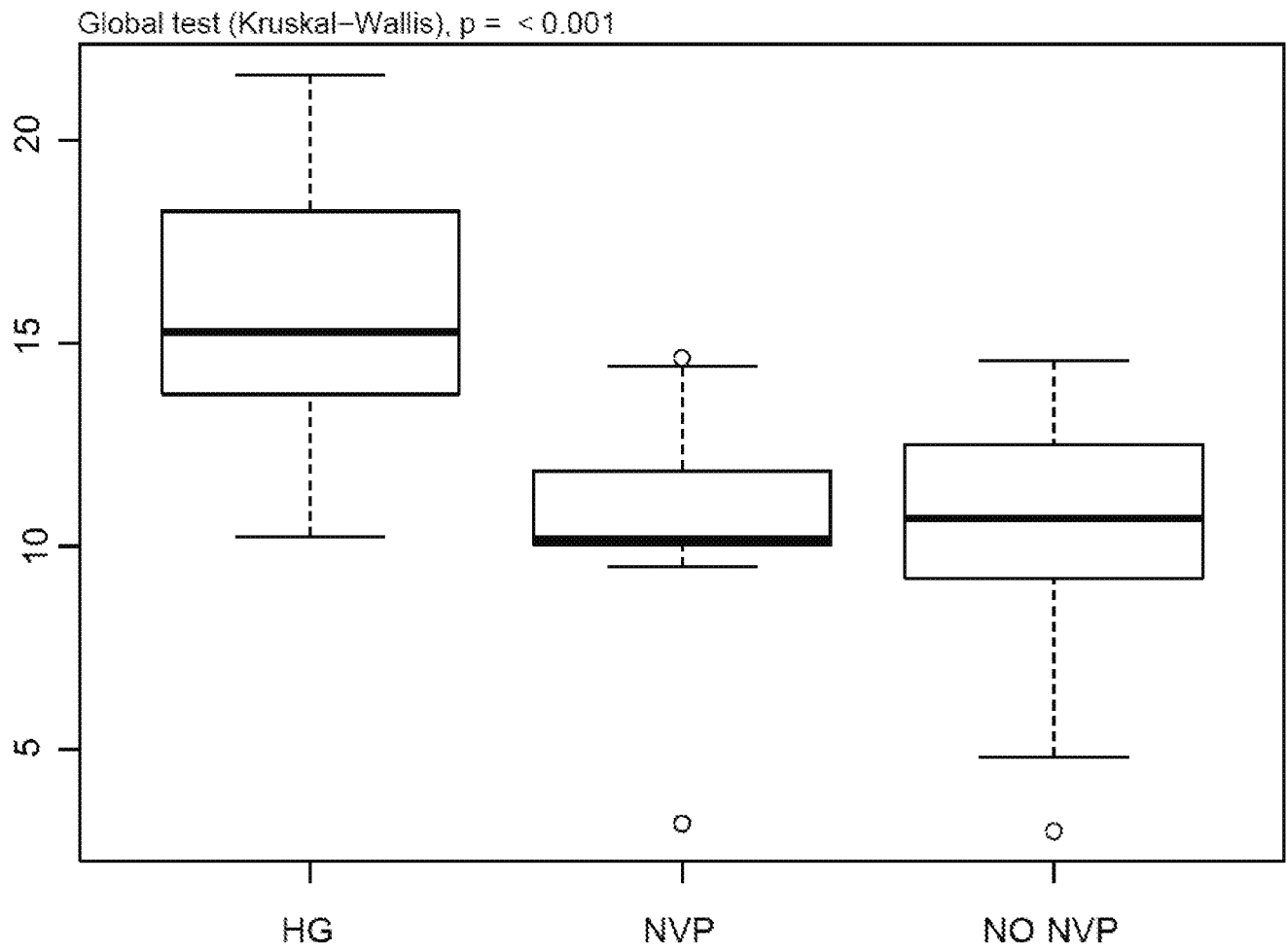


FIGURE 3

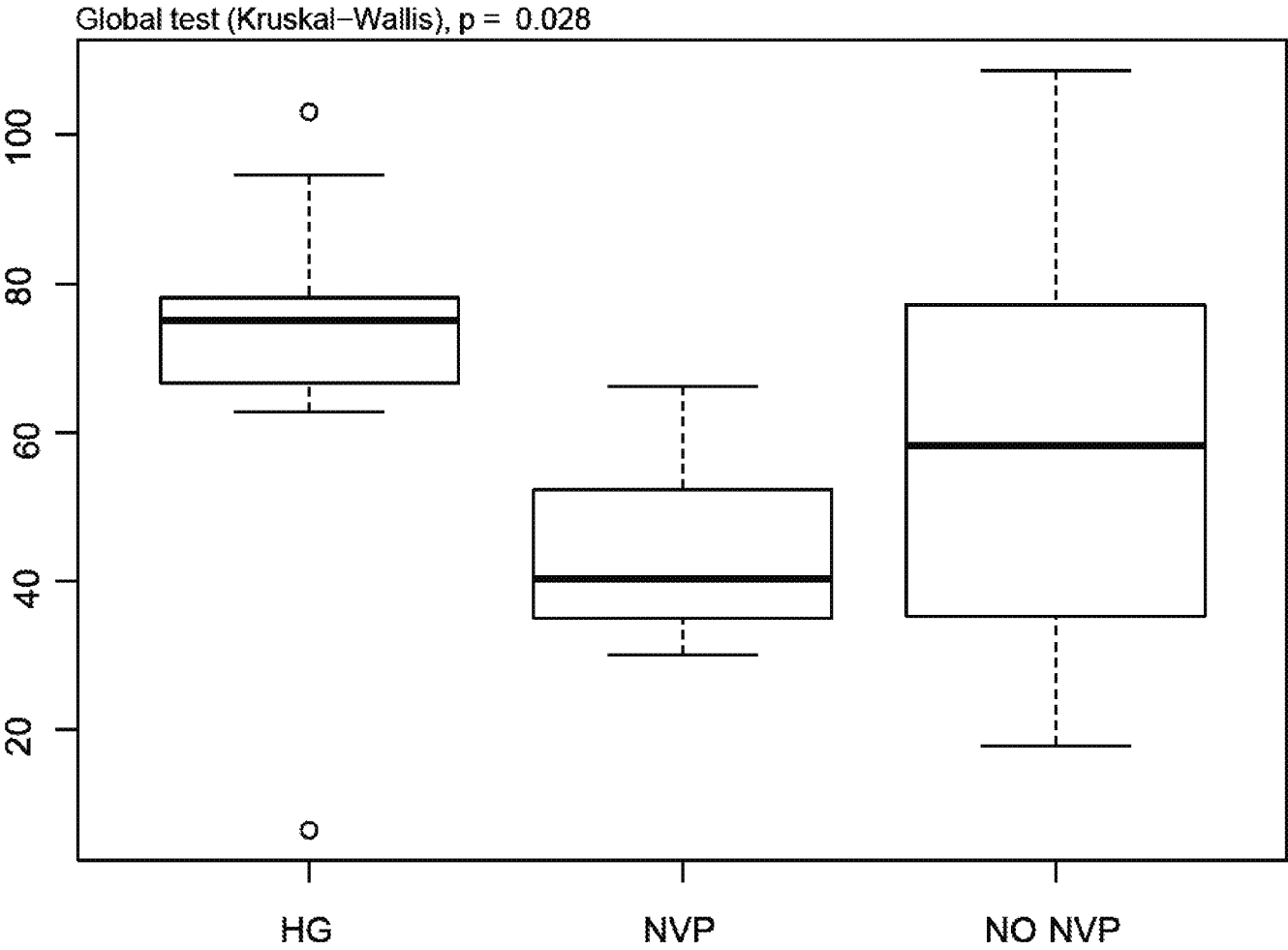


FIGURE 4

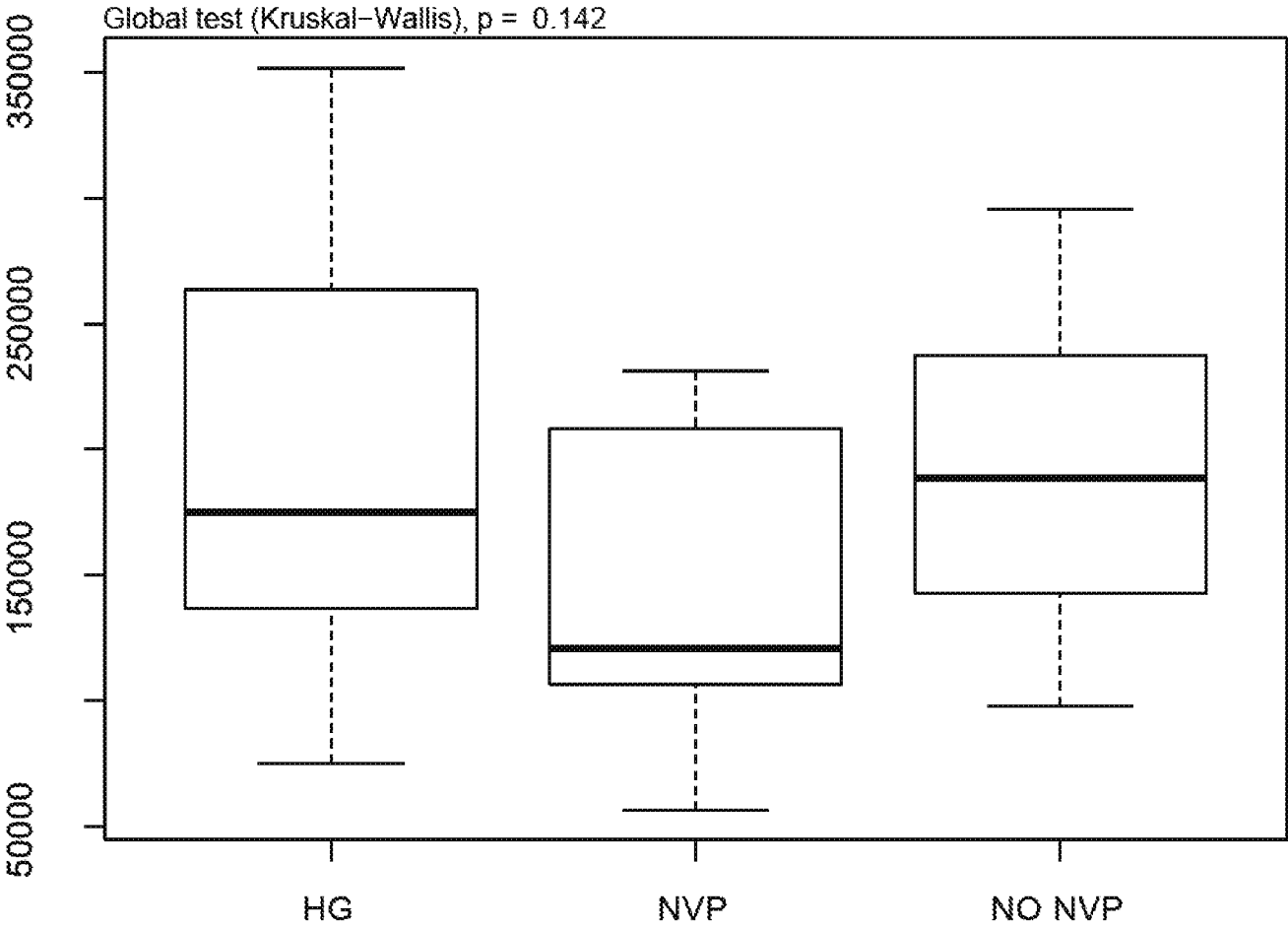


FIGURE 5

6/9

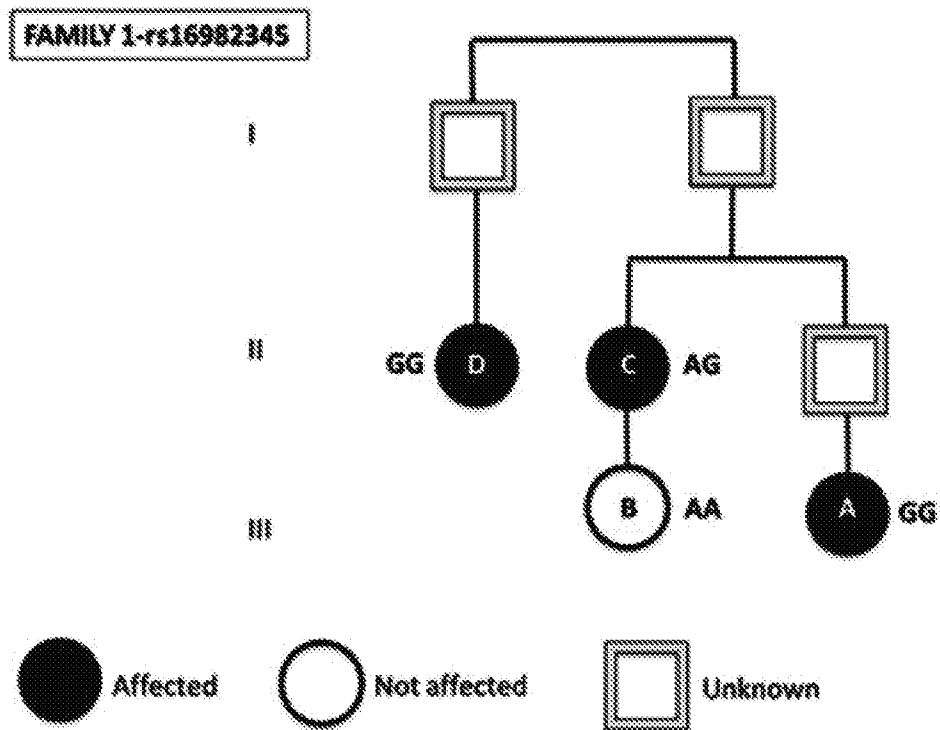


FIGURE 6A

