Journal of Assisted Reproduction and Genetics (2019) 36:1539–1548
https://doi.org/10.1007/s10815-019-01499-6

REVIEW

Genetic diagnosis in first or second trimester pregnancy loss using exome sequencing: a systematic review of human essential genes

Sarah M. Robbins[1,2] · Matthew A. Thimm[3] · David Valle[1] · Angie C. Jelin[4]

Received: 18 December 2018 / Accepted: 29 May 2019 / Published online: 4 July 2019

Abstract

Purpose Non-aneuploid recurrent pregnancy loss (RPL) affects approximately 100,000 pregnancies worldwide annually. Exome sequencing (ES) may help uncover the genetic etiology of RPL and, more generally, pregnancy loss as a whole. Previous studies have attempted to predict the genes that, when disrupted, may cause human embryonic lethality. However, predictions by these early studies rarely point to the same genes. Case reports of pathogenic variants identified in RPL cases offer another clue. We evaluated known genetic etiologies of RPL identified by ES.

Methods We gathered primary research articles from PubMed and Embase involving case reports of RPL reporting variants identified by ES. Two authors independently reviewed all articles for eligibility and extracted data based on predetermined criteria. Preliminary and amended analysis isolated 380 articles; 15 met all inclusion criteria.

Results These 15 articles described 74 families with 279 reported RPLs with 34 candidate pathogenic variants in 19 genes (*NOP14*, *FOXP3*, *APAF1*, *CASP9*, *CHRNA1*, *NLRP5*, *MMP10*, *FGA*, *FLT1*, *EPAS1*, *IDO2*, *STIL*, *DYNC2H1*, *IFT122*, *PADI6*, *CAPS*, *MUSK*, *NLRP2*, *NLRP7*) and 26 variants of unknown significance in 25 genes. These genes cluster in four essential pathways: (1) gene expression, (2) embryonic development, (3) mitosis and cell cycle progression, and (4) inflammation and immunity.

Conclusions For future studies of RPL, we recommend trio-based ES in cases with normal parental karyotypes. In vitro fertilization with preimplantation genetic diagnosis can be pursued if causative variants are found. Utilization of other sequencing technologies in concert with ES should improve understanding of the causes of early embryonic lethality in humans.

Keywords Pregnancy loss · Embryonic lethal · Recurrent pregnancy loss · Exome sequencing

Introduction

According to the World Health Organization, 211 million pregnancies occur each year [1]. Pregnancy loss before 20 weeks, or miscarriage, occurs in an estimated 15–25% of pregnancies [2]. About 50% of losses are secondary to chromosomal abnormalities, including monosomy and trisomy [3]. In many cases, these errors are sporadic and do not predict future pregnancy losses. Recurrent pregnancy loss (RPL) is defined as the loss of two or more consecutive clinically recognized pregnancies before 20 weeks of gestation [4]. RPL is much less common than singular pregnancy loss, occurring in only about 1 in 100 pregnancies [4]. RPL is suggestive of a genetic etiology. Known risk factors in RPL include maternal endocrine or anatomical abnormalities, maternal inherited thrombophilia, and parental genetic abnormalities such as balanced translocations. The American Society for Reproductive Medicine recommends a workup for RPL after the second clinically recognized miscarriage [5]. The workup includes maternal and paternal karyotype, tests for lupus anticoagulant, antiphospholipid antibodies, prolactin level, hemoglobin A1c, and a sonohysterogram to evaluate uterine anatomy; additionally, a chromosomal microarray on the product of conception (POC) is recommended to determine if unexplained sporadic aneuploidy accounts for both miscarriages [5]. Popescu et al. recently

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10815-019-01499-6) contains supplementary material, which is available to authorized users.

✉ Sarah M. Robbins
srobbi10@jhmi.edu

[1] McKusick-Nathans Institute in the Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[2] Predoctoral Training Program in Human Genetics, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[3] Johns Hopkins University School of Medicine, Baltimore, MD, USA

[4] Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

estimated that 95% of RPL can be explained after the recommended workup [5]. Therefore, 5% of RPL, or approximately 100,000 miscarriages per year worldwide, remain completely unexplained by known causes, either maternal or fetal.

Mendelian causes of pregnancy loss have long been hypothesized. Certain genetic disorders, like Barth syndrome (OMIM 302060), lethal multiple pterygium syndrome (OMIM 253290), and homozygous achondroplasia (OMIM 100800), are known to be associated with fetal demise in utero [6–8]. The association of these disorders with pregnancy loss was recognized in natural history studies of the primary disorder. Often, POCs in these cases carry the same variants as their living affected siblings but, because of variable expressivity, do not survive intrauterine life [9]. Maternal causes of pregnancy loss may also include single-gene disorders, such as Factor V Leiden thrombophilia [10]. Prior to the advent of exome sequencing (ES), no completely penetrant Mendelian causes of RPL had been described.

Many challenges hinder the investigation of genetic factors contributing to RPL, including acquisition of adequate samples, public stigma of pregnancy loss, and an expected high degree of genetic heterogeneity. POC samples are often degraded or contaminated with maternal tissues, making it difficult to determine variants that are present only in the POC [11]. Mothers may not report pregnancy losses due to fear that they may be at fault or that they may be blamed for the loss. Additionally, losses may occur in the very early stages of gestation, before the mother is aware of her pregnancy. Lastly, variants in many genes are likely responsible for fetal demise. In mice, roughly 30% of gene knockouts result in utero lethality, strongly suggesting that other mammals will have a similarly large fraction of essential genes [12]. ES offers a new and promising method for the identification of human essential genes. By interrogating the genetic sequence of all protein-coding parts of the genome, ES has revolutionized diagnosis of genetic disorders, allowing researchers to associate thousands of genes with the disorders they cause [13]. Finding causative variants in candidate genes for a Mendelian disorder creates paths for disease research, genetic counseling, and treatment, giving hope to patients and their families. Successful research using ES depends on reliable variant interpretation (ClinVar, Human Gene Mutation Database, dbSNP) and previous knowledge of the function of the encoded proteins, as well as robust frequency information allowing researchers to narrow their variant search. Disorders in which the causative variant and gene are not discovered by ES alone may also be solved by coordination of ES with complementary technologies, like genome sequencing (GS) and RNA sequencing.

This review summarizes current literature describing the use of ES for the diagnosis of causes of pregnancy loss. We highlight 19 genes now known to be causes of human embryonic lethality and make recommendations for the ongoing use of ES in the obstetrics or fertility clinic.

Materials and methods

We performed this review according to the PRISMA guidelines for systematic reviews. We summarized current reports of the use of ES in cases with known recurrent pregnancy loss before 20 weeks gestation. Our final inclusion criteria focused specifically on ES in an effort to report only genes and variants associated with pregnancy loss and RPL. To date, GS studies have reported only copy number variants or aneuploidy and are often undertaken as an alternative to array comparative genomic hybridization, the current standard for aneuploidy detection. Variants in non-coding sequences with potential regulatory consequences can be detected by GS, but our ability to analyze the consequence of these variants is, to date, limited.

Search strategy and study selection

We conducted a systematic literature search for studies written in English from January 2010 to April 2018 on PubMed and Embase databases. We restricted the date range to begin in 2010 based on the date of the first successful application of ES for the discovery of the cause of a Mendelian disorder [14]. ES is still a relatively new technology but has significantly changed the field of Mendelian genetics [14]. Our search strategy identified reports of families or cases including phenotype information and known fetal losses, with documented use of ES.

Two reviewers (S.R. and M.T.) screened all articles independently in April 2018. We included all studies that (1) used ES, (2) included novel patient reports, (3) had human subjects, (4) were in English, and (5) included cases of RPL (< 20 weeks gestation). We excluded reports that did not meet all of these criteria ($n = 251$, Figure S1) (Table 1). The final list of reports was reviewed by all authors of this report. We also reviewed the references of the retrieved publications to identify any additional studies that may not have been identified in our original search. For this review, we included all reports of at least 1 family with 2 or more pregnancy losses before 20 weeks gestation, or at least one pregnancy loss with other medically terminated pregnancies due to fetal anomalies, or two or more families with one or more pregnancy loss (Table 2). Because we were searching for monogenic causes of RPL, we excluded carriers of complex chromosomal rearrangements, insertions, deletions, or translocations. The search returned 309 references, which we reduced to 262 after removal of duplicate articles. After screening of titles and abstracts, we assessed full-text publications for eligibility and selected 10 publications that fulfilled all inclusion/exclusion criteria. Upon review, we performed an amended search, which recovered articles excluded from our original search and identified an additional 118 candidate articles, which we reduced to a final list of 15 publications (Fig. 1; [15–29]).

Table 1 Eligibility criteria for this review. PubMed and Embase searches were completed in April 2018

Inclusion criteria	Exclusion criteria
1. Recurrent (2+) first or second trimester pregnancy loss (< 20 weeks gestation)	1. Exome sequencing methods with no patients reported
2. Exome sequencing utilized	2. Review
3. Novel reports of a patient or family (i.e., a case study or case series)	3. Genome sequencing, SNP array, or karyotype only
4. Human subjects	4. Parents are translocation carriers
5. English language	
6. Full text available	
7. Found in PubMed and/or Embase	

Data extraction

Two authors extracted data using a standardized set of tables created during the filtering process. Following discussion with all authors, we assessed all articles using these tables (Tables 2, 3, S1). The data includes information on ES platform and capture, patient phenotype, and pregnancy information, as well as variant filtering method and final candidate genes. We also noted other functional data performed by the study authors.

Data analysis

We summarized the data in tabular and narrative format and highlighted studies with particularly interesting approaches to variant filtering and any functional data that lends additional support to smaller studies. We also included studies with additional families with related phenotypes in an effort to expand understanding of what phenotypes and pathways may also lead to pregnancy loss or RPL in extreme cases.

Results

Study characteristics

We report key characteristics of the selected studies in Tables 2 and 3. Maternal age in all cases was earlier than age 40 at the time of her first miscarriage; however, some reports only noted age after multiple miscarriages. Because aneuploidy is responsible for such a large percentage of fetal loss in early pregnancy, all parents in these studies were confirmed to be euploid before ES.

The most common ES method used included exome selection with Agilent SureSelect exome capture reagents followed by paired-end sequencing on Illumina HiSeq machines (Table S1). DNA from maternal, paternal, and family member samples were extracted from leukocytes in 11 studies, and proband samples were obtained from fetal or placental tissue in 7 studies. Importantly, studies where fetal tissue was not available to be sequenced directly often included multiple other family members or other families to test their hypotheses.

Pregnancy loss candidate genes

The authors of these studies used multiple databases and algorithms to help prioritize the tens of thousands of variants that are found in any sequenced exome. These included large collections of human genomes and exomes to assess variant frequency: ExAC, gnomAD, BRAVO, 1000 Genomes, and Exome Variant Server [30–34]. We also used algorithms based on nucleotide and amino acid conservation in related vertebrates to evaluate nucleotide variants and prediction scores based on protein function (GERP++, PolyPhen2, SIFT, MutationTaster, and others) to predict the consequences of missense variants [35–38].

In all of the papers we selected, cases were filtered based on a presumed de novo autosomal dominant or autosomal recessive mode of inheritance, which identified candidate variants that were either homozygous (in consanguineous families) or compound heterozygous with one variant being inherited from each parent. Confirmatory functional testing was completed in some but not all cases; all functional testing revealed that the identified variants produced loss of function. In one case, authors discovered a heterozygous variant inherited from the mother in a paternally imprinted gene, thereby leading to monoallelic expression of the impaired allele. Authors also used functional scores, such as probability of loss-of-function intolerance (pLI), and mouse models to predict lethality [12, 31]. We found a strong correlation between pLI > 0.9 and an embryonic lethal mouse phenotype (Figure S2) [39].

A total of 19 strong candidate genes were reported by 14 studies, with 3 studies reporting 2 or more different candidate genes in 2 different families with a similar function, as reported in GeneCards (genecards.org) and OMIM (omim.org) (Table 3). One study included in these 14 and one other study also reported other genes that did not meet these criteria, which we classify as genes of unknown significance (Table S2). These 19 genes all have known functions, many of which cluster in 4 categories: development, gene expression and regulation, immunity and inflammation, and cell cycle progression (Fig. 2). Genes encoding proteins already known to be involved or essential in vertebrate development, such as *APAF1*, *CASP9*, *CSPP1*, *IFT122*, *NLRP5*, and *PADI6*, are likely to have a conserved function in human development

Table 2 Pedigree information for all reports included in this review

Reference	Number of RPL families	Ethnicity	Number of pregnancy losses	Age of mother at first pregnancy loss	Number of weeks at pregnancy loss	Diagnosis given
Suzuki et al. 2018	2	Iranian	9, 2		First and second trimester	RPL^{M}
Reichert et al. 2016	1	Cuban, Spanish	7	37	17 weeks, 22 weeks	$Hydrops^{POC}$
Spellicy et al. 2018	2	Caucasian	2+, 6	25, 27	21, 20, 20, 18 weeks	Neural tube $defects^{POC}$
Shamseldin et al. 2013	1	Saudi	3	(less than 24)	19 weeks	$NIHF^{POC}$
Docherty et al. 2015	5		6, 4, 0 (× 3)			Multi-locus imprinting $disorder^{POC}$
Quintero-Ronderos et al. 2017	49	Caucasian	2 (× 3), 3 (× 34), 5 (× 1), 18 (× 1)		131 SAB >10 weeks, 32 SAB < 10 weeks but > 20 weeks	RPL^{M}
Cristofoli et al. 2017	1	Caucasian	5 + 2 medically terminated		7–11 weeks, 21 weeks	Fetal $microcephaly^{POC}$
Filges et al. 2015*	1	Caucasian	36	34 (at 18th SAB)	5.5–12 weeks	$Triploidy^{POC}$
Qiao et al. 2016**	4	Caucasian	2 (× 3), 5	29, 37, 33, 39	Embryonic stage	RPL^{M}
Tsurusaki et al. 2014	1		5 + 2 medically terminated		6, 8, 8, 7, 21 weeks SAB, 13, 21 weeks MT	Fetal $hydrops^{POC}$, skull $deformity^{POC}$, acromelic limb $shortening^{POC}$
Qian et al. 2018	1	Han Chinese	5 + 1 HM		7, 7, 7, 7, 10	Non-molar RPL^{M}
Pan et al. 2019	1	Chinese	9 (5, 2, 2)	(less than 30, 26, 23)		RPL^{M}
Wilbe et al. 2015	1		6	30s	6, 17, 18, 30, 27	Lethal fetal akinesia deformation $sequence^{POC}$
Shehab et al. 2017	1		19		< 20 in all cases	Intrauterine fetal $demise^{POC}$
Begemann et al. 2018	3	German, UK	3, > 2, 2 + 1 MT		8, 24, 36; not given; 4, 4, 19	Multi-locus imprinting $disorder^{POC}$
Total	74		279	Average 30.3	Average 16.3+	

The column "Number of pregnancy losses" notes the number of pregnancy losses for each family; each number denotes a separate family reported with parenthesis denoting number of families with that number of SAB. The columns "Ethnicity" and "Age of mother at first pregnancy loss" are only reported in some cases, as shown; if not given, any information was included here in parentheses with additional qualifier. The column "Number of weeks gestation at pregnancy loss" has the number of weeks gestation for each pregnancy loss if given; aggregate information is given for larger studies. Diagnosis may be for mother experiencing pregnancy loss (M) or fetuses (POC), as denoted by superscript. +Average was calculated using only losses with exact week number given. *RPL* recurrent pregnancy loss, *NIHF* non-immune hydrops fetalis, *HM* hydatidiform mole, *also reported in Manokhina et al. 2013, **also reported in Rajcan-Separovic 2010

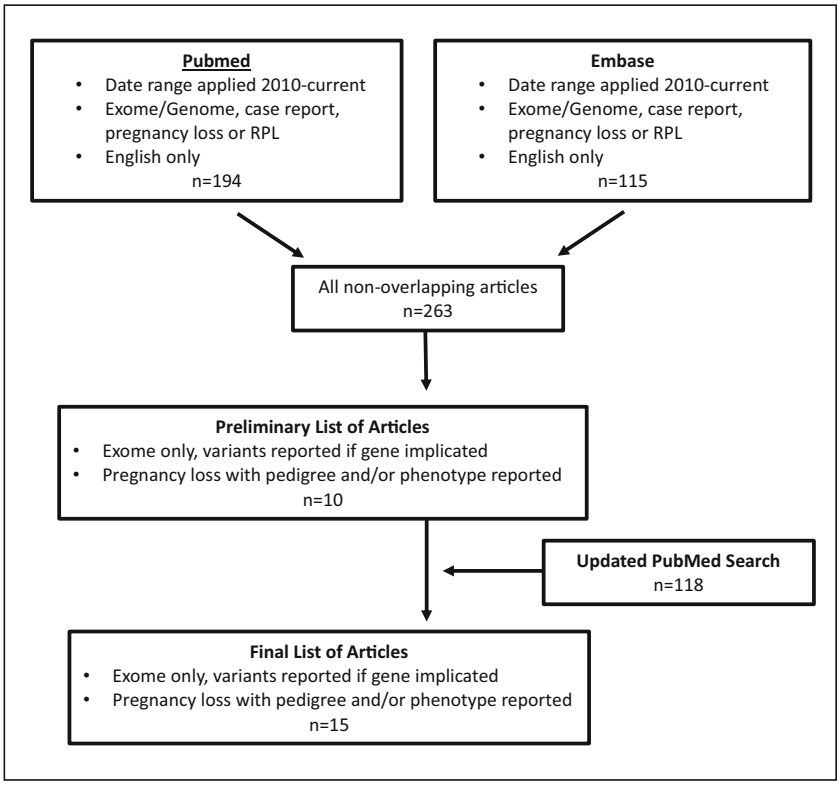

Fig. 1 Flow diagram of article selection for this review. Article selection was completed according to PRISMA guidelines for systematic reviews

