

## **Endometrial transcriptome in recurrent miscarriage and recurrent implantation failure**

HUANG, Jin, QIN, Hao, YANG, Yihua, CHEN, Xiaoyan, ZHANG, Jiamiao, LAIRD, Susan <<http://orcid.org/0000-0003-4020-9020>>, WANG, Chi Chiu, CHAN, Ting Fung and LI, Tin Chiu

Available from Sheffield Hallam University Research Archive (SHURA) at:

<http://shura.shu.ac.uk/15428/>

---

This document is the author deposited version. You are advised to consult the publisher's version if you wish to cite from it.

### **Published version**

HUANG, Jin, QIN, Hao, YANG, Yihua, CHEN, Xiaoyan, ZHANG, Jiamiao, LAIRD, Susan, WANG, Chi Chiu, CHAN, Ting Fung and LI, Tin Chiu (2017). Endometrial transcriptome in recurrent miscarriage and recurrent implantation failure. *Reproduction*, 153 (6).

---

### **Copyright and re-use policy**

See <http://shura.shu.ac.uk/information.html>

1 **TITLE**

2 A comparison of transcriptomic profiles in endometrium during window of implantation between women  
3 with unexplained recurrent implantation failure and recurrent miscarriage

4

5 **RUNNING TITLE**

6 Endometrial transcriptome in recurrent miscarriage and recurrent implantation failure

7

8 **AUTHORS**

9 Jin Huang<sup>1\*</sup>, Hao Qin<sup>2\*</sup>, Yihua Yang<sup>1,3</sup>, Xiaoyan Chen<sup>1</sup>, Jiamiao Zhang<sup>1,3</sup>, Susan Laird<sup>4</sup>, Chi Chiu Wang<sup>1,5,</sup>  
10 <sup>6</sup>, Ting Fung Chan<sup>2</sup>, Tin Chiu Li<sup>1</sup>.

11 \*These authors contributed equally to this work.

12

13 **ADDRESS**

14 <sup>1</sup> Department of Obstetrics and Gynaecology, the Chinese University of Hong Kong, Hong Kong SAR,  
15 China

16 <sup>2</sup> School of Life Sciences, The Chinese University of Hong Kong, Hong Kong SAR, China

17 <sup>3</sup> Reproductive Medicine Centre of the Affiliated Hospital, Guilin Medical University, Guilin, Guangxi,  
18 China.

19 <sup>4</sup> Biomolecular Sciences Research Centre, Sheffield Hallam University, Sheffield, UK.

20 <sup>5</sup> Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong, Hong Kong SAR,  
21 China

22 <sup>6</sup> School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong SAR, China

23

24 **DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST:**

25 Direct Grant for Research (Ref No. 2014.1.042) from the Chinese University of Hong Kong. HQ is  
26 supported by a General Research Fund (GRF14102014) from the Research Grant Committee, Hong Kong  
27 Special Administration region to TFC. The authors have no other relevant conflicts of interest to disclose.

28

29 **CORRESPONDING AUTHOR:**

30 Tin Chiu Li, MD, PhD.  
31 Professor  
32 Department of Obstetrics and Gynaecology  
33 The Chinese University of Hong Kong  
34 Prince of Wales Hospital  
35 Shatin, New Territories  
36 Hong Kong  
37 Tel: (852) 2632 3199  
38 Fax: (852) 2636 0008  
39 Email: [tinchiu.li@cuhk.edu.hk](mailto:tinchiu.li@cuhk.edu.hk)  
40  
41

42 **ABSTRACT**

43 The endometrium becomes receptive to the embryo only in the mid-luteal phase, but not other stages of  
44 the menstrual cycle. Endometrial factors play an important role in implantation. Women with recurrent  
45 miscarriage and recurrent implantation failure have both been reported to have altered expression of  
46 receptivity markers during the window of implantation. We aimed to compare the gene expression  
47 profiles of the endometrium in the window of implantation among women with unexplained recurrent  
48 implantation failures (RIF) and unexplained recurrent miscarriages (RM) by RNA sequencing (RNA-Seq).  
49 In total 20 patients (9 RIF and 11 RM) were recruited. In addition 4 fertile subjects were included as  
50 reference. Endometrium samples were precisely timed on the 7th day after luteal hormone surge (LH+7).  
51 All the 24 endometrium samples were extracted for total RNA. The transcriptome was determined by  
52 RNA-Seq in first 14 RNA samples (5 RIF, 6 RM, and 3 fertile). Differentially expressed genes between  
53 RM and RIF were validated by quantitative real time PCR (qPCR) in all 24 RNA samples (9 RIF, 11 RM  
54 and 4 fertile). Transcriptomic profiles of RM and RIF, but not control samples, were separated from each  
55 other by principle component analysis (PCA) and support vector machine (SVM). Complementary and  
56 coagulation cascades pathway was the significantly up-regulated in RIF while down-regulated in RM.  
57 Differentially expressed genes C3, C4, C4BP, DAF, DF and SERPING1 in complement and coagulation  
58 cascade pathway between RM and RIF were further validated by qPCR. This study compared endometrial  
59 transcriptome among patients with RIF and RM in the window of implantation; it identified differential  
60 molecular pathways in endometrium between RIF and RM, which potentially affect the implantation  
61 process.

