Expression of paralogs of cytochrome P_{450}21a1 pseudogene (Cyp21a1-ps) and endogenous retrovirus SC1 (SC1) in the rat ovary during the ovulatory process

Lawrence L Espey, Rebecca A Garcia, Haruhiro Kondo¹, Bunpei Ishizuka¹, Shinya Yoshioka², Shingo Fujii², Stephen Hampton and JoAnne S Richards³

Department of Biology, One Trinity Place, Trinity University, San Antonio, Texas 78212, USA
¹Department of Obstetrics and Gynecology, St Marianna University School of Medicine, Kawasaki 216-8511, Japan
²Department of Obstetrics and Gynecology, Kyoto University School of Medicine, Kyoto 606- 8507, Japan
³Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA
(Correspondence should be addressed to L L Espey; Email: lespey@trinity.edu)

Abstract

This study assesses the relatively high incidence of the expression of paralogs of several pseudogenes within the cascade of expression of functional genes in the rat ovary in response to an ovulation-stimulating dose of gonadotropin. Immature Wistar rats were primed with 10 IU equine chorionic gonadotropin subcutaneously, and 48 h later the 12-h ovulatory process was initiated by 10 IU hCG subcutaneously. Ovarian RNA was extracted at 0, 2, 4, 8, 12, and 24 h after injecting the animals with hCG. The RNA extracts were used for RT-PCR differential display to detect gene expression in the ovarian tissue. Sequence analyses of differentially expressed cDNAs revealed that ~27% (i.e. 22/82 clones) of the transcripts were fragments of paralogs of known pseudogenes. Out of the 22 clones reported here, 12 have high sequence similarity to the cytochrome P450 pseudogene Cyp21a1-ps, and 5 have high sequence similarity to both the Cyp21a1-ps and the aldo-keto reductase gene Akr1c6. The remaining five clones were paralogs of the endogenous retrovirus SC1 that has heavily infested the rat genome. Northern analyses reveal that peak expression of all the 22 paralogs occurs at 4–8 h into the ovulatory process. In situ hybridization shows that expression of these pseudogenes is primarily in the granulosa layer of ovulatory follicles. In summary, the results reveal that ovarian expression of Cyp21a1-ps- and SC1-like pseudogenes occurs concurrently with the ovulatory process.


Introduction

Ovulation is a distinct biological process that is initiated when a surge in luteinizing hormone (LH) acts on mature ovarian follicles. This hormonally induced process has been studied by various biophysical and biochemical procedures for more than a century (Espey & Lipner 1994). In recent years, the work on ovulation has focused on LH-induced gene expression and on the manner in which the protein products of these genes contribute to the decomposition and rupture of ovarian follicles (Espey et al. 2000a, 2004, Robker et al. 2000, Richards et al. 2002, Jo & Curry 2004, Espey 2006, Espey & Richards 2006, Miyakoshi et al. 2006). Initially, it was assumed that the ovulatory process might involve expression of only a few genes for proteolytic degradation of follicular connective tissues. However, recent analyses of ovulation-specific gene expression by a number of investigators have made it evident that the ovulatory surge in LH induces an extensive assortment of gene expression in the mammalian ovary (Espey & Richards 2002, 2006, Jo et al. 2004, Rae et al. 2004, McNatty et al. 2005, Hernandez-Gonzalez et al. 2006, Hourvitz et al. 2006, Richards 2007). It is now clear this ‘cascade’ of gene expression contributes to a transition in ovarian steroidogenesis, the induction of an acute inflammatory reaction, the suppression of inflammation-induced oxidative stress, and other less distinct metabolic events in the ovary at the time of ovulation.

The present study focuses on multiple paralogs of several ovarian pseudogenes that are expressed during ovulation (by definition, ‘paralogs’ are homologous sequences that underwent evolutionary separation from a common ancestor following a gene duplication event (Fitch 2000)). In particular, this study examines a number of paralogs of cytochrome P_{450}21a1 pseudogenes (Cyp21a1-ps) that are transcribed by multiple chromosomes during the ovulatory process. In addition, this study assesses the ovulation-related expression of several pseudogenes that are similar to segments of Cyp21a1-ps, but have distinct sequence similarity to a
functional aldo-keto reductase (Akr1c6). Also, this study presents data on the ovulation-related expression of multiple forms of the endogenous retrovirus (ERV) SC1 in the rat ovary. Collectively, these genes represent several varieties of pseudogenes.

The ovulation-related expression of Cyp21a1-ps and SC1 paralogs was detected during RT-PCR differential display (DD) experiments conducted during the past 12 years. Routinely, the DD procedure detects only random gene expression in an experimental tissue. Therefore, the pseudogenes reported here are not based on any predetermined interest or bias. Out of a total of 82 different ovulation-related genes that were discovered, 22 of the sequences were identified as pseudogenes of the above types. None of the remaining 60 ovulation-related genes were of the pseudogene category. Thus, the unique expression of ovarian Cyp21a1-ps and SC1 merits closer examination because their transcription is a common characteristic of ovulation in the rat.

Materials and Methods

Animal tissue and animal injections

Immature Wistar rats were induced to superovulate by injections of equine chorionic gonadotropin (eCG) and hCG as described previously (Espey et al. 2000b). Ovarian RNA was extracted at the periovulatory intervals of 0, 2, 4, 8, 12, and 24 h after hCG. These extracts of nucleic acid were used for DD, and differentially expressed cDNA bands were used for northern analyses, in situ hybridization, and sequence analyses. The ovulation rate in the experimental animals was determined by a procedure that also has been described previously (Espey et al. 2000b). The rats were killed by exposure to CO2 in order to extract ovarian RNA and to determine ovulation rate. The number of ova in the oviducts of ten of the rats at 24 h after hCG was 70±6 ±3.0. This work was conducted in accordance with the accepted standards of humane animal care, and the experimental procedures conformed with The UFAW Handbook on the Care and Management of Laboratory Animals and the approval of the institutional committee on animal care, i.e. the Animal Research Committee of Trinity University.

RNA extraction

Total RNA was extracted from whole ovaries that were extirpated at 0, 2, 4, 8, 12, and 24 h after the ovulatory process had been initiated by a dose of 10 IU of hCG. The 0-h control group did not receive any hCG. At each of the six designated intervals, the ovaries from seven to ten rats were immediately frozen on blocks of dry ice and then pooled together to provide a total of ~0.5 g ovarian tissue for RNA extraction. The RNA was extracted by a standard guanidine isothiocyanate/cesium chloride procedure that has been described previously (Espey et al. 2000b). Different RNA extractions for DD RT-PCR and northern blotting were performed on numerous occasions over the course of this 12-year study of ovarian gene expression. The 12 Cyp21a1-ps clones were obtained from ten different RNA extractions, while the five Akr1c6 and the five SC1 clones were obtained from ten other RNA extractions.

DD protocols that led to detection of pseudogenes

The steps of the DD procedure were carried out as described previously (Espey et al. 2000b). In brief, RNA was extracted by a standard guanidine isothiocyanate/cesium chloride procