Guha, Pokhraj; Roychoudhury, Shubhadeep; Singha, Sobita; Kalita, Jogen C.; Kolesarova, Adriana; Jamal, Qazi Mohammad Sajid; Jha, Niraj Kumar; Kumar, Dhruv; Ruokolainen, Janne; Kesari, Kavindra Kumar

A comparative cross-platform meta-analysis to identify potential biomarker genes common to endometriosis and recurrent pregnancy loss

Published in:
Applied Sciences (Switzerland)

DOI:
10.3390/app11083349

Published: 02/04/2021

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Guha, P., Roychoudhury, S., Singha, S., Kalita, J. C., Kolesarova, A., Jamal, Q. M. S., Jha, N. K., Kumar, D., Ruokolainen, J., & Kesari, K. K. (2021). A comparative cross-platform meta-analysis to identify potential biomarker genes common to endometriosis and recurrent pregnancy loss. Applied Sciences (Switzerland), 11(8), [3349]. https://doi.org/10.3390/app11083349
A Comparative Cross-Platform Meta-Analysis to Identify Potential Biomarker Genes Common to Endometriosis and Recurrent Pregnancy Loss

Pokhraj Guha 1, Shubhadeep Roychoudhury 2,*, Sobita Singha 2, Jogen C. Kalita 3, Adriana Kolesarova 4, Qazi Mohammad Sajid Jamal 5,*, Niraj Kumar Jha 6, Dhruv Kumar 7, Janne Ruokolainen 8 and Kavindra Kumar Kesari 8,*

Citation: Guha, P.; Roychoudhury, S.; Singha, S.; Kalita, J.C.; Kolesarova, A.; Jamal, Q.M.S.; Jha, N.K.; Kumar, D.; Ruokolainen, J.; Kesari, K.K. A Comparative Cross-Platform Meta-Analysis to Identify Potential Biomarker Genes Common to Endometriosis and Recurrent Pregnancy Loss. Appl. Sci. 2021, 11, 3349. https://doi.org/10.3390/app11083349

Abstract: Endometriosis is characterized by unwanted growth of endometrial tissue in different locations of the female reproductive tract. It may lead to recurrent pregnancy loss, which is one of the worst curses for the reproductive age group of human populations around the world. Thus, there is an urgent need for unveiling any common source of origin of both these diseases and connections, if any. Herein, we aimed to identify common potential biomarker genes of these two diseases via in silico approach using meta-analysis of microarray data. Datasets were selected for the study based on certain exclusion criteria. Those datasets were subjected to comparative meta-analyses for the identification of differentially expressed genes (DEGs), that are common to both diagnoses. The DEGs were then subjected to protein-protein networking and subsequent functional enrichment analyses for unveiling their role/function in connecting two diseases. From the analyses, 120 DEGs are reported to be significant out of which four genes have been found to be prominent. These include the \textit{CTNNB1}, \textit{HNRNP AB}, \textit{SNRPF} and \textit{TWIST2} genes. The significantly enriched pathways based on the above-mentioned genes are mainly centered on signaling and developmental events. These findings could significantly elucidate the underlying molecular events in endometriosis-based recurrent miscarriages.

Keywords: endometriosis; recurrent pregnancy loss; meta-analysis; functional enrichment; TWIST2 gene

1. Introduction

Endometriosis is commonly known as a chronic condition that has been characterized by the growth of endometrial tissue in sites other than the endometrium [1]. This may result in the abnormal growth of endometrial cells outside the uterus and cause a painful condition. According to NHS-UK, symptoms include severe pelvic pain during periods, sex, urination and defecation. Major symptoms could be constipation, diarrhea, and even blood during urination. Women also face difficulties in getting pregnant...
Appl. Sci. 2021, 11, 3349

(https://www.nhs.uk/conditions/endometriosis/, accessed on 20 July 2020). After several years of research, the pathogenesis of endometriosis is still not clear [2]. The existence of endometriosis has been found from Müllerian or non-Müllerian stem cells, which may include those from bone marrow, the endometrial basal layer, the peritoneum, or Müllerian remnants [3]. In addition, scientists believe that dysregulation of the canonical Wnt/β-signaling pathway could be responsible for the endometriotic lesions leading to the endometriosis condition [2]. Wnt/β-catenin signaling also has a role in governing the endometrial cells regulated by estrogen and progesterone. Any changes in the expression of estrogen and progesterone receptors may cause progesterone resistance in endometriosis patients [4]. Infertility problems may be caused due to recurrent pregnancy loss which has been found a major issue in endometriosis patients. Indeed, the loss of two or more pregnancies has also been reported by the European Society of Human Reproduction and Embryology Recurrent Pregnancy Loss (RPL) [5], where ectopic pregnancy and molar pregnancy has been excluded. Endometriosis-associated infertility could be identified by potential markers, such as inflammatory cytokines, iron and oxidative stress, oxidant-antioxidant imbalance, and iron-dependent progression of endometriosis [6]. A recent literature review suggested that endometrial immune dysregulation could be responsible for RPL and may also lead to endometriosis [7]. Thus far, the exact reason for endometriosis is still not clear and, therefore meta-data analysis may provide further knowledge to solve the molecular pathogenesis complexity of such condition(s). To find the genes involved in the loss of the hormonal functions and association with endometriosis, Sapkota et al. [8] performed a large scale, 11 genome-wide case-control dataset meta-analysis and found that FN1, CCDC170, ESR1, SYNE1, and FSHB are the 5 genes that could be responsible for the endometriosis risk. Therefore, computational system biology plays a major role in meta-data analysis. In combination with machine learning, many biomarker genes have been identified, including NOTCH3, SNAPC2, B4GALNT1, SMAP2, DDB2, GTF3C5, and PTOV1 from the transcriptomic data analysis, and TRPM6, RASSF2, TNIP2, RP3-522J7.6, FGD3, and MFSD14B from the methylomic data analysis [9]. The latest metadata investigation related to polymorphisms and endometriosis tried to find the genetic level reason behind endometriosis, where five polymorphisms have been associated with endometriosis [10]. They were glutathione S-transferase pi 1 (GSTPI) rs1695, interferon-gamma (IFNG) (CA) repeat, wingless-type MMTV integration site family member 4 (WNT4) rs16826658, rs2235529, and glutathione S-transferase mu 1 (GSTM1) null genotype. The present study aimed to identify the genes that are differentially expressed in endometriosis and RPL conditions, and to elucidate their involvement in protein–protein interactions, as well as their functional importance in biological pathways as potential biomarkers common to both endometriosis and RPL.