62

63 **KEY WORDS:**

64 Transcriptome; Endometrium; Window of Implantation; Recurrent Miscarriage; Recurrent Implantation  
65 failure

## 66 **Introduction**

67       Implantation is a process when the embryo attaches to the endometrium, followed by migration and  
68 invasion into the deeper layer of the endometrium to become embedded, which involves a complex  
69 sequence of cellular and molecular changes. There are two well defined categories of reproductive failure  
70 attributable to implantation disorders, namely recurrent miscarriage (RM) and recurrent implantation  
71 failure (RIF). Recurrent implantation failure refers to failure to achieve a clinical pregnancy after transfer  
72 of at least four good-quality embryos in a minimum of three cycles in a woman under the age of 40 years  
73 (Coughlan *et al.* 2014). The failure usually occurs at an earlier stage, resulting in complete failure to  
74 implant or failure to establish the pregnancy. On the other hand, recurrent miscarriage is defined as three  
75 or more consecutive clinical miscarriages (Saravolos & Li 2012). It commonly manifests as pregnancy  
76 loss later on in the pregnancy, often around 6-8 weeks gestation.

77       A number of earlier studies suggested that the endometrium in unexplained RM and RIF shared some  
78 common pathological changes. For example, both uNK cell count and interleukin 15 expressions have  
79 been reported to be increased in the two conditions (Tuckerman *et al.* 2007, Tuckerman *et al.* 2010) On  
80 the other hand, certain molecules have been found to be deranged in one condition but not the other. A  
81 notable example is beta3 integrin, which is down regulated in RM (Germeyer *et al.* 2014) but not in RIF  
82 (Coughlan *et al.* 2013); whereas leukaemia inhibitory factor is down regulated in RIF (Mariee *et al.* 2012),  
83 but not in RM (Xu *et al.* 2012, Karaer *et al.* 2014). In most of these earlier studies, only one specific  
84 marker was examined, with the exception of a few which examined up to 3 specific markers at the same  
85 time (Xu *et al.* 2012). The study of a single or a few biomarkers has a limited value especially in the case  
86 of implantation as it is a rather complex process involving several well recognized steps (apposition,  
87 adhesion, and invasion including angiogenesis) (Fitzgerald *et al.* 2008), each of which involves many  
88 molecules.

89       An alternative approach which enables the simultaneous study of all different molecules involved in  
90 the implantation process is transcriptomic study by using micro-array analysis or RNA sequencing (RNA-  
91 Seq). Several studies have used this approach to examine the endometrium in the peri-implantation period  
92 but they focused either on the changing transcriptome profiles before and during the window of  
93 implantation (Diaz-Gimeno *et al.* 2011, Hu *et al.* 2014), or in a specific population such as PCOS (Qiao *et al.*  
94 2008), RIF (Koler *et al.* 2009, Altmae *et al.* 2010, Ruiz-Alonso *et al.* 2013, Koot *et al.* 2016) or RM  
95 (Othman *et al.* 2012, Kosova *et al.* 2015), or under the impact of different hormonal treatment (Mirkin *et al.*  
96 2004, Haouzi *et al.* 2009). However, none of the earlier studies employed the RNA sequencing to  
97 compare and contrast the transcriptome profiles of endometrium in unexplained RM and RIF.

98       More recently, Brosen et al (Teklenburg *et al.* 2010, Brosens *et al.* 2014, Macklon & Brosens 2014)  
99 hypothesized that RM is associated with an over-receptive endometrium which would allow defective or  
100 abnormal embryos to implant and in turn leads to super-fertility, but followed by an increased risk of  
101 miscarriage of an abnormal embryo. In contrast, in women with RIF, implantation often fails to take place  
102 despite the replacement of many good quality embryos, implying that the defect is in the endometrium  
103 which is less receptive. The underlying molecular mechanism of altered endometrium receptivity during  
104 window of implantation in RM and RIF are still unclear.

105 In this study, we wish to directly compare the transcriptome profiles of RM and RIF, on precisely  
106 timed endometrial specimens obtained seven days after the LH surge (LH+7) with a view to establishing  
107 to what extent RM and RIF represent two ends of the spectrum of implantation disorder.

108

## 109 **Materials and Methods**

### 110 **Subjects**

111 Subjects were recruited from the Prince of Wales Hospital, Chinese University of Hong Kong. The  
112 inclusion criteria of all subjects recruited include: age no more than 40 years, with regular cycles (25-35  
113 day), had not used steroid hormone in the preceding 2 months. Women with one or more of the following  
114 situations were excluded: peripheral blood showing chromosomal anomaly, tested positive for  
115 anticardiolipin antibody or lupus anticoagulant, abnormal thyroid function test, uncorrected uterine  
116 anomalies, intra-uterine device in situ, intrauterine adhesions or serious systematic disease. Women with  
117 unexplained RIF was defined as failure to achieve a clinical pregnancy after transfer of at least four good-  
118 quality embryos in a minimum of three cycles in a woman under the age of 40 years, in whom routine  
119 investigations had not uncover any obvious cause (Coughlan *et al.* 2014). Unexplained RM was defined  
120 as three or more consecutive miscarriages before 24 weeks of gestation, with no identifiable cause after  
121 routine investigations according to an established protocol (Sarvelos & Li 2012). Fertile control subjects  
122 referred to women who had one or more live birth following spontaneous conception, stopped  
123 breastfeeding for more than 6 months, and without any history of spontaneous miscarriage, were also  
124 included as reference. In total, 14 women were recruited for transcriptome sequencing as screening, 5  
125 women with unexplained RIF, 6 women with unexplained RM, and 3 fertile. For validation, additional 10  
126 women were included, with 4 unexplained RIF, 5 unexplained RM and 1 fertile.

### 127 **Endometrial sample**

128 In the cycle of study, all subjects started daily urine LH test from day 9 of the cycle onwards until the  
129 LH surge had been identified. An endometrial biopsy was obtained on day LH+7 as an outpatient  
130 procedure with the use of a Pipelle® sampler. The samples were immediately snap-frozen and stored in  
131 liquid nitrogen for later processing.

### 132 **RNA extraction and expression calculation**

133 For the first batch of 14 recruited samples (5 with RIF, 6 with RM, 3 fertile), total RNA was extracted  
134 from endometrium by TRIzol according to manufacturer's protocols (Invitrogen). RNA quality and  
135 integrity was confirmed by NanoDrop 2000 (Thermo Scientific) and Bioanalyzer 2100 Eukaryote Total  
136 RNA Pico (Agilent Tech, Inc), respectively. All 14 extracted RNA samples were rRNA depleted by  
137 Ribozero (Illumina) and the paired-ends strand-specific libraries were prepared by TrueSeq Stranded  
138 Total RNA Library Prep Kit (Illumina). All samples were sequenced by Illumina HiSeq2000. After  
139 sequencing, low quality reads whose sequencing quality below 20 were trimmed. All reads were mapped  
140 to human genome hg38 by Tophat2(Kim *et al.* 2013) with default parameters. The Reads Per Kilobase  
141 Per Million Reads (RPKM) of gene expression was calculated based on the GENCODE v23 annotation  
142 (Harrow *et al.* 2012). All expressions were normalized by quantile normalization method using median

143 (Risso *et al.* 2014). The differential expressed genes (DEGs) were determined by two criteria: (a) the fold  
144 change between the means of groups was higher than 1.5; and (b) the p-value calculated from pooled T-  
145 test was smaller than 0.05.

## 146 **Hierarchical clustering and principle component analysis (PCA)**

147 The expressions of each gene were firstly scaled as follow:

$$148 \quad S_{ij} = \frac{R_{ij}}{\max(R_i)}$$

149 where  $S_{ij}$  is the scaled expression of gene  $i$  in sample  $j$ ,  $R_{ij}$  is the raw normalized expression of gene  $i$  in  
150 sample  $j$ ,  $\max(R_i)$  is the largest value of gene  $i$  among all samples from RM, RIF and fertile groups.  
151 Afterwards the scaled expressions will be used for unsupervised hierarchical clustering and then PCA by  
152 R packages gplots (Warnes *et al.* 2009) and prcomp.

153 After PCA was done, the vector of each principle component was calculated. In the space constructed  
154 by any two principle components  $V_p$  and  $V_q$ , the direction of classification which was vertical to the  
155 calculated boundary  $aP + bQ = c$  by SVM is  $V_D = \begin{bmatrix} a \\ b \end{bmatrix}$ , where  $a^2 + b^2 = 1$ . Thus the contribution of  
156 each gene to the classification direction was calculated by:

$$157 \quad aV_p + bV_q = \mathbf{X}$$

158 where  $\mathbf{X}$  contains the contribution of the corresponding genes to the classification direction. Support  
159 vector machine (SVM) was performed by Python library sklearn 0.17.0 (Pedregosa 2011). Genes whose  
160 absolute values of the contribution scores were larger or equal to 0.01 were considered to have significant  
161 contribution.

## 162 **Gene ontology and pathway analysis**

163 Only genes with significant differential expression were retrieved for gene ontology (GO) and pathway  
164 analysis. Pathway enrichment was analyzed by DAVID 6.7 (the Database for Annotation, Visualization  
165 and Integrated Discovery) (Risso *et al.* 2014). The pathways whose correlated p-value (q-value) smaller  
166 than 0.05 were considered significantly enriched.

## 167 **Quantitative RT-PCR**

168 In additional to the 14 sequenced samples, extra 10 independent samples (4 RIF, 5 RM, and 1 fertile)  
169 were added to measure relative gene expression using quantitative real-time RT-PCR (qPCR) for  
170 validation. TATA-box binding protein (TBP) and ribosomal protein L13a (RPL13A) were used as  
171 reference genes for expression normalization. Total RNA was extracted and quality checked as above.  
172 One microgram of total RNA was used for reverse transcription. Quantitative PCR was performed by  
173 using SYBR® Green PCR Master Mix (Applied Biosystems) with Roche LightCycler® 480 II. Primer  
174 sequences can be found in Supplementary Table 1. Wilcoxon test was used to exam the statistical  
175 significance between RIF and RM.

## 176 **Ethics**

177 This study was approved by the Joint Chinese University of Hong Kong – New Territories East Cluster  
178 Clinical Research Ethics Committee. Written consent was obtained from all participants.

179

## 180 **Results**

181 The demographics of the recruited subjects are summarized in Table 1. There was no significant  
182 difference in age, BMI, cycle length and endometrium thickness at the time of biopsy amongst RIF, RM  
183 and fertile groups ( Table 1).

184 The reads mapping of all 14 samples for RNA-Seq were satisfactory (Supplementary Figure 1). All  
185 samples had over 80% reads mapped to the human genome hg38. The raw sequencing data was uploaded  
186 to NCBI with reference BioProject ID: PRJNA314429. Firstly, we explored whether RM, RIF and fertile  
187 samples could be separated in the transcriptome profiling. After normalization and scaling of gene  
188 expressions, un-supervised hierarchical clustering was performed to all samples (Figure 1). Most RIF and  
189 RM samples were clustered to two sub-trees, while fertile samples could not be grouped and were  
190 clustered within RIF or RM samples. There were 661 genes significantly up-regulated in RIF compared  
191 with RM; whilst 301 genes up-regulated in RM compared with RIF. To further compare and contrast the  
192 differences between RIF and RM, fertile samples were excluded and principle component analysis (PCA)  
193 was performed. RIF and RM samples showed distinct spatial distribution in the three-dimensional space  
194 constructed by the first three components. In the space constructed by the first component and the third  
195 component, RIF and RM were perfectly linear separated (Figure 2). The boundary between RIF and RM  
196 could be further learnt by SVM, where the classification direction which was vertical to the linear  
197 boundary gave the best resolution to distinguish RIF and RM. To identify the genes that contribute the  
198 most to the classification direction, the contribution score for each gene was calculated (see Method). The  
199 genes with positive contribution scores showed higher expressions in RM, while genes with negative  
200 contribution scores showed higher expressions in RIF (Figure 3). Genes whose absolute values of  
201 contribution scores were larger or equal to 0.01 were considered to have significant contribution to the  
202 classification of RIF and RM, where 183 genes had significant positive contribution and 380 had  
203 significant negative contribution.

204 To investigate which biological and molecular pathway contributed most to the differential  
205 transcriptomic pattern between RIF and RM, pathway enrichment analysis was applied on genes with  
206 significant contribution on both directions (Figure 3). Pathways with  $-\log(q\text{-value}) > 2$  were considered  
207 as significant. The localization of proteins encoded by genes in both directions showed high enrichment in  
208 extracellular regions indicated by GO cellular component terms. However, the pathways where they were  
209 involved were distinct, which could be revealed by GO biological process terms and KEGG pathways  
210 (Figure 4). On negative direction, many responses to wounding and inflammatory genes were  
211 predominately enriched, including the most enriched complement and coagulation cascades in KEGG  
212 (Figure 5). Several central components of complement and coagulation cascades have significantly higher  
213 expressions in RIF than those in RM, with significant t-test p-values less than 0.05. We chose the top six  
214 over-expressed genes in RIF from complement cascade, namely C3, C4, C4BP, DAF, DF and SERPING1,  
215 for real time PCR validation (Figure 6). In contrast, on positive direction, genes that were involved in  
216 extracellular structure organization and biological adhesion by GO biological process terms and



217 neuroactive ligand-receptor interaction and calcium signaling pathways in KEGG, were significantly  
218 enriched, but only few (2/78) significant differentially expressed genes were identified.

219 Real time PCR was performed on all 24 samples to validate the differential expression of 6 genes in  
220 complement cascade (Figure 6). All the gene expressions were significantly up-regulated in RIF group  
221 compared with that of RM with Wilcoxon test, p values less than 0.05. However, the real time PCR  
222 results of control group showed great variation, and most of the difference are not significant.

223

## 224 Discussion

225 In this study, we have found that the transcriptome profiles of the two groups of patients studied (RM  
226 and RIF) are distinctively different from one another. In addition, we have found significant amounts of  
227 differentially expressed genes (661 up-regulated in RIF and 301 up-regulated in RM) and one  
228 distinctively and validated pathway between women with RIF and RM.

229 Among all the enriched pathways, the *complement and coagulation cascades pathway* was the most  
230 significantly affected pathway with  $-\log(q\text{-values}) > 4$  in KEGG pathway analysis in RIF. This particular  
231 pathway was upregulated in RIF patients compared with RM patients, the genes involved in this pathway  
232 included C3, C4, C4BP, DAF, DF and SERPING1. All of these 6 up-regulated genes had been validated  
233 with qPCR, which confirmed the up-regulation generally existed in patients with unexplained RIF.  
234 Reference to Figure 6 showed that C3 expression in RIF was significantly higher than that of control,  
235 whereas the expression in RM was not different to that of control subjects. The complement system,  
236 represented by complement component 3 (C3), is a proteolytic cascade in plasma and an upstream  
237 mediator of innate immunity. It is known that human chorionic gonadotropin (hCG) has positive effects  
238 on endometrial C3 expression (Palomino *et al.* 2013). Whilst the adverse effect of decreased C3  
239 expression on placental development and fetal development has been shown in C3 deficient mice (Chow  
240 *et al.* 2009), and variants in FOXD1 that enhance the expression of C3 were associated with miscarriage  
241 in humans and mice (Laissue *et al.* 2016). ¶The possible adverse effect of over expression of C3 has not  
242 been previously reported. It seems therefore a fine balance is necessary; both under expression as well as  
243 over expression of C3 may be detrimental. Our observation that the C3 was over expressed in RIF but not  
244 in RM suggested that the two conditions affect uterine receptivity in different ways. One of the major  
245 immune functions of C3 pathway is to form the membrane attack complex, leading to cell lysis (Ricklin *et*  
246 *al.* 2010). While DAF (complement-protective protein decay-accelerating factor, also known as CD55)  
247 was considered as an inhibitor to the increased complement activity (Young *et al.* 2002) and expression of  
248 DAF was minimal in the proliferative and early secretory phase in endometrium, increased to a maximum  
249 on LH+7, and decreased until next cycle. Endometrial C3 and DAF expression was associated with  
250 human chorionic gonadotropin, indicating its roles in early embryo development (Palomino *et al.* 2013).  
251 Whilst there has not been any study in the literature which reported on the expression of any of the genes  
252 involved in this pathway in the endometrium of RIF or RM at the time of implantation, previous genetic  
253 association studies found the loss of functional mutation of some genes in this pathway were associated  
254 with RM (Mohlin *et al.* 2013) or other adverse pregnancy outcome, such as preeclampsia (Salmon *et al.*  
255 2011). Although both C3 and DAF were upregulated in RIF when compared with RM, the increased C3  
256 expression (3.5 folds) was higher than the increased DAF expression (2.8 folds), suggesting the inhibitory

257 complement system in RIF may be more likely a reactive response. Further studies of its inhibitory  
258 mechanism and subsequent downstream innate immune response in RIF are needed.

259 On the other hand, though the neuroactive ligand-receptor interaction pathway and calcium signaling  
260 pathway were enriched in RM according to the positively contributed gene list from SVM analysis, we  
261 did not consider them as important as the complement and coagulation cascades pathway as discussed  
262 above. Firstly, most of the genes with positive contribution in these two pathways were expressed at very  
263 low expression level. Furthermore, almost no differentially expressed genes were identified in those two  
264 pathways. One explanation for lack of significantly expressed genes in calcium pathway in our study  
265 could be that this activity could be prominent in endometrial epithelium cells (Thie & Denker 2002,  
266 Brosens *et al.* 2014, Ruan *et al.* 2014), which might be diluted if sequencing endometrium tissue as a  
267 whole. However, it has long been known that  $\text{Ca}^{2+}$  channels involves in a variety of implantation  
268 processes and increased  $\text{Ca}^{2+}$  mobilization can assist blastocyst-endometrium adhesion (Thie & Denker  
269 2002, Brosens *et al.* 2014, Ruan *et al.* 2014). Brosen *et al.* found that competent and low-quality embryos  
270 elicited different  $\text{Ca}^{2+}$  channel responses in vitro, which indicate the active role of endometrial selective  
271 function of human embryos (Brosens *et al.* 2014). And this selection was impaired in the endometrium of  
272 RM subjects (Teklenburg *et al.* 2010). In this study, the genes with positive contribution in calcium  
273 signaling pathway may suggest the higher activity of the  $\text{Ca}^{2+}$  channel in the endometrium of RM  
274 compared with that of RIF. It might indicate that the endometrium during WOI is more favorable for  
275 implantation in RM compared with RIF, which would also be in consistent with in vitro study carried out  
276 by Brosens *et al.* that pattern of  $\text{Ca}^{2+}$  signals was associated with the implantation results.

277 According to the hypothesis put forward by Brosen and Macklon *et al.* (Teklenburg *et al.* 2010, Brosens  
278 *et al.* 2014, Macklon & Brosens 2014), women with unexplained RM would be superfertile because the  
279 endometrium is over-receptive, less able to discern and prevent the abnormal embryos from implantation,  
280 in contrast to that of women with RIF, in which the abnormality makes it difficult for even the normal  
281 embryo to implant. Whilst our findings do not directly confirm or refute the Brosen hypothesis that the  
282 endometrium is over-receptive in women with RM, our finding regarding the differential regulation of the  
283 pathways between the two groups of women may provide insight into the molecular mechanism  
284 controlling the implantation process in the endometrium to make it under-receptive (as in women with  
285 RIF) or over-receptive (as in women with RM).

286 The transcriptome pattern of fertile women seems dispersedly distributed among RIF and RM subjects,  
287 but the sample size in our study indeed is very small to make any conclusion. The dispersed distribution  
288 may be due to the heterogeneity of endometrial receptivity status. Although they were classified as fertile  
289 controls as they had previously successful pregnancy in early years, unfortunately it does not necessarily  
290 imply they will still be able to achieve successful pregnancy if conceived. This is one of the limitations of  
291 our study.

292 A particular strength of our study is the precise timing of the endometrial specimen, all obtained on day  
293 LH+7. Whilst some earlier studies did time the specimen precisely on a single day (Diaz-Gimeno *et al.*  
294 2011, Hu *et al.* 2014) others obtained the specimen over two or more days (Ledee *et al.* 2011, Koot *et al.*  
295 2016). Given that the endometrium changes very rapidly around the time of implantation, the inclusion of  
296 samples collected on different days after the LH surge could introduce a significant source of variance to  
297 the results. It may help to explain why, in a previous study by Ledee *et al.* which studied similar subjects

298 groups as in our study but with biopsies obtained over a three-day period from days LH+7 to +9, they  
299 could find gene expression differences between RM and RIF, consistent with our findings, but not able to  
300 identify the pathways (Ledee *et al.* 2011).

301 Another strength of this study is that we used RNA-seq rather than micro-array to analyze the  
302 specimens. Earlier transcriptome studies of the endometrium used micro-array analysis (Ledee *et al.* 2011,  
303 Othman *et al.* 2012, Ruiz-Alonso *et al.* 2013, Koot *et al.* 2016) although 2 recent studies did use  
304 sequencing techniques (Hu *et al.* 2014, Kosova *et al.* 2015). It is now well accepted that sequencing  
305 technique is more comprehensive in coverage and precise in quantification of global gene expression  
306 profiles (McGettigan 2013). Furthermore, in our study we have used more straightforward and more  
307 comprehensive methods to mine the features which contributed to the classification of the two groups of  
308 women, and thus identified genes and pathways that were differentially expressed between RIF and RM.  
309 The chosen testing platforms and analysis methods could also greatly contributed to the identification of  
310 significant pathways.

311 One possible limitation of our study is the relatively small sample size (RIF=9, RM=11, control=4) and  
312 so the conclusions reached in this study should be considered preliminary, especially in view of the  
313 potential heterogeneity of the study populations.

314 To conclude, we have identified that the complement and coagulation cascades pathway are  
315 significantly different between women with RM and RIF. The identified pathways provide an insight into  
316 how the process of implantation in these two types of implantation disorder differs from one other.

317

## 318 **Authors' roles**

319 J.H. and T.C.L. designed the study, prepared the samples, interpreted the data and wrote the manuscript.  
320 H.Q. and T.F.C. performed the bioinformatics analysis and prepared the figures. Y.Y. and J.Z. performed  
321 the qPCR validation. X.C. helped on patient recruitment and sample collection. C.C.W. and S.L.  
322 contributed to the interpretation of data and the manuscript preparation. All of the authors contributed to  
323 finalizing of the manuscript.

324 **References**

- 325 **Altmae S, Martinez-Conejero JA, Salumets A, Simon C, Horcajadas JA & Stavreus-Evers A** 2010  
326 Endometrial gene expression analysis at the time of embryo implantation in women with  
327 unexplained infertility. *Mol Hum Reprod* **16** 178-187.
- 328 **Brosens JJ, Salker MS, Teklenburg G, Nautiyal J, Salter S, Lucas ES, Steel JH, Christian M, Chan YW,**  
329 **Boomsma CM, Moore JD, Hartshorne GM, Sucurovic S, Mulac-Jericevic B, Heijnen CJ, Quenby S,**  
330 **Koerkamp MJ, Holstege FC, Shmygol A & Macklon NS** 2014 Uterine selection of human  
331 embryos at implantation. *Sci Rep* **4** 3894.
- 332 **Chow WN, Lee YL, Wong PC, Chung MK, Lee KF & Yeung WS** 2009 Complement 3 deficiency impairs  
333 early pregnancy in mice. *Mol Reprod Dev* **76** 647-655.
- 334 **Coughlan C, Ledger W, Wang Q, Liu F, Demirolo A, Gurgan T, Cutting R, Ong K, Sallam H & Li TC** 2014  
335 Recurrent implantation failure: definition and management. *Reprod Biomed Online* **28** 14-38.
- 336 **Coughlan C, Sinagra M, Ledger W, Li TC & Laird S** 2013 Endometrial integrin expression in women with  
337 recurrent implantation failure after in vitro fertilization and its relationship to pregnancy  
338 outcome. *Fertil Steril* **100** 825-830.
- 339 **Diaz-Gimeno P, Horcajadas JA, Martinez-Conejero JA, Esteban FJ, Alama P, Pellicer A & Simon C** 2011 A  
340 genomic diagnostic tool for human endometrial receptivity based on the transcriptomic  
341 signature. *Fertil Steril* **95** 50-60, 60 e51-15.
- 342 **Fitzgerald JS, Poehlmann TG, Schleussner E & Markert UR** 2008 Trophoblast invasion: the role of  
343 intracellular cytokine signalling via signal transducer and activator of transcription 3 (STAT3).  
344 *Hum Reprod Update* **14** 335-344.
- 345 **Germeyer A, Savaris RF, Jauckus J & Lessey B** 2014 Endometrial beta3 integrin profile reflects  
346 endometrial receptivity defects in women with unexplained recurrent pregnancy loss. *Reprod*  
347 *Biol Endocrinol* **12** 53.
- 348 **Haouzi D, Assou S, Mahmoud K, Tondeur S, Reme T, Hedon B, De Vos J & Hamamah S** 2009 Gene  
349 expression profile of human endometrial receptivity: comparison between natural and  
350 stimulated cycles for the same patients. *Hum Reprod* **24** 1436-1445.
- 351 **Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A,**  
352 **Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-**  
353 **Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J,**  
354 **Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M,**  
355 **Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigo R & Hubbard TJ** 2012  
356 GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22**  
357 1760-1774.
- 358 **Hu S, Yao G, Wang Y, Xu H, Ji X, He Y, Zhu Q, Chen Z & Sun Y** 2014 Transcriptomic changes during the  
359 pre-receptive to receptive transition in human endometrium detected by RNA-Seq. *J Clin*  
360 *Endocrinol Metab* **99** E2744-2753.
- 361 **Karaer A, Cigremis Y, Celik E & Urhan Gonullu R** 2014 Prokineticin 1 and leukemia inhibitory factor  
362 mRNA expression in the endometrium of women with idiopathic recurrent pregnancy loss. *Fertil*  
363 *Steril* **102** 1091-1095 e1091.
- 364 **Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R & Salzberg SL** 2013 TopHat2: accurate alignment of  
365 transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14** R36.
- 366 **Koler M, Achache H, Tsafirir A, Smith Y, Revel A & Reich R** 2009 Disrupted gene pattern in patients with  
367 repeated in vitro fertilization (IVF) failure. *Hum Reprod* **24** 2541-2548.
- 368 **Koot YE, van Hooff SR, Boomsma CM, van Leenen D, Groot Koerkamp MJ, Goddijn M, Eijkemans MJ,**  
369 **Fauser BC, Holstege FC & Macklon NS** 2016 An endometrial gene expression signature  
370 accurately predicts recurrent implantation failure after IVF. *Sci Rep* **6** 19411.

371 **Kosova G, Stephenson MD, Lynch VJ & Ober C** 2015 Evolutionary forward genomics reveals novel  
372 insights into the genes and pathways dysregulated in recurrent early pregnancy loss. *Hum*  
373 *Reprod* **30** 519-529.

374 **Laissue P, Lakhali B, Vatin M, Batista F, Burgio G, Mercier E, Santos ED, Buffat C, Sierra-Diaz DC,**  
375 **Renault G, Montagutelli X, Salmon J, Monget P, Veitia RA, Mehats C, Fellous M, Gris JC,**  
376 **Cocquet J & Vaiman D** 2016 Association of FOXD1 variants with adverse pregnancy outcomes in  
377 mice and humans. *Open Biol* **6**.

378 **Ledee N, Munaut C, Aubert J, Serazin V, Rahmati M, Chaouat G, Sandra O & Foidart JM** 2011 Specific  
379 and extensive endometrial deregulation is present before conception in IVF/ICSI repeated  
380 implantation failures (IF) or recurrent miscarriages. *J Pathol* **225** 554-564.

381 **Macklon NS & Brosens JJ** 2014 The human endometrium as a sensor of embryo quality. *Biol Reprod* **91**  
382 98.

383 **Mariee N, Li TC & Laird SM** 2012 Expression of leukaemia inhibitory factor and interleukin 15 in  
384 endometrium of women with recurrent implantation failure after IVF; correlation with the  
385 number of endometrial natural killer cells. *Hum Reprod* **27** 1946-1954.

386 **McGettigan PA** 2013 Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol* **17** 4-11.