[15]. For example, Spellicy et al. showed that two genes involved in neural tube closure in mice, *APAF1* and *CASP9*, encode proteins that function in humans similar to in mouse [15]. In their study of two families, compound heterozygous variants in one of two apoptosis genes, *APAF1* and *CASP9*, lead to multiple pregnancy losses in the second trimester. All of the affected fetuses displayed neural tube defects, including craniorachischisis and vertebral segmentation anomalies [15]. Variants in these genes were rare (MAF < 0.03 in dbSNP, 1000 Genomes, ExAC, and ESP) and involved in a functional domains in the encoded protein. Both genes encode proteins known to be involved in the apoptotic pathway. Functional studies confirmed that the variants were loss of function, with patient fibroblasts showing significantly less apoptotic activity in a TUNEL immunofluorescence assay [15]. Key regulatory genes, such as *FOXP3* and *NOP14*, may also have widespread effects in human development. For example, Reichert et al. demonstrated that heterozygous loss of function of the transcription factor *FOXP3* caused fetal demise and hydrops in 7 fetuses in one family [16]. Additionally, progression through the cell cycle and normal mitotic function is important for a developing embryo, as evidenced by the involvement of *DYNC2H1* and *STIL*. Starting from a single-cell fertilized egg, embryonic growth is dependent upon normal mitosis. Cristofoli et al. describe a case of loss-of-function variants in *STIL*, a core centrosome protein needed for maintenance of the mitotic spindle leading to multiple pregnancy losses with all POCs displaying microcephaly [17]. *STIL* loss-of-function variants were inherited from both parents in affected POCs, while both unaffected siblings tested were heterozygous with one functional allele [17]. Lastly, effective immune response, as controlled by *FOXP3*, *IDO2*, *NLRP2*, *NLRP5*, and *NLRP7*, is needed to combat infection in the fetus and to ensure that the maternal immune system does not reject the fetus. The four pathways identified in these studies are promising candidates for further follow-up in other cases.

Pregnancy loss genes of unknown significance

In two studies, authors reported genes of unknown significance, which did not meet our criteria for candidate pregnancy loss genes. These genes of unknown significance were found in only one family without functional testing and require additional functional testing to determine their inclusion in RPL genes. Filges et al. presented a list of 8 candidates from a family with 36 pregnancy losses, where all 5 POCs tested were fully triploid [18]. However, without additional families or performing additional studies, they were unable to confirm which of these was the causative gene (Table S2). Lastly, Quintero-Ronderos et al. reported a cohort of 49 women experiencing RPL and found candidate genes in 22 cases (Table S2) [19]. Filtering criteria isolated candidates that have functions related to spontaneous abortion, implantation, uterine biology, or mouse lethality in multiple databases including PubMed, GeneCards (http://www.genecards.org), and Mouse Genome Informatics database [40]. The final candidate list in

Table 3 Strong pregnancy loss candidate genes found in 2 or more families or candidate genes where functional studies were performed

Reference	Gene	No. tested	pLI	Alleles	Variant	Function	Mouse model
Suzuki et al. 2018	*NOP14*	2	0.02	Homozygous	c.136C>G	18S rRNA processing and 40S ribosome biogenesis	None available
Reichert et al. 2016	*FOXP3*	2	0.95	Het (paternal imprinting)	c.1009C>T (p. R337X)	Transcriptional regulation and autoimmunity, embryonic development	Abnormal immune function, inflammation, abnormal reproductive development (MGI: 1891436)
Spellicy et al. 2018	*APAF1*	2	0.0	Cmpd. het	p.L261F;p.C450W *in trans*	Neural tube closure	Neural tube defects, abnormal neural development excencephaly, perinatal lethality (MGI: 1306796)
	CASP9	2	0.0	Cmpd. het	p.H237P;p.G309Wfs*3 *in trans*	Neural tube closure	Neural tube defects, abnormal neural development, excencephaly, perinatal lethality (MGI: 1277950)
Shamseldin et al. 2013	*CHRNA1*	1	0.0	Homozygous	p.R254C	Acetylcholine receptor subunit alpha	Neuromuscular defects, kyphosis, neonatal and juvenile lethality (MGI: 87885)
Docherty et al. 2015	*NLRP5*	4	0.0	Cmpd. het	p.G555V; p.C774R *in trans* p.Q785X, p.L947P *in trans*	Zygote development and maternal imprinting	Female infertility (MGI: 1345193)
Quintero-Ronderos et al. 2017	*MMP10*	1	0.0	Het	p.D199N	Extracellular matrix organization	Lung inflammation and hemorrhage (MGI: 97007)
	FGA	1	0.0	Het	p.F685C	Coagulation	Abnormal coagulation, hemorrhage in multiple tissues, female infertility (MGI: 1316726)
	FLT1	2	1.0	Het	p.S318L (1), p.R812Q (1)	Angiogenesis	Abnormal vascular and heart development, embryonic lethality (MGI: 95558)
	EPAS1	3	0.5	Het	p.L605V (1), p.Y488C (2)	Angiogenesis	Abnormal angiogenesis/blood cell counts, embryonic lethality, respiratory distress (MGI: 109169)
	IDO2	2	0.0	Het	p.F180C (1), p.F250S (1)	Immunity	Decreased immune cell counts (MGI: 2142489)
Cristofoli et al. 2017	*STIL*	2	0.12	Cmpd. het	p.H411D, p.M124V *in trans*	Centrosome integrity and mitotic spindle formation	Holoprosencephaly, neural tube defects, complete embryonic lethality (MGI: 107477)
Qiao et al. 2016	*DYNC2H1*	2	0.0	Cmpd. het	p.Y2016C, p.D2184V *in trans*	Dynein involved in ciliary transport	Holoprosencephaly, congenital heart disease, complete embryonic lethality (MGI: 107736)
Tsurusaki et al. 2014	*IFT122*	3	0.0	Cmpd. het	p. E370fs*51, p.G546R *in trans*	Bone and embryonic development	Exencephaly, neural tube defects, hemorrhage, complete embryonic lethality (MGI: 1932386)
Qian et al. 2018	*PADI6*	1		Cmpd. het	p.N598S, p.R682Q *in trans*	Post-translational modifier of arginine, involved in early embryo development	Abnormal embryonic development, female infertility, abnormal oocyte morphology (MGI: 2655198)
Pan et al. 2019	*CAPS*	3	0.0	Homozygous	p.L127Wfs	Calcium binding, ion transport	Abnormal neural development, neonatal lethality, respiratory failure (MGI: 1350922)
Wilbe et al. 2015	*MUSK*	5	0.0	Homozygous	c.40A>AA	Muscle-specific tyrosine kinase receptor	Neuromuscular defects, neonatal lethality, respiratory failure (MGI: 103581)
Shehab et al. 2017	*FOXP3*	5	0.99	Hemizygous	p.D303fs*87	Transcriptional regulation and autoimmunity, embryonic development	Abnormal immune function, inflammation, abnormal reproductive development (MGI: 1891436)
Begemann et al. 2018	*NLRP2*	2	0.0	Homozygous Cmpd. het	p.R493Sfs*32 *in trans*	Immune response regulator, inflammation	None available
	NLRP7	2	0.0	Cmpd. het	p.P105L *in trans* with deletion p.R721W,p.I858T *in trans*	Immune response regulator, inflammation	None available