2. Materials and Methods

2.1. Microarray Data

Suitable gene expression microarray samples were obtained from the NCBI Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/, accessed on 20 July 2020) [11]. A thorough search was performed of the GEO database from July 2020 to September 2020 (3 months) using the keywords “Endometriosis AND Recurrent Pregnancy Loss”. The GEO datasets that were included in our study are GSE7305, GSE23339, GSE26787, GSE58178 and GSE111974 subject to their fulfillment of certain criteria. The gene expression profiling was based on endometrial tissue and each dataset contained sufficient data to perform a meta-analysis. The following inclusion criteria were imposed while selecting the datasets for the meta-analyses: (i) the sample type must be endometrial tissue only, (ii) datasets should not contain overlapping sample sets, (iii) datasets must not have been generated from the same research laboratory, and (iv) they are heterogeneous in terms of microarray platform (Table 1). The datasets that met these inclusion criteria were selected for the present study.
2.2. DEG Screening and Meta-Analyses

Analyses of microarray expression data were performed using the ExAtlas meta-analyses software [12]. The expression profiles of the 5 GEO datasets that were included in our study were extracted from the GEO database.

Normalization of the data was carried out using the quantile method. Each dataset was saved separately and later combined using the batch normalization method. Gene-specific batch normalization can be used to combine two or more datasets. If two datasets include the same tissue or organ then median expression levels for this common tissue/organ are equalized in the two datasets using this method.

ExAtlas uses the same algorithm for statistical analysis as NIA Array Analysis [13]. Gene expression values are log-transformed and used for ANOVA [13], which is modified for the multiple hypotheses testing case. Additionally, the false discovery rate (FDR) [2] is used to assess the significance of gene expression change instead of p-values. Later meta-analyses were performed on the saved datasets using a random effect method and lists of differentially-expressed genes were saved as a gene set file. The random-effects method takes into account the variance of heterogeneity between studies, which is added to the variance of individual effects. Here, term effect means the log ratio of gene expression change/difference compared to control or study-wide mean or median.

In a parallel manner, the same raw datasets were analyzed with another software named Network analyst 3.0 [13]. Upon combining the datasets after normalization, 17,347 matched feature numbers were recognized, which were then subjected to batch effect adjustment using Combat. Then, meta-analyses were carried out on the combined dataset using a random effect model with the p-value set to less than 0.05 and FDR to less than or equal to 2. While FDR can be a great indicator of the strength of a study, the p-value can be more useful for statistical power analyses in future studies. The Limma package [14] was used for the identification of differentially expressed genes (DEGs).

Furthermore, gene expression analyses were performed on all the datasets individually using Geo2R [3]. Quantile normalization was performed and the Benjamini and Hochberg false discovery rate method was selected by default for Geo2R analyses because it is the most commonly used adjustment for microarray data and provides a good balance between the discovery of statistically significant genes and limitation of false positives.

2.3. Comparative Analyses

The DEGs from both the analyses were then compared and then the common genes were marked. These genes have the annotation set to official gene symbol, which was corrected using db2db tool of the Biological Database Network [15]. Furthermore, the gene expression outputs of all the datasets generated using Geo2R [11] were compared and the common DEGs were recorded, which were also compared with the output of ExAtlas and Network Analyst 3.0. The DEGs were then used to construct a heatmap using the ComplexHeatmap package of R [16].

2.4. Protein–Protein Interaction Network Construction and Pathway Enrichment Analyses

2.4.1. Protein–Protein Network Interaction

Additionally, DEGs have also been used to study the protein–protein interactions using the STRING app [17] of Cytoscape [18]. The protein–protein interaction network was developed in the STRING app. The meaning of the network edges was set to evidence-based analyses. The second shell interactors were added to the network to ensure or visualize connections between our target proteins, which were too weak to be found. The 1st shell interactors were the proteins directly associated with the input protein(s) while the 2nd shell of interactors were the proteins associated with the proteins from the 1st shell. It can be the case that a 2nd shell protein can be directly connected to an input protein(s), but it will usually have a weaker association and therefore it would not show up among the specified number of 1st shell interactors. The 2nd shell proteins are always grey. The generated network was then analyzed using the Network Analyzer function of Cytoscape.
2.4.2. Pathway Enrichment Analysis

Furthermore, the biological processes that are involved with the DEGs and the functional enrichment analysis were also studied using the BINGO app [19] of Cytoscape. A hypergeometric test was carried out using Benjamini and Hochberg FDR correction. The GO Biological process was selected as the ontology file for executing enrichment analyses. The generated network was then analyzed using the network analyzer function of Cytoscape.

Table 1. List of the datasets that have been included in the study.

| Sl. No. | GEO Accession | Patient | Control | Total | Sample | Analytical Platform | Patient Type | Reference |
|---------|---------------|---------|---------|-------|--------|---------------------|--------------|-----------|
| 1       | GSE58178      | 6       | 6       | 12    | Endometrial tissue | GPL6947 (Illumina Human HT-12 v3.0 Expression Beadchip) | Endometriosis | [20]      |
| 2       | GSE23339      | 10      | 9       | 19    | Endometrial tissue | GPL6102 (Illumina Human-6 v2.0 Expression Beadchip) | Endometriosis | [21]      |
| 3       | GSE7305       | 10      | 10      | 20    | Endometrial tissue | GPL570 [HG-U133_Plus_2] (Affymetrix Human Genome U133 plus 2.0 Array) | Endometriosis | [22]      |
| 4       | GSE111974     | 24      | 24      | 48    | Endometrial tissue | GPL17077 (Agilent-039494 SurePrint G3 Human GE v2 8 × 60K Microarray) | Recurrent Pregnancy Loss | [23]      |
| 5       | GSE26787      | 10      | 5       | 15    | Endometrium | GPL570 [HG-U133_Plus_2] (Affymetrix Human Genome U133 Plus 2.0 Array) | Recurrent Pregnancy Loss | [24]      |

GEO—Gene Expression Omnibus.

The overall presentation of the methods used in this study is present in Figure 1.

Figure 1. A diagram illustrating the workflow for the identification of potential biomarker genes common to endometriosis and recurrent pregnancy loss, introducing an in silico approach.
3. Results

Five microarray datasets met the inclusion criteria and have been included in our study namely, GSE58178, GSE23339, GSE7305, GSE111974 and GSE26787 (Table 1). Altogether, these 5 datasets consisted in 114 samples, of which 54 were controls, and the remaining 60 were patient samples (34 EMS and 26 RPL subjects). Box plots representing the value distribution of these five datasets, which were constructed using Geo2R. The plot shows that the log₂ expression values are normalized across all the samples of each dataset with the median line having more or less equal distribution for each dataset (Figure 2).