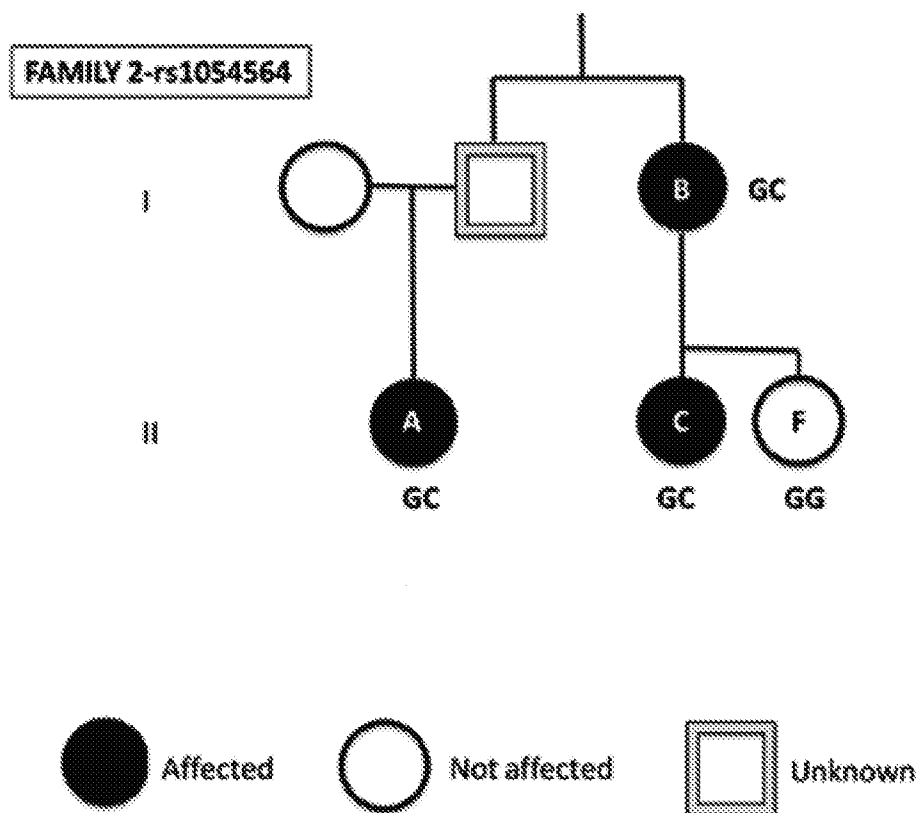


FIGURE 6B

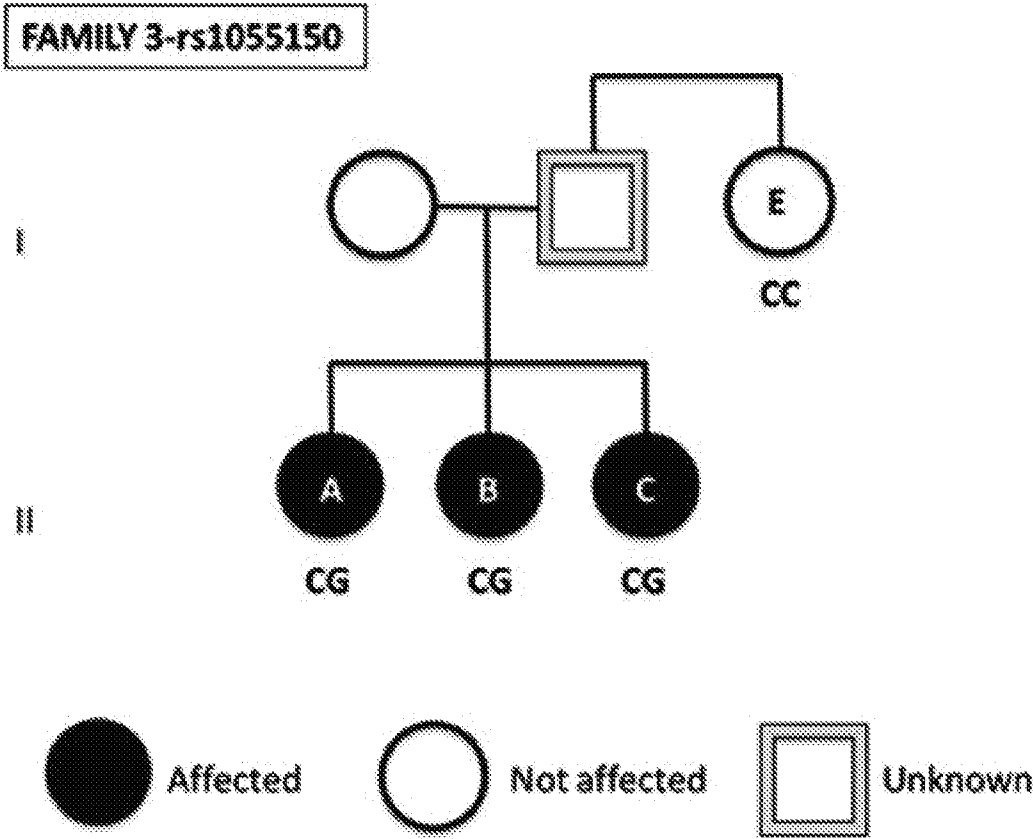


FIGURE 6C

Table 1.2

Cytoband	Marker	Position	Alleles	MAF	SCAN1 p-value	SCAN1 OR	SCAN1 95% CI	SCAN2 p-value	SCAN2 effect size	SCAN2 95% CI	Gene context
19p13.11	rs45543339	19:18503194	C/T	0.25	1.9×10^{-14}	0.668	[0.601,0.743]	5.7×10^{-38}	-0.104	[-0.120, -0.890]	[LRRC25]
4q12	rs143409503	4:58351064	-/AGC	0.34	4.5×10^{-10}	0.752	[0.687,0.824]	6.8×10^{-25}	-0.073	[-0.087, -0.0587]	[GFBP7-[]]
22q12.3	rs5998706	22:33440890	C/T	0.49	1.6×10^{-7}	0.786	[0.717,0.860]	3.5×10^{-5}	-0.03	[-0.044, -0.016]	[SYN3]
2q23.3	rs114571265	2:150910809	G/C	0.024	8.3×10^{-7}	0.388	[0.253,0.595]	3.6×10^{-4}	-0.087	[-0.133, -0.039]	MMADHC- []---RND3
9q31.2	rs56108151	9:109514426	A/G	0.095	9.2×10^{-7}	0.677	[0.575,0.797]	5.9×10^{-4}	-0.034	[-0.062, -0.017]	TMEM38B- []---ZNF462