The number of fetal losses or affected mothers for each family tested for the variant is noted in column "No. tested." Variants given as reported by study authors. Homozygous variants are present in two copies in trans; compound heterozygous variants are present in one copy in trans with another variant in the same gene. Heterozygous variants are present in 1 copy in each family, with number of families with that variant in parentheses. Function is as reported by study authors and may not be comprehensive; additional studies are needed to elucidate all gene functions. Mouse model is an abbreviated description of phenotypes as given by Mouse Genome Informatics (MGI), with MGI ID given in parentheses [30]. *Cmpd.* compound, *het* heterozygous, *pLI* probability of loss of function intolerance [14]

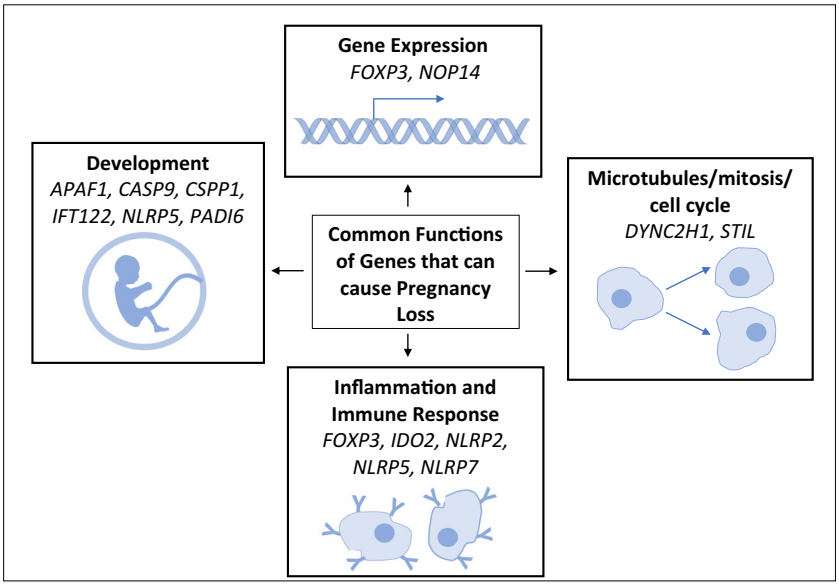

Fig. 2 Functions of genes essential to human development identified in these studies. Essential cellular processes, like mitotic division and gene expression, were common functions of 6 genes, while 5 others are known to be involved in fetal development in mice or humans

this report was not confirmed by further functional studies for 17 of the genes; however, variants were found in more than one family for 3 of their reported genes (Table 3) [19].

Because these studies were limited to a single family and lacked functional studies to confirm causality, the involvement of these genes is still uncertain (Table S2). Additional functional studies or families would likely solve this problem. These genes should be prioritized in future cases where pregnancy loss ES is pursued.

Discussion

We systematically reviewed the literature to determine if whole exome sequencing is an effective strategy for finding Mendelian causes of spontaneous abortion. We present 15 studies that utilized ES to determine the cause of pregnancy loss. These studies found 19 strong candidate genes, along with 25 other preliminary candidate genes, associated with embryonic lethality in humans. These genes are involved in essential pathways such as transcription, mitotic and meiotic cell division, immune response, and embryonic development. For further clinical study of pregnancy loss and RPL, we recommend trio sequencing of POC and both parents, along with frequency and algorithm-based variant prioritization. In cases where more study is needed, we recommend functional tests, submission to GeneMatcher, and additional sequencing studies [41]. Because de novo mutations may contribute to the basis of pregnancy loss, it is extremely important in future studies to include both parents in addition to POC in sequencing studies. Pregnancy loss and RPL sequencing studies can help uncover the core set of genes for human life.

How do genetic defects cause RPL?