![Box plots of gene expression values for different datasets](image)

**Figure 2.** Value distribution (box plots) performed in GEO2R of the 5 datasets of endometriosis (GSE7305 (A), GSE58178 (C), GSE23339 (D)) and recurrent pregnancy loss (GSE26787 (B) and GSE111974 (E)) displaying the distribution of expression values of each sample within a dataset. The plot is useful for determining whether the dataset is normalized, i.e., the value distributions are median-centered across samples.

3.1. Expression of Up- and Down-Regulated Genes

Meta-analyses of selected microarray datasets using ExAtlas software estimated 207 significant genes using a random-effect model, of which 109 genes were down-regulated and 98 genes were up-regulated in the patients (both endometriosis and RPL patients taken together) compared to healthy controls. Figure 3 shows clustered heatmaps of the five datasets comprising the expression of the up-regulated and down-regulated DEGs. Based on the expression values of the DEGs, the datasets are clearly clustered into two groups, namely endometriosis and RPL. It is evident from Figure 3 that both the groups—endometriosis and RPL—have a similar pattern of expression of genes. In Figure 3, effect value refers to the log ratio of gene expression change/difference compared to control or study-wide mean or median.

![Clustered heatmaps of gene expression values](image)
NA analysis revealed 685 DEGs, of which 236 were up-regulated and 449 were down-regulated. When the results of both EA and NA were compared, 120 genes were found to be common. The top 25 DEGs from the above-mentioned 120 genes are listed in Table 2 based on their fold change (FC) values along with their Entrez ID, log-ratio combined and FDR value. Interestingly, among all the DEGs, the TWIST2 gene was found to possess the highest fold change value (3.494), which can be considered as a significant observation since the same gene has been found to have the highest fold change value in the case of NA analyses. Among these top 25 DEGs, 60% were down-regulated as evident from their log-ratio combined value while the rest 40% were up-regulated (Table 2). Thus, the down-regulated genes outweighed the scale as compared to the up-regulated genes.

**Table 2.** Top 25 up-regulated and down-regulated genes of the microarray meta-analyses along with their fold change values.

| Gene Symbol | Entrez ID | Log Ratio Combined | Fold Change | FDR * |
|-------------|-----------|-------------------|-------------|-------|
| 1.1 TWIST2  | Twist Family Bhhl Transcription Factor 2 | -0.5434 | 3.494 | $8.79 \times 10^{-11}$ |
| CA12        | Carboxylic Anhydrase XII | -0.5111 | 3.244 | 0.001487 |
| PGK6      | PiggyBac Transposable Element Derived 5 | -0.4782 | 3.007 | 0.002422 |
| H19        | H19, Imprinted Maternally Expressed Transcript (Non-Protein Coding) | -0.4696 | 2.948 | 0.000894 |
| SGCD        | Sarcoglycan Delta | 0.4523 | 2.833 | 0.0 |
| ANO4        | Anoatin 4 | -0.4227 | 2.647 | $2.42 \times 10^{-5}$ |
| CHN2        | Chimerin 2 | 0.4002 | 2.513 | $8.01 \times 10^{-7}$ |
| MLPH        | Melanophilin | -0.3955 | 2.486 | $3.27 \times 10^{-6}$ |
| PLPP1      | Phospholipid Phosphatase 1 | -0.3872 | 2.439 | 0.004665 |
| NR4A2      | Nuclear Receptor Subfamily 4 Group A Member 2 | 0.3829 | 2.415 | 0.0217 |
| DACH1      | Dachshund Family Transcription Factor 1 | -0.3827 | 2.414 | $3.07 \times 10^{-8}$ |
| ADAMTS19   | ADAM Metallopeptidase With Thrombospondin Type 1 Motif 19 | -0.3787 | 2.392 | 0.004645 |
| VLDLR      | Very Low-Density Lipoprotein Receptor | 0.3534 | 2.256 | 0.007674 |
| NF1B       | Nuclear Factor 1/B | 0.3519 | 2.249 | $4.80 \times 10^{-6}$ |
| PCSK6      | Proprotein Convertase Subtilisin/Kexin Type 6 | 0.3468 | 2.223 | 0.0154 |
| GALNT10    | N-Acetylgalactosaminyltransferase 10 | 0.334 | 2.158 | 0 |
| TGM2       | Transglutaminase 2 | -0.3236 | 2.107 | 0.006722 |
| CREG1      | Cellular Repressor Of E1A-Stimulated Genes 1 | 0.3113 | 2.048 | 0.0175 |
In a parallel workflow, all the target GEO datasets were analyzed using Geo2R. The expression profiles contained genes that were significantly expressed in comparison to the control. Following this, the expression profiles of all the datasets were overlapped using the Venn diagram (Figure 4A); it was seen that only 19 significantly overexpressed genes were common among the five datasets. Interestingly, when these 19 genes were compared with the differentially expressed genes from EA and NA analysis results (Figure 4B), then, surprisingly, only a single gene, i.e., TWIST2, was found to be commonly present among all the three analyses, viz. EA, NA and Geo2R. This outcome shifted our focus towards the TWIST2 gene and triggered our interest in exploring the biological role of this marker, especially in the context of human reproductive health. It should be noted here, with respect to Figure 4B, that all the genes that are considered for comparative analyses between the three different software-based approaches demonstrated significant fold change in the patient sample compared to the control.

In Table 2, we have shown the list of genes that were found to be differentially expressed across the three software programs, ExAtlas, Network Analyst and Geo2R. Only 1 gene was found common among the individual results amongst the 5 datasets; (obtained across the 3 programs—TWIST2). The genes were ranked based on their fold change and statistical significance (FDR). The top three genes with the highest fold change were HOXD11 (2.009), NDRG2 (1.995) and R-Spondin 3 (2.014). Additionally, the genes with the lowest fold change were Carboxypeptidase X (M14 Family), Member 1 (0.2865), Fibulin 7 (0.2862) and Homeobox D11 (0.2822).