FIGURE 7

Table 1.4

Cytoband	Marker	Position	Alleles	MAF	p-value	β	95% CI	SCAN1 p-value	SCAN1 OR	SCAN1 95% CI	Gene context
19p13.11	rs16982345	19:18500722	G/A	0.25	2.4×10^{-41}	-0.104	[-0.120, -0.089]	6.81×10^{-16}	0.68	[0.615, 0.751]	GDF15]-LRRC25
4q12	rs143409503	4:56351064	-/AGC	0.34	9.2×10^{-24}	-0.072	[-0.086, -0.058]	1.86×10^{-10}	0.752	[0.688, 0.822]	IGFBP7---[]
11q22.1	rs2508362	11:101260798	A/G	0.199	1.7×10^{-12}	0.06	[0.043, 0.076]	1.06×10^{-04}	0.819	[0.741, 0.904]	PGR---[]-TRPC6
6q15	rs4707680	6:92294371	T/C	0.427	2.8×10^{-10}	0.043	[0.030, 0.056]	5.79×10^{-02}	0.924	[0.851, 1.00]	MAP3K7-[]
22q11.23	rs201838815	22:24753023	T/-	0.02	5.8×10^{-9}	-0.159	[-0.212, -0.105]	2.88×10^{-02}	0.685	[0.480, 0.979]	[SPECCL,SPECCIADORA2A]
6p12.1	rs7761177	6:55151343	T/C	0.495	6.2×10^{-9}	-0.039	[-0.052, -0.026]	1.70×10^{-06}	0.82	[0.755, 0.889]	HCRT2-[]-GFRAL
4q31.21	rs4690766	4:144030040	G/C	0.358	1.3×10^{-8}	-0.04	[-0.054, -0.026]	1.02×10^{-01}	0.932	[0.856, 1.01]	INPP4B---[]-USP38

FIGURE 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/26588

A. CLASSIFICATION OF SUBJECT MATTER

IPC - A61K 38/195; C07K 16/24; C12Q 1/6883; 1/6886 (2018.01)

CPC - A61K 38/195; C07K 16/24; C12Q 1/6883; 1/6886; G01N 33/6887

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	US 2014/0065648 A1 (ROCHE DIAGNOSTICS OPERATIONS INC) March 6, 2014; abstract	15 ---
Y	WO 2001/81928 A1 (ST VINCENT'S HOSPITAL SYDNEY LIMITED) November 1, 2001; page 21, lines 18-23	16, 17/15-16
Y	WO 2001/81928 A1 (ST VINCENT'S HOSPITAL SYDNEY LIMITED) November 1, 2001; page 21, lines 18-23	16, 17/15-16
A	US 2016/0258018 A1 (CHINESE UNIV HONG KONG NT) September 8, 2016; paragraphs [0008], [0049]; claim 1	1-12, 13/1-12, 14/13/1-12
A	US 2016/0048632 A1 (GENOMIC HEALTH, INC) February 18, 2016; claim 1	1-12, 13/1-12, 14/13/1-12
A	BLITZ et al. Gonadotropin-Releasing Hormone And Luteinizing Hormone Receptor Polymorphisms And Risk Of Hyperemesis Gravidarum. American Journal of Obstetrics & Gynecology, 2009, Vol. 201, No. 6, page S296 column 2 fourth – fifth paragraphs	1-12, 13/1-12, 14/13/1-12
A	US 6,197,329 B1 (HERMELIN et al.) March 6, 2001; claim 53	18/17/15-16, 19/17/15-16, 20-21
PX	FEJZO, M et al. Placenta and Appetite Genes GDF15 and IGFBP7 are Associated with Hyperemesis Gravidarum. Nature Communications. 21 March 2018, Vol. 9, No. 1178, pages 1-9; entire document. doi:10.1038/s41467-018-03258-0	1-12, 13/1-12, 14/13/1-12, 15-16, 17/15-16, 18/17/15-16, 19/17/15-16, 20-21

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

21 May 2018 (21.05.2018)

Date of mailing of the international search report

11 JUN 2018

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774