Known Mendelian disorders affect every organ system, and we hypothesize that phenotypes causing early lethality will be similarly distributed. In some cases, the genetic defects were clearly linked to a lethal phenotype, like neural tube defects or fetal akinesia deformation sequence [15, 20]. In others, the fetal loss was unexplained but could be due to dysfunction of any organ or developmental system. Here, we focus specifically on RPL cases less than 20 weeks gestation, with reported cases occurring as early as 4 weeks. We expect that genes involved in essential cellular processes, such as mitosis and transcription, are more likely to cause lethality early in pregnancy, as opposed to defects affecting later processes that required later in fetal development. The molecular mechanisms downstream of each variant reported here are currently unclear. Table 2 annotates the general function of each gene as reported in the current literature. These RPL-associated genes point toward pathways like development, gene expression and regulation, immunity and inflammation, and cell cycle progression. Interestingly, variants in three genes in the NLR family of proteins all caused fetal lethality; these genes are known players in the pro-inflammatory cascade [42]. Inflammation is known to mediate the implantation of the fetus and its development in a semi-allogenic environment [30]. The remaining 17 genes on our list are currently the only gene in their family associated with RPL, but more RPL-associated genes, especially those that play a part in the 4 pathways we highlighted, may come to light as additional families with RPL pursue ES or GS. Gene knockout mouse models also offer a window into the function of orthologous RPL genes in humans (Table 3). Of the 19 RPL candidate genes, 16 have a published mouse model, 11 of which have a severe phenotype with either embryonic or perinatal lethality (*Apaf1*, *Casp9*, *Chrna1*, *Epas1*, *Stil*, *Dync2h1*, *Ift22*, *Caps*, and *Musk*) or infertility (*Nlrp5* and

Fga) [40]. Affected organ systems in the fetus included the respiratory, the vascular, and the central nervous system. Fetal brain anomalies associated with RPL, such as exencephaly and holoprosencephaly, were reported in 5 models (*Apaf1*, *Casp9*, *Stil*, *Dync2h1*, and *Ift122*). Knockouts of two genes, *Apaf1* and *Casp9*, resulted in neural tube defects and exencephaly, similar to the reported neural tube defects in the corresponding human RPL examples [15]. In each instance, these anatomic phenotypes may explain the embryonic or fetal demise, or there may be some underlying chemical imbalance not apparent to gross and/or microscopic evaluation. Also, for 4 genes (*Foxp3*, *Nlrp5*, *Fga*, and *Padi6*), the abnormality was in oocyte maturation or in the reproductive tract of affected females. These observations emphasize the need to consider fetal or maternal causes or a combination of both in the causation of RPL. Three of the 19 RPL candidate genes (*Nlrp2*, *Nlrp7*, and *Nop14*) have no reported mouse model and would benefit from further investigation.

Exome sequencing in the clinic

The American Society for Reproductive Medicine estimates that a clinical workup to evaluate parental karyotype, uterine anomalies, and other maternal health issues will explain approximately 50% of RPL cases [5]. In conjunction with POC karyotype, up to 95% of cases can be explained [5]. However, for the approximately 100,000 de novo cases annually in which pregnancy loss cannot be explained using these tests, we recommend trio ES of POC and both parents. Another case where trio ES should be pursued is consistent, abnormal POC karyotypes not found in either parent, as in complete triploidy in all tested POCs [18]. In a reported case, the abnormal karyotype resulted from loss of a crucial gene for meiotic division. ES of trios enhances the diagnostic rate of ES by 10% or more over proband-only ES [43].

Because many genes may cause embryonic lethality and because many biological functions are required for successful development of the fetus, variant filtering may be difficult [12]. Collection of parental DNA for trio sequencing is critical to assess the presence of rare variants identified in the POC. DNA from other family members also experiencing RPL or their POCs should also be collected if possible. Information from additional family member can often help researchers to identify alleles that are now found in trans with another variant allele or are de novo in the POC. De novo variants contribute to a surprisingly large proportion of Mendelian disease, since highly deleterious alleles are often not passed on to offspring [44]. Before functional studies, variant prediction algorithms, such as GERP++, SIFT, PolyPhen2, PROVEAN, pLI, CADD, VEST, REVEL, and Primate-DL, can be utilized for prioritization [31, 35–38, 45–49]. Lethal or severe phenotypes in model systems are also extremely valuable predictors, as essential function may be conserved in humans. Genes associated with severe Mendelian phenotypes in humans may also be prioritized, especially if relatives are diagnosed with those conditions. Because there could be many reasons why a pregnancy might end in fetal loss, we encourage the use of GeneMatcher and the Matchmaker Exchange to find additional families with variants in candidate genes matching the above criteria [41, 50]. Matching can be considered even if patients do not have the exact phenotype, since phenotypic heterogeneity may lead to the severe consequence of embryonic lethality.

In all cases, genetic counseling is warranted to help couples with RPL understand challenges and uncertainties associated with ES, especially for a highly heterogeneous condition like pregnancy loss. Several studies have shown diagnostic rates of 25–35% for ES in pediatric patients with genetic disorders [13]. In our review of the literature, only 19 genes have been associated with pregnancy loss or RPL in the literature according to this systematic review. We expect pregnancy loss may have a lower diagnostic yield than other Mendelian conditions where all genes associated with the disorder are known. However, Vora et al. recently found that prenatal exome sequencing in anomalous fetuses yielded possible diagnoses in 7 of 15 cases, an even higher rate than normally discussed for pediatric ES [51]. Counseling should emphasize that using ES or GS for diagnosis of pregnancy loss is currently experimental, with unknown diagnostic yield [51]. As with all ES or GS studies, there is a risk of secondary findings, including non-paternity and variants of uncertain significance [52, 53]. These experimental ES studies are currently limited in their clinical utility, but the number of clinically actionable findings will continue to increase as ES and GS are used in both research and clinical settings. However, if a variant in a causative gene is discovered, in vitro fertilization (IVF) with preimplantation genetic diagnosis (PGD) can prevent further pregnancy loss by ensuring only healthy embryos are implanted. PGD screens embryos for pathogenic or likely pathogenic variants in clinical actionable genes that are present in the parents. PGD can also be used in conjunction with aneuploidy testing to select non-aneuploid embroys, eliminating another miscarriage risk factor [54]. As in other Mendelian conditions, diagnostic rates will only improve in the future as more genes are associated with pregnancy loss.