### Table 2. Cont.

| Gene Symbol | Entrez ID | Log Ratio Combined | Fold Change | FDR * |
|-------------|-----------|--------------------|-------------|-------|
| NDRG2       | NDRG Family Member 2 | 0.31              | 2.042       | 1.71 × 10⁻⁵ |
| H4C3        | H4 Clustered Histone 3 | −0.304           | 2.014       | 4.67 × 10⁻⁷ |
| RSP03       | R-Spondin 3 | −0.3029           | 2.009       | 0.004831 |
| TSPAN2      | Tetraspanin 2 | 0.2999           | 1.995       | 0.0251   |
| CPXM1       | Carboxypeptidase X (M14 Family), Member 1 | −0.2865 | 1.934 | 4.13 × 10⁻⁶ |
| FBLN7       | Fibulin 7 | −0.2862           | 1.933       | 5.63 × 10⁻⁶ |
| HOXD11      | Homeobox D11 | −0.2822 | 1.915 | 0.0406   |

* FDR: False discovery rate.
3.2. Protein–Protein Interaction (PPI) Network

The PPI network for the DEGs is illustrated in Figure 5. The size of the node indicates the connection degree value. Centrality is an important parameter in a signaling network since it helps us to estimate the importance of a node/edge in the flow of information. It is considered an important parameter while exploring drug targets. The degree of the nodes can be used as a rough estimate of centrality. The top 20 query nodes, based on the descending order of their degree of centrality, are listed in Table 3, along with their respective betweenness centrality, closeness centrality, and the average shortest path length. A small nuclear ribonucleoprotein F (SNRPF), had the highest degree of node (84) followed by Catenin Beta 1 (CTNNB1) and Heterogeneous Nuclear Ribonucleoprotein A/B (HNRNPAB) with their degrees of nodes being 54 and 50, respectively.

![Figure 5. Protein–protein interaction (PPI) network of endometriosis and recurrent pregnancy loss genes, performed using the STRING app of Cytoscape. The size of the node indicates the connection degree value. Colored nodes represent the most common 120 genes.](image-url)
Table 3. List of top 20 interactions from protein–protein analyses using the STRING app.

| Name      | Average Shortest Path Length | Betweenness Centrality | Closeness Centrality | Clustering Coefficient | Degree |
|-----------|------------------------------|------------------------|----------------------|------------------------|--------|
| SNRPF     | 2.145228                     | 0.001357               | 0.466151             | 0.883247               | 84     |
| CTNNB1    | 1.929461                     | 0.065249               | 0.51828              | 0.26485                | 54     |
| HNRNPAB   | 2.373444                     | 2.87E-05               | 0.421329             | 0.980408               | 50     |
| RBBP4     | 2.394191                     | 0.002108               | 0.417678             | 0.642105               | 20     |
| WNT2      | 2.481328                     | 0.000697               | 0.40301              | 0.760234               | 19     |
| PRKAP1    | 2.489627                     | 0.003624               | 0.401667             | 0.79085                | 18     |
| GNAQ      | 2.556017                     | 0.009064               | 0.391234             | 0.33333                | 18     |
| GL12      | 2.456432                     | 0.002979               | 0.407095             | 0.698529               | 17     |
| RRAGD     | 2.697095                     | 0.000347               | 0.370769             | 0.95                   | 16     |
| MITF      | 2.448133                     | 0.010795               | 0.408475             | 0.549451               | 14     |
| NES       | 2.53112                      | 0.000352               | 0.395082             | 0.769231               | 13     |
| TLE4      | 2.385892                     | 0.000377               | 0.41913              | 0.709691               | 11     |
| RND3      | 2.53112                      | 0.00226                | 0.395082             | 0.490909               | 11     |
| PRL       | 2.585062                     | 0.001092               | 0.386838             | 0.472277               | 11     |
| IL2RB     | 2.742739                     | 0.000934               | 0.364599             | 0.644444               | 10     |
| F2RL2     | 2.73029                      | 0.003432               | 0.366261             | 0.527778               | 9      |
| TWIST2    | 2.809129                     | 0.000309               | 0.359582             | 0.527778               | 9      |
| TRIO      | 2.622407                     | 0.008446               | 0.381329             | 0.535714               | 8      |
| EPS15     | 2.705394                     | 0.002056               | 0.369632             | 0.642857               | 8      |

Betweenness centrality is a measure of information flow in a network system. Nodes with a high betweenness centrality are crucial for a network since they can control information flow in a biological network and can be considered as targets for drug discovery. It is basically defined as the number of shortest paths in a graph that pass through the node, divided by the total number of shortest paths. Among the top three genes with the highest degrees of centrality, CTNNB1 has a comparatively higher betweenness centrality value than the other two, i.e., SNRPF and HNRNPAB. Another important measure that estimates how fast the flow of information would be through a given node to other nodes is closeness centrality. Among the three top genes in Table 3, CTNNB1 (0.51828) has the highest value followed by SNRPF (0.46615) and HNRNPAB (0.42133), respectively. Average shortest-path length may be defined as the average number of steps along the shortest paths for all possible pairs of network nodes. It measures how efficiently information or mass transport occurs on a network. This list has also been topped by CTNNB1 (1.92946) followed by SNRPF (2.14523) and HNRNPAB (2.37344), respectively.

The colored nodes represent the first shell interactors or the query proteins (120 DEGs) while the white nodes represent the second shell interactors or the proteins that are not included in the input file and have been included for analytical purposes only. The maximum number of white nodes that was allowed in our PPI analyses was set to 50. In the inset, the 20 proteins that were listed in Table 3 are represented via protein–protein interactions without any secondary interactors.

### 3.3. Pathway Enrichment Analyses

In the GO functional enrichment analyses using the BINGO plugin of Cytoscape (Figure 6), the yellow nodes are significantly over-represented while the white nodes are not significantly over-represented and are included only to show the yellow nodes in the context of the GO hierarchy. The size of a node is proportional to the number of query genes that are annotated to the corresponding GO category. The top 20 GO categories based on their respective node sizes, which are significantly over-represented in our study, are listed in Table 4. Among these significantly over-represented categories, the highest node size was reported for the biological regulation pathway followed by regulation of biological processes and regulation of cellular processes. Neighborhood connectivity was found to be highest for regulation of the signaling pathway, followed by biological regulation, organ morphogenesis, and skeletal development. It is interesting to find that among the
first 20 over-represented pathways, CTNNB1 was found to be present in all the pathways. This shows the importance of this gene in the flow of information in reference to the pathophysiology of both the diseases. The HNRNPAB protein was found to be involved in 15 pathways, thereby demonstrating its role in disease occurrence. TWIST2 protein has also been found to be present in 12 pathways. These observations definitely point towards the probable involvement of the TWIST2 gene in endometriosis and RPL etiology. The SNRPF protein was found to only be linked to the cellular component organization pathway in spite of having the highest degree of centrality in the case of the protein–protein interaction network.

**Figure 6.** Enrichment network of the shared DEGs based on biological processes. Biological process network of differentially expressed genes of endometriosis and recurrent pregnancy loss patients using the BINGO app of Cytoscape. Large nodes indicate more genes involved and the size of a node is proportional to the number of targets in the GO category. Yellow nodes indicate the 120 genes playing a significant role in endometriosis and RPL promotion: \( p \)-value < 0.05.

**Table 4.** List of top 20 significantly overrepresented GO categories derived from BINGO analysis output, based on our data. The list has been arranged in ascending order of node size.

| Name | Description                       | Average Shortest Path Length | Betweenness Centrality | Closeness Centrality | Neighborhood Connectivity | Node Size | No. of Genes | Adjusted \( p \)-Value |
|------|-----------------------------------|------------------------------|------------------------|----------------------|---------------------------|-----------|--------------|---------------------|
| 65007 | biological regulation             | 3.72                         | 0.138077               | 0.268817             | 8.333333                  | 16.12452  | 65           | 0.00348             |
| 50789 | regulation of the biological process | 2.68254                    | 0.263324               | 0.372781             | 4.090909                  | 15.87451  | 63           | 0.0027              |
| 50794 | regulation of the cellular process | 2.605263                    | 0.131545               | 0.383838             | 4.2                       | 15.74802  | 62           | 0.0024              |
### Table 4. Cont.

| Name     | Description                                              | Average Shortest Path Length | Betweenness Centrality | Closeness Centrality | Neighborhood Connectivity | Node Size | No. of Genes | Adjusted p-Value |
|----------|-----------------------------------------------------------|------------------------------|------------------------|-----------------------|----------------------------|-----------|--------------|------------------|
| 31323    | regulation of the cellular metabolic process             | 0                            | 0                      | 0                     | 7                          | 12        | 36           | 0.0449           |
| 23052    | signaling                                                 | 2.625                        | 0.012372               | 0.380952              | 5.666667                   | 12        | 36           | 0.00789          |
| 32502    | developmental process                                     | 2.858824                     | 0.112328               | 0.349794              | 7                          | 11.6619   | 34           | 0.0216           |
| 7275     | multicellular organismal development                      | 3.904762                     | 0.175373               | 0.256098              | 5                          | 11.3137   | 32           | 0.0216           |
| 10468    | regulation of gene expression                            | 1.333333                     | 0.015726               | 0.75                  | 3                          | 11.1353   | 31           | 0.0299           |
| 48856    | anatomical structure development                          | 2.426471                     | 0.066427               | 0.412121              | 5.428571                   | 10.7733   | 29           | 0.0295           |
| 16043    | cellular component organization                           | 2.625                        | 0.017086               | 0.380952              | 5.2                        | 10.7733   | 29           | 0.0207           |
| 48731    | system development                                        | 3.531915                     | 0.30602                | 0.283133              | 5.125                      | 10.5831   | 28           | 0.0194           |
| 23033    | signaling pathway                                         | 1.5                          | 0.007868               | 0.666667              | 2.5                        | 10.3923   | 27           | 0.0113           |
| 48869    | cellular developmental process                            | 2.774194                     | 0.143077               | 0.360465              | 6.166667                   | 9.591663  | 23           | 0.0143           |
| 48523    | negative regulation of cellular process                  | 3                            | 0.074143               | 0.333333              | 5.25                       | 9.380832  | 22           | 0.0371           |
| 30154    | cell differentiation                                      | 2.615385                     | 0.069689               | 0.382353              | 4                          | 9.380832  | 22           | 0.0184           |
| 48513    | organ development                                         | 2.827586                     | 0.199668               | 0.353659              | 5.375                      | 9.165151  | 21           | 0.0482           |
| 7166     | cell surface receptor linked signaling pathway            | 1                            | 0.005863               | 1                     | 1.5                        | 8.944272  | 20           | 0.00875          |
| 51239    | regulation of the multicellular organismal process        | 1.9375                       | 0.058335               | 0.516129              | 4.285714                   | 8.485281  | 18           | 0.0083           |
| 35466    | regulation of signaling pathway                           | 0                            | 0                      | 0                     | 11                         | 7.745967  | 15           | 0.0299           |

### 4. Discussion

A large number of works have been carried out in the past decades to identify genetic markers for both endometriosis and RPL [25–33]. However, a trustworthy molecular marker having significant prognostic value has not yet been determined. Moreover, the lack of potential drug targets is also one of the probable causes for several unsuccessful
attempts to ameliorate the diseases. Therefore, there is an urgent need for the identification of potential biomarkers for the two diseases. This study is one of the pioneers in finding a common potential biomarker for the two diseases for successful diagnostic purposes and for effective drug delivery systems.

The literature survey provided epidemiological evidence to establish a probable link between endometriosis and RPL [34]. A recent investigation by Santulli et al. demonstrated an increased rate of spontaneous miscarriages in endometriosis-affected females [35]. Another interesting study in 2017 claimed mild endometriosis to be a potential risk for miscarriages [36]. Later in 2019, a study claimed that endometriosis affected the efficacy of assisted reproductive technology by increasing the risk of miscarriage [37].

More recently, Poli-Neto et al. identified the NOLC1 gene as the most common gene in the phase I and II endometriosis and affects menstruation, while in phases III and IV, the genes CDKN1B, DLD, ELOVL5, H2AFZ, ID11, ME1, MTHFD2, NOLC1, and SOD1 play a major role. These reports prompted the authors to explore the relationship between endometriosis and RPL through the identification of any potential biomarkers common to both the diseases. The present study, in relation to Poli-Neto et al. [38], extends the identification of CTNNB1, HNRNP A2B1, SNRPF, and TWIST2 genes as major markers, while the TWIST2 gene was identified as the most prominent marker for the exploration of endometriosis and RPL. Although, authors investigated and predicted several parameters reporting challenges in treating the diseases [38].

It is clearly evident from Figure 3 that both the diseases have similar gene expression patterns, thereby providing a clear indication for some common markers for the two diseases. It is also evident from Figure 3 that there exists a clear distinction between the patient and the control groups of each dataset in terms of the expression profiles of the genes. This observation partially supports the idea that the above-mentioned 207 genes may be considered as signatory markers for both EM and RPL. It is clearly evident from Table 2 that the TWIST2 gene has the highest fold change value from both the EA and NA analyses. This observation clearly indicates that TWIST2 has a significant role to play as a potential diagnostic marker for endometriosis-based recurrent miscarriages. TWIST2 has a very important role to play in reproduction. The TWIST2 gene is proved to play a very significant role in embryo implantation in mice. Embryo implantation is a very important event for a successful pregnancy. Suppression of the TWIST2 gene impaired the embryo implantation by suppressing endothelial-mesenchymal transition (EMT) during embryo implantation [39]. A recent clinical study reported Setlleis syndrome in a child with a novel mutation in the TWIST2 gene [40]. Another study in 2014 by Huang et al. showed that haploinsufficiency of TWIST2 results in reduced bone formation [41]. Franco et al. highlighted TWIST2 as a molecular switch during gene transcription [42]. Furthermore, sequestration of E-proteins by increased TWIST2 levels functions to inhibit muscle-specific gene activation [42–44]. TWIST2 requires Histone Deacetylases for Myoblast Determination Protein 1-Myocyte Enhancer Factor 2 inhibition [43]. TWIST2 is also known to regulate osteoblast differentiation, however its involvement occurs temporally after TWIST1 [41,45]. The transcription factor RUNX2 is considered a master regulator of the osteogenic program due to its indispensable role in the regulation of most of the genes that give rise to the mature osteoblast phenotype [41,46]. Both TWIST1 and TWIST2 can also regulate RUNX2 at the protein level by physically interacting with RUNX2 and inhibiting its ability to bind DNA [41,46]. TWIST2 also acts as an important key negative regulator of myeloid lineage development, as manifested by marked increases in mature myeloid populations of macrophages, neutrophils, and basophils in TWIST2-deficient mice [41,47]. Therefore, on converging our findings with the above-mentioned published investigations, it is clearly evident that downregulation of the TWIST2 gene may have a very potent role in early embryonic developmental events, rendering it as a potential clinical marker for endometriosis based RPL. Another gene, CA XII (Carbonic Anhydrase XII), also has a high log fold change value, as evident from Table 2. CA XII has been found to have prominent expression during mouse embryonic development [48]. However, in this article, we did
not focus on other genes in Table 2 since, on overlapping our intersected gene list from EA and NA output with the Geo2R results, only the TWIST2 gene was found to be in common. In other words, only the TWIST2 gene was found to be present in all the three analyses and therefore was considered to be an important clinical marker.

In the case of protein–protein interaction analyses, the top three genes participating in the network were SNRPF, CTNNB1 and HNRNPAB, based on their degree of centrality. Small Nuclear Ribonucleoprotein Polypeptide F (SNRPF) plays role in pre-mRNA splicing and also as a component of the spliceosomal U1, U2, U4 and U5 small nuclear ribonucleoproteins (snRNPs), the building blocks of the spliceosome [49–57]. The SNRPF gene was found to be downregulated in our study samples and therefore may serve as a valid target for disease-based research.

CTNNB1 or Catenin Beta 1 is an important downstream component of the canonical Wnt signaling pathway [58–65]. The Wnt signaling pathway is known for its role in embryonic development, where it actively participates in body axis patterning, cell fate specification, cell proliferation, and cell migration events [66]. These developmental processes are essential for proper tissue formations, including bone, heart, and muscles. CTNNB1 protein is also a part of a protein complex that forms cell–cell junctions in epithelial and endothelial tissues [67]. Additionally, β-catenin 1 also promotes neurogenesis by maintaining sympathetic neuroblasts within the cell cycle [68]. Surprisingly, β-catenin has also been associated with endometrial cancer onset and recurrence [69]. Therefore, it is evident from the above-mentioned studies that CTNNB1 has an important role in early developmental pathways and inter- and intracellular recognitions. Interestingly, this gene was found to be upregulated in our study, rendering it an important marker for disease-based research and for exploring its role in disease prognosis.

Located on chromosome 5q35.3, HNRNPAB or Heterogeneous Nuclear Ribonucleoprotein A/Bis is a member of a subfamily of ubiquitously expressed heterogeneous nuclear ribonucleoproteins (hnRNPs). They are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism and transport. HNRNPAB has also been found to be associated with ankyloblepharon-ectodermal defects–cleft lip/palate syndrome [70–72]. Surprisingly, this gene is also a member of the preimplantation embryo pathway (WP3527) [73]. In our study, this gene is downregulated, similar to the SNRPF gene. Considering the above-mentioned facts, it can be hypothesized that HNRNPAB has a definitive role in disease prognosis via pre-mRNA processing or preimplantation embryo pathways and can be an essential diagnostic marker for endometriosis-based RPL.

It is clearly evident from Table 4 that the top 20 pathways of the pathway enrichment analysis based on the overlapping genes of the EA–NA analyses are mainly concerned with signaling pathways and developmental biology, thereby indicating the combined inclination of the genes towards functioning in the arena of developmental signaling events during embryogenesis. When we tried to explore the involvement of our potential biomarkers in the biological pathways, it was seen that the SNRPF protein is involved in the cellular component organization pathway. Interestingly, CTNNB1 is involved in all 20 pathways. HNRNPAB is involved in 15 pathways and TWIST2 in 13 of the pathways. CTNNB1, HNRNPAB and TWIST2 are commonly involved in 11 out of 20 major pathways, that are shown in Supplementary Table S1, while SNRPF and CTNNB1 share only one pathway in common.

5. Conclusions

In conclusion, our work has identified 120 DEGs in the five profile datasets based on ExAtlas and Network Analyst results. A handful of biomarkers were found common to both endometriosis and RPL, and can have a diagnostic role in the case of endometriosis-based RPL. Notable among these markers are CTNNB1, HNRNPAB, SNRPF and TWIST2. The 120 DEGs, when compared with the cumulative output of Geo2R software, showed
only one gene (TWIST2) to be common among the three analytical approaches. Therefore, our study also claims the TWIST2 gene as a prominent marker of choice for the diseases. The significantly enriched pathways based on the above-mentioned genes are mainly centered on signaling and developmental events. These findings could significantly improve our understanding of the cause and underlying molecular events in endometriosis-based recurrent miscarriages. However, further downstream validation of these markers is a needed for quantitating their potentiality and establishing their efficacy as a potential drug target(s).

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10.3390/app11083349/s1, Table S1: List of top 20 pathways as an outcome of pathway enrichment analysis using the target genes by BINGO app of Cytoscape.

**Author Contributions:** Conceptualization, S.R.; methodology, software, data curation, analysis P.G., S.R., S.S., Q.M.S.J. and K.K.K.; writing—original draft preparation, P.G., S.R., S.S.; writing—review and editing, S.R., K.K.K., J.C.K., A.K., Q.M.S.J., N.K.J., D.K. and J.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data used to arrive at the findings of the study are openly available in GEO datasets at https://www.ncbi.nlm.nih.gov/gds [11].

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**
1. Farquhar, C. Endometriosis. *BMJ* 2007, 334, 249–253. [CrossRef]
2. Klemmt, P.A.B.; Starzinski-Powitz, A. Molecular and Cellular Pathogenesis of Endometriosis. *Curr. Womens Health Rev.* 2018, 14, 106–116. [CrossRef]
3. Laganà, A.S.; Garzon, S.; Götte, M.; Viganò, P.; Franchi, M.; Ghezzi, F.; Martin, D.C. The Pathogenesis of Endometriosis: Molecular and Cell Biology Insights. *Int. J. Mol. Sci.* 2019, 20, 5615. [CrossRef]
4. Pazhohan, A.; Amidi, F.; Akbari-Asbagh, F.; Seyedrezazadeh, E.; Farzadi, L.; Khodarahmin, M.; Mehidinejadiani, S.; Sebbani, A. The Wnt/β-catenin signaling in endometriosis, the expression of total and active forms of β-catenin, total and inactive forms of glycogen synthase kinase-3β, WNT7a and DICKKOPF-1. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 2018, 220, 1–5. [CrossRef]
5. RPL (Recurrent Pregnancy Loss): Guideline of the European Society of Human Reproduction and Embryology. ESHRE Early Pregnancy Guideline Development Group. 2017, pp. 1–153. Available online: https://www.eshre.eu/Guidelines-and-Legal/Guidelines/Recurrent-pregnancy-loss.aspx (accessed on 18 July 2020).
6. Imanaka, S.; Maruyama, S.; Kimura, M.; Nagayasu, M.; Kobayashi, H. Towards an understanding of the molecular mechanisms of endometriosis-associated symptoms (Review). *World Acad. Sci. J.* 2020, 2, 12. [CrossRef]
7. Ticoni, C.; Pietropolli, A.; Di Simone, N.; Piccione, E.; Fazleabas, A. Endometrial Immune Dysfunction in Recurrent Pregnancy Loss. *Int. J. Mol. Sci.* 2019, 20, 5332. [CrossRef] [PubMed]
8. Sapkota, Y.; iPSYCH-SSI-Broad Group; Steinthorsdottir, V.; Morris, A.P.; Fassbender, A.; Rahmioglu, N.; De Vivo, I.; Buring, J.E.; Zhang, F.; Edwards, T.L.; et al. Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. *Nat. Commun.* 2017, 8, 15539. [CrossRef]
9. Akter, S.; Xu, D.; Nagel, S.C.; Bromfield, J.J.; Pelch, K.; Wilshire, G.B.; Joshi, T. Machine Learning Classifiers for Endometriosis Using Transcriptomics and Metabolomics Data. *Front. Genet.* 2019, 10, 766. [CrossRef]
10. Mear, L.; Herr, M.; Fauconnier, A.; Pineau, C.; Vialard, F. Polymorphisms and endometriosis: A systematic review and meta-analyses. *Hum. Reprod. Update* 2020, 26, 73–102. [CrossRef]
11. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002, 30, 207–210. [CrossRef] [PubMed]
12. Sharov, A.A.; Schlessinger, D.; Ko, M.S. ExAtlas: An interactive online tool for meta-analysis of gene expression data. *J. Bioinform. Comput. Biol.* 2015, 13, 1550019. [CrossRef]
13. Zhou, G.; Soufan, O.; Ewald, J.; Hancock, R.E.W.; Basu, N.; Xia, J. NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* 2019, 47, W234–W241. [CrossRef]
14. Limma Powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. Available online: https://pubmed.ncbi.nlm.nih.gov/25605792/ (accessed on 23 March 2021).
44. Spicer, D.B.; Rhee, J.; Cheung, W.L.; Lassar, A.B. Inhibition of myogenic bHLH and MEF2 transcription factors by the bHLH protein Twist. *Science* **1996**, *272*, 1476–1480. [CrossRef]

45. Lee, M.S.; Lowe, G.; Flanagan, S.; Kuchler, K.; Glackin, C.A. Human Dermo-1 has attributes similar to twist in early bone development. *Bone* **2000**, *27*, 391–402. [CrossRef]

46. Bialek, P.; Kern, B.; Yang, X.; Schrock, M.; Sosic, D.; Hong, N.; Wu, H.; Yu, K.; Ornitz, D.M.; Olson, E.N.; et al. A twist code determines the onset of osteoblast differentiation. *Dev. Cell* **2004**, *6*, 423–435. [CrossRef]

47. Sharabi, A.B.; Aldrich, M.; Sosic, D.; Olson, E.N.; Friedman, A.D.; Lee, S.H.; Chen, S.Y. Twist-2 controls myeloid lineage development and function. *PLoS Biol.* **2008**, *6*, e316. [CrossRef] [PubMed]

48. Kallio, H.; Pastorekova, S.; Pastorek, J.; Waheed, A.; Sly, W.S.; Mannisto, S.; Heikinheimo, M.; Parkkila, S. Expression of carbonic anhydrases IX and XII during mouse embryonic development. *BMC Dev. Biol.* **2006**, *6*, 22. [CrossRef] [PubMed]

49. Agafonov, D.E.; Kastner, B.; Dybkov, O.; Hofele, R.V.; Liu, W.T.; Urlaub, H.; Luhrmann, R.; Stark, H. Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* **2016**, *351*, 1416–1420. [CrossRef] [PubMed]

50. Chari, A.; Golas, M.M.; Klingenheger, M.; Neuenkirchen, N.; Sander, B.; Englbrecht, C.; Sickmann, A.; Stark, H.; Fischer, U. An assembly chaperone collaborates with the SMN complex to generate spliceosomal SnRNPs. *Cell* **2008**, *135*, 497–509. [CrossRef] [PubMed]

51. Grimm, C.; Chari, A.; Pelz, J.P.; Kuper, J.; Kisser, C.; Diederichs, K.; Stark, H.; Schindelin, H.; Fischer, U. Structural basis of assembly chaperone-mediated snRNP formation. *Mol. Cell* **2013**, *19*, 692–703. [CrossRef] [PubMed]