387 **Mirkin S, Nikas G, Hsiu JG, Diaz J & Oehninger S** 2004 Gene expression profiles and structural/functional  
388 features of the peri-implantation endometrium in natural and gonadotropin-stimulated cycles. *J*  
389 *Clin Endocrinol Metab* **89** 5742-5752.

390 **Mohlin FC, Mercier E, Fremeaux-Bacchi V, Liszewski MK, Atkinson JP, Gris JC & Blom AM** 2013 Analysis  
391 of genes coding for CD46, CD55, and C4b-binding protein in patients with idiopathic, recurrent,  
392 spontaneous pregnancy loss. *Eur J Immunol* **43** 1617-1629.

393 **Othman R, Omar MH, Shan LP, Shafiee MN, Jamal R & Mokhtar NM** 2012 Microarray profiling of  
394 secretory-phase endometrium from patients with recurrent miscarriage. *Reprod Biol* **12** 183-199.

395 **Palomino WA, Argandona F, Azua R, Kohen P & Devoto L** 2013 Complement C3 and decay-accelerating  
396 factor expression levels are modulated by human chorionic gonadotropin in endometrial  
397 compartments during the implantation window. *Reprod Sci* **20** 1103-1110.

398 **Pedregosa** 2011 Scikit-learn: Machine Learning in Python. *JMLR* **12**, pp2825-2830.

399 **Qiao J, Wang L, Li R & Zhang X** 2008 Microarray evaluation of endometrial receptivity in Chinese women  
400 with polycystic ovary syndrome. *Reprod Biomed Online* **17** 425-435.

401 **Ricklin D, Hajishengallis G, Yang K & Lambris JD** 2010 Complement: a key system for immune  
402 surveillance and homeostasis. *Nat Immunol* **11** 785-797.

403 **Risso D, Ngai J, Speed TP & Dudoit S** 2014 Normalization of RNA-seq data using factor analysis of  
404 control genes or samples. *Nat Biotechnol* **32** 896-902.

405 **Ruan YC, Chen H & Chan HC** 2014 Ion channels in the endometrium: regulation of endometrial  
406 receptivity and embryo implantation. *Hum Reprod Update* **20** 517-529.

407 **Ruiz-Alonso M, Blesa D, Diaz-Gimeno P, Gomez E, Fernandez-Sanchez M, Carranza F, Carrera J, Vilella F,**  
408 **Pellicer A & Simon C** 2013 The endometrial receptivity array for diagnosis and personalized  
409 embryo transfer as a treatment for patients with repeated implantation failure. *Fertil Steril* **100**  
410 818-824.

411 **Salmon JE, Heuser C, Triebwasser M, Liszewski MK, Kavanagh D, Roumenina L, Branch DW, Goodship T,**  
412 **Fremeaux-Bacchi V & Atkinson JP** 2011 Mutations in complement regulatory proteins  
413 predispose to preeclampsia: a genetic analysis of the PROMISSE cohort. *PLoS Med* **8** e1001013.

414 **Saravelos SH & Li TC** 2012 Unexplained recurrent miscarriage: how can we explain it? *Hum Reprod* **27**  
415 1882-1886.

416 **Teklenburg G, Salker M, Heijnen C, Macklon NS & Brosens JJ** 2010 The molecular basis of recurrent  
417 pregnancy loss: impaired natural embryo selection. *Mol Hum Reprod* **16** 886-895.

418 **Thie M & Denker H-W** 2002 In vitro studies on endometrial adhesiveness for trophoblast: cellular  
419 dynamics in uterine epithelial cells. *Cells Tissues Organs* **172** 237-252.

420 **Tuckerman E, Laird SM, Prakash A & Li TC** 2007 Prognostic value of the measurement of uterine natural  
421 killer cells in the endometrium of women with recurrent miscarriage. *Hum Reprod* **22** 2208-2213.

422 **Tuckerman E, Mariee N, Prakash A, Li TC & Laird S** 2010 Uterine natural killer cells in peri-implantation  
423 endometrium from women with repeated implantation failure after IVF. *J Reprod Immunol* **87**  
424 60-66.

425 **Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M,**  
426 **Magnusson A & Moeller S** 2009 gplots: Various R programming tools for plotting data. *R*  
427 *package version 2*.

428 **Xu B, Sun X, Li L, Wu L, Zhang A & Feng Y** 2012 Pinopodes, leukemia inhibitory factor, integrin-beta3,  
429 and mucin-1 expression in the peri-implantation endometrium of women with unexplained  
430 recurrent pregnancy loss. *Fertil Steril* **98** 389-395.

431 **Young SL, Lessey BA, Fritz MA, Meyer WR, Murray MJ, Speckman PL & Nowicki BJ** 2002 In vivo and in  
432 vitro evidence suggest that HB-EGF regulates endometrial expression of human decay-  
433 accelerating factor. *J Clin Endocrinol Metab* **87** 1368-1375.

434