Barriers to solving pregnancy loss with ES

Even with ES, not every cause of pregnancy loss will be identified. Because exome capture biases read coverage, copy number variants are extremely hard to detect using ES. Exome capture reagents are not efficient in difficult areas of the genome, so not even the entire protein-coding portion of the genome (a mere 1–2%) is covered with ES. Non-coding variants will not be sequenced with ES. Importantly, ES analysis often focuses on Mendelian phenotypes, making it difficult to find genetic interactions. Environmental variables, like infection during pregnancy or trauma, or gene by environment interactions will also be missed in most cases. Perhaps, the most difficult barrier in

pregnancy loss ES is sample acquisition. Miscarriage may occur even before the mother is aware of her pregnancy. Public stigma keeps many mothers from reporting their miscarriages [55]. Similar to many oncology providers, obstetrics and gynecology providers may not be well-versed in genetic technologies and the interpretation of ES [56]. The small number of studies reported here is only a small fraction of the total number of genes estimated to cause human lethality [39]. More samples and larger cohort studies are needed to create a more comprehensive map of the human essential gene network. This effort will yield insight into the biological systems that are central to human reproduction and the beginnings of life.

Extending the exome: finding human essential genes

ES diagnostic rates can be increased by adding additional data, including other sequencing technologies, like RNA-seq or whole genome sequencing. RNA-seq technologies can add variants not detected by ES due to uneven capture or add important interpretation, such as expression, splicing, or tissue-specific changes. Ideally, tissue from the POC would be collected to perform RNA-seq, but these tissues may be difficult to acquire. GS can extend the findings of ES by interrogating non-coding regions, where interpretation is difficult and by capturing coding regions where capture is difficult, such as CG-rich regions or repeat regions. Long-read GS may further resolve difficult regions for short-read sequencing, which is the most common method in ES today.

Studies in human cell lines, mouse models, and computational algorithms have predicted that over 5000 genes may be essential in humans [38]. Due to ethical and legal restrictions on human embryo studies, natural human knockouts are the only way to confirm human gene essentiality and resolve disparities in these predictive methods. Human essential genes represent the core set of biological information at the cellular, tissue, and organismal level. Collectively, interrogation of pregnancy loss cases is a promising and informative method for discovering the keys to human development.

Funding information This review was funded by the Baylor-Hopkins Center for Mendelian Genomics (NHGRI/NHLBI UM1HG006542) and the Johns Hopkins Women's Health Scholarship.

References

1. World Health Organization. Not every pregnancy is welcome. Accessed at: http://www.who.int/whr/2005/chapter3/en/index3.html. Accessed 8 Aug 2018.
2. Practice Committee Opinion of the ASRM. Evaluation and treatment of recurrent pregnancy loss: a committee opinion. Fertil Steril. 2012;98(5):1103–11.
3. Page JM, Silver RM. Genetic causes of recurrent pregnancy loss. Clin Obstet Gynecol. 2016;59(3):498–508.
4. Shahine L, Lathi R. Recurrent pregnancy loss: evaluation and treatment. Obstet Gynecol Clin N Am. 2015;42(1):117–34.
5. Popescu F, Jaslow CR, Kutteh WH. Recurrent pregnancy loss evaluation combined with 24-chromosome microarray of miscarriage tissue provides a probable or definite cause of pregnancy loss in over 90% of patients. Hum Reprod. 2018;33(4):579–87.
6. Steward CG, Newbury-Ecob RA, Hastings R, Smithson SF, Tsai-Goodman B, Quarrell OW, et al. Barth syndrome: an X-linked cause of fetal cardiomyopathy and stillbirth. Prenat Diagn. 2010;30(1):970–6.
7. Kareminijad A, Ghaderi-Sohi S, Hossein-Nejad NH, et al. Lethal multiple pterygium, the extreme end of the RYR1 spectrum. BMC Musculoskelet Disord. 2016;17:109.
8. McKusick VA, Kelly TE, Dorst JP. Observations suggesting allelism of the achondroplasia and hypochondroplasia genes. J Med Genet. 1973;10(1):11–6.
9. Hammarsjö A, Wang Z, Vaz R, Taylan F, Sedghi M, Girisha KM, et al. Novel KIAA0753 mutations extend the phenotype of skeletal ciliopathies. Sci Rep. 2017;7(1):15585.
10. Rai R, Backos M, Elgaddal S, Shlebak A, Regan L. Factor V Leiden and recurrent miscarriage – prospective outcome of untreated pregnancies. Hum Reprod. 2002;17(2):442–5.
11. Lathi RB, Gustin SL, Keller J, Maisenbacher MK, Sigurjonsson S, Tao R, et al. Reliability of 46,XX results on miscarriage specimens: a review of 1,222 first-trimester miscarriage specimens. Fertil Steril. 2014 Jan;101(1):178–82.
12. Dickinson ME, Flenniken AM, Ji X, et al. High-throughput discovery of novel developmental phenotypes. Nature. 2016;537(7621):508–14.
13. Chong JX, Buckingham KJ, Jhangiani SN, Boehm C, Sobreira N, Smith JD, et al. The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. Am J Hum Genet. 2015;97(2):199–215.
14. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, et al. Exome sequencing identifies the cause of a mendelian disorder. Nat Genet. 2010;42:30–5.
15. Spellicy CJ, Norris J, Bend R, Bupp C, Mester P, Reynolds T, et al. Key apoptotic genes APAF1 and CASP9 implicated in recurrent folate-resistant neural tube defects. Eur J Hum Genet. 2018;26(3):420–7.
16. Reichert SL, McKay EM, Moldenhauer JS. Identification of a novel nonsense mutation in the FOXP3 gene in a fetus with hydrops–expanding the phenotype of IPEX syndrome. Am J Med Genet. 2016;170A(1):226–32.
17. Cristofoli F, De Keersmaecker B, De Catte L, et al. Novel STIL compound heterozygous mutations cause severe fetal microcephaly and centriolar lengthening. Mol Syndromol. 2017;8(6):282–93.
18. Filges I, Manokhina I, Peñaherrera MS, McFadden DE, Louie K, Nosova E, et al. Recurrent triploidy due to a failure to complete maternal meiosis II: whole-exome sequencing reveals candidate variants. Mol Hum Reprod. 2015;21(4):339–46.
19. Quintero-Ronderos P, Mercier E, Fukuda M, González R, Suárez CF, Patarroyo MA, et al. Novel genes and mutations in patients affected by recurrent pregnancy loss. PLoS One. 2017;12(1):e0186149.
20. Pan H, Xiang H, Wang J, Wei Z, Zhou Y, Liu B, et al. CAPS mutations are potentially associated with unexplained recurrent pregnancy loss. Am J Pathol. 2019;189(1):124–31.
21. Suzuki T, Behnam M, Ronasian F, Salehi M, Shiina M, Koshimizu E, et al. A homozygous NOP14 variant is likely to cause recurrent pregnancy loss. J Hum Genet. 2018;63(4):425–30.
22. Shamseldin HE, Swaid A, Alkuraya FS. Lifting the lid on unborn lethal Mendelian phenotypes through exome sequencing. Genet Med. 2013;15(4):307–9.