52. Jurica, M.S.; Licklider, L.J.; Gygi, S.R.; Grigorieff, N.; Moore, M.J. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* **2002**, *8*, 426–439. [CrossRef]

53. Kondo, Y.; Oubridge, C.; van Roon, A.M.; Nagai, K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5′ splice site recognition. *Elife* **2015**, *4*, e04986. [CrossRef] [PubMed]

54. Pomeranz Krummel, D.A.; Oubridge, C.; Leung, A.K.; Li, J.; Nagai, K. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* **2009**, *458*, 475–480. [CrossRef]

55. Zhang, X.; Yan, C.; Hang, J.; Finci, L.I.; Lei, J.; Shi, Y. An Atomic Structure of the Human Spliceosome. *Cell* **2017**, *169*, 918–929.e14. [CrossRef] [PubMed]

56. Bertram, K.; Agafonov, D.E.; Dybkov, O.; Haselbach, D.; Leelaram, M.N.; Will, C.L.; Urlaub, H.; Kastner, B.; Luhrmann, R.; Stark, H. Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. *Cell* **2017**, *170*, 701–713.e11. [CrossRef] [PubMed]

57. Bertram, K.; Agafonov, D.E.; Liu, W.T.; Dybkov, O.; Will, C.L.; Hartmuth, K.; Urlaub, H.; Kastner, B.; Stark, H.; Luhrmann, R. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **2017**, *542*, 318–323. [CrossRef]

58. Lillehoj, E.P.; Lu, W.; Kiser, T.; Goldblum, S.E.; Kim, K.C. MUC1 inhibits cell proliferation by a beta-catenin-dependent mechanism. *Biochim. Biophys. Acta* **2007**, *1773*, 1028–1038. [CrossRef]

59. Weiske, J.; Albring, K.F.; Huber, O. The tumor suppressor Fhit acts as a repressor of beta-catenin transcriptional activity. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20344–20349. [CrossRef]

60. Bahmanyar, S.; Kaplan, D.D.; Deluca, J.G.; Giddings, T.H., Jr.; O'Toole, E.T.; Winey, M.; Sloan, E.D.; Casey, P.J.; Nelson, W.J.; Barth, A.I. beta-Catenin is a Nek2 substrate involved in centrosome separation. *Genes Dev.* **2008**, *22*, 91–105. [CrossRef] [PubMed]

61. Li, H.; Ray, G.; Yoo, B.H.; Erdogan, M.; Rosen, K.V. Down-regulation of death-associated protein kinase-2 is required for beta-catenin-induced anoikis resistance of malignant epithelial cells. *Genes Dev.* **2008**, *22*, 2206–2215. [CrossRef] [PubMed]

62. Fiset, A.; Xu, E.; Bergeron, S.; Marette, A.; Pelletier, G.; Siminovitch, K.A.; Olivier, M.; Beauchemin, N.; Faure, R.L. Compartmentalized CDK2 is connected with SHP-1 and beta-catenin and regulates insulin internalization. *Cell. Signal.* **2011**, *23*, 911–919. [CrossRef]

63. Satow, R.; Shitashige, M.; Jigami, T.; Fukami, K.; Honda, K.; Kitabayashi, I.; Yamada, T. beta-catenin inhibits promyelocytic leukemia protein tumor suppressor function in colorectal cancer cells. *Gastroenterology* **2012**, *142*, 572–581. [CrossRef]

64. Genovese, G.; Ghosh, P.; Li, H.; Rettino, A.; Sioletic, S.; Cittadini, A.; Sgambato, A. The tumor suppressor HINT1 regulates MITF and beta-catenin transcriptional activity in melanoma cells. *Cell Cycle* **2012**, *11*, 2206–2215. [CrossRef] [PubMed]

65. Yu, Y.; Wu, J.; Wang, Y.; Zhao, T.; Ma, B.; Liu, Y.; Fang, W.; Zhu, W.G.; Zhang, H. Kindlin 2 forms a transcriptional complex with beta-catenin and TCF4 to enhance Wnt signalling. *EMBO Rep.* **2012**, *13*, 750–758. [CrossRef]

66. Bellows, T.S.; Fisher, T.W. *Handbook of Biological Control: Principles and Applications of Biological Control*; Academic Press: San Diego, CA, USA, 1999; p. xxiii.

67. Brembeck, F.H.; Rosario, M.; Birchmeier, W. Balancing cell adhesion and Wnt signaling, the key role of beta-catenin. *Curr. Opin. Genet. Dev.* **2006**, *16*, 51–59. [CrossRef]

68. Joksimovic, M.; Patel, M.; Taketo, M.M.; Johnson, R.; Awatramani, R. Ectopic Wnt/beta-catenin signaling induces neurogenesis in the spinal cord and hindbrain floor plate. *PLoS ONE* **2012**, *7*, e30266. [CrossRef]

69. Kurnit, K.C.; Kim, G.N.; Fellman, B.M.; Urbauer, D.L.; Mills, G.B.; Zhang, W.; Broadus, R.R. CTNNB1 (beta-catenin) mutant mutation identifies low grade, early stage endometrial cancer patients at increased risk of recurrence. *Mod. Pathol.* **2017**, *30*, 1032–1041. [CrossRef]

70. Yoh, K.; Prywes, R. Pathway Regulation of p63, a Director of Epithelial Cell Fate. *Front. Endocrinol.* **2015**, *6*, 51. [CrossRef] [PubMed]

71. Fete, M.; vanBokhoven, H.; Clements, S.E.; McKeon, F.; Roop, D.R.; Koster, M.I.; Missero, C.; Attardi, L.D.; Lombillo, V.A.; Ratovitski, E.; et al. International Research Symposium on Ankyloblepharon-Ectodermal Defects-Cleft Lip/Palate (AEC) syndrome. *Am. J. Med. Genet. A* **2009**, *149A*, 1885–1893. [CrossRef]
72. Fomenkov, A.; Huang, Y.P.; Topaloglu, O.; Brechman, A.; Osada, M.; Fomenkova, T.; Yuriditsky, E.; Trink, B.; Sidransky, D.; Ratovitski, E. P63 alpha mutations lead to aberrant splicing of keratinocyte growth factor receptor in the Hay-Wells syndrome. *J. Biol. Chem.* 2003, 278, 23906–23914. [CrossRef] [PubMed]

73. Yan, L.; Yang, M.; Guo, H.; Yang, L.; Wu, J.; Li, R.; Liu, P.; Lian, Y.; Zheng, X.; Yan, J.; et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 2013, 20, 1131–1139. [CrossRef] [PubMed]