23. Shaheen R, Shamseldin HE, Loucks CM, Seidahmed MZ, Ansari S, Ibrahim Khalil M, et al. Mutations in CSPP1, encoding a core centrosomal protein, cause a range of ciliopathy phenotypes in humans. Am J Hum Genet. 2014;94(1):73–9.
24. Qiao Y, Wen J, Tang F, Martell S, Shomer N, Leung PCK, et al. Whole exome sequencing in recurrent early pregnancy loss. Mol Hum Reprod. 2016;22(5):364–72.
25. Tsurusaki Y, Yonezawa R, Furuya M, Nishimura G, Pooh RK, Nakashima M, et al. Whole exome sequencing revealed biallelic IFT122 mutations in a family with CED1 and recurrent pregnancy loss. Clin Genet. 2014;85:592–4.
26. Qian J, Nguyen NMP, Rezael M, et al. Biallelic PADI6 variants linking infertility, miscarriages, and hydatidiform moles. Eur J Hum Genet. 2018;26:1007–13.
27. Wilbe M, Ekvall S, Eurenius K, Ericson K, Casar-Borota O, Klar J, et al. MuSK: a new target for lethal fetal akinesia deformation sequence (FADS). J Med Genet. 2015;52:195–202.
28. Shehab O, Tester DJ, Ackerman NC, Cowchock FS, Ackerman MJ. Whole genome sequencing identifies etiology of recurrent male intrauterine fetal death. Prenat Diagn. 2017;37:1040–5.
29. Begemann M, Rezwan FI, Beygo J, Docherty LE, Kolarova J, Schroeder C, et al. Maternal variants in NLRP and other maternal effect proteins are associated with multilocus imprinting disturbance in offspring. J Med Genet. 2018;55:497–504.
30. Chavan AR, Griffith OW, Wagner GP. The inflammation paradox in the evolution of mammalian pregnancy: turning foe into friend. Curr Opin Genet Dev. 2017;47:24–32.
31. Exome Aggregation Consortium, Lek M, Karczewski KJ, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285–91.
32. BRAVO Top-Med Database. Accessed at: https://bravo.sph.umich.edu/freeze5/hg38/about. Accessed 25 Oct 2018.
33. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68–74.
34. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle WA. http://evs.gs.washington.edu/EVS/. Accessed 25 Oct 2018.
35. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol. 2010;6(1):e1001025.
36. Adzuhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248–9.
37. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073–81.
38. Schwarz JM, Rodelsperger C, Scheulke M, et al. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods. 2010;7:575–6.
39. Bartha I, di Iulio J, Venter JC, Telenti A. Human gene essentiality. Nat Rev Genet. 2018;19(1):51–62.
40. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36(1):928–30.
41. Tschopp J, Martinon F, Burns K. NALPs: a novel protein family involved in inflammation. Nat Rev Mol Cell Biol. 2003;4:94–105.
42. Bult CJ, Blake JA, Smith CL, Kadin JA, Richardson JE, the Mouse Genome Database Group. Mouse Genome Database (MGD) 2019. Nucleic Acids Res. 2019;47(D1):D801–6.
43. Farwell KD, Shahmirzadi L, El-Khechen D, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet Med. 2017;17:578–86.
44. Veltman J, Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet. 2012;13(8):565–75.
45. Choi Y, Sims GE, Murphy S, Miller JR, Chain AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One. 2012;7(1):e46688.
46. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310–5.
47. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics. 2013;14(Suppl 3):S3.
48. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. Am J Hum Genet. 2016;99(4):877–85.
49. Sundaram L, Gao H, Padigepati SR, McRae JF, Li Y, Kosmicki JA, et al. Predicting the clinical impact of human mutation with deep neural networks. Nat Genet. 2018;50(8):1161–70.
50. Philippakis AA, Azzariti DR, Beltran S, Brookes AJ, Brownstein CA, Brudno M, et al. The Matchmaker Exchange: a platform for rare disease gene discovery. Hum Mutat. 2015;36(10):915–21.
51. Vora NL, Powell B, Brandt A, Strande N, Hardisty E, Gilmore K, et al. Prenatal exome sequencing in anomalous fetuses: new opportunities and challenges. Genet Med. 2017;19(11):1207–16.
52. Jelin AC, Vora N. Whole exome sequencing: applications in prenatal genetics. Obstet Gynecol Clin N Am. 2018;45(1):69–81.
53. Quinlan-Jones E, Kilby MD, Green S, et al. Prenatal whole exome sequencing: the views of clinicians, scientists, genetic counsellors and patient representatives. Prenat Diagn. 2016;36(1):10–9.
54. Capalbo A, Romanelli V, Cimadomo D, Girardi L, Stoppa M, Dovere L, et al. Implementing PGD/PGD-A in IVF clinics: considerations for the best laboratory approach and management. J Assist Reprod Genet. 2016;33(1):1279–86.
55. Hanschmidt F, Linde K, Hilbert A, Riedel- Heller SG, Kersting A. Abortion stigma: a systematic review. Perspect Sex Reprod Health. 2016;48(4):169–77.
56. Gray SW, Park ER, Najita J, Martins Y, Traeger L, Bair E, et al. Oncologists' and cancer patients' views on whole- exome sequencing and incidental findings: results from the CanSeq study. Genet Med. 2016;18:1011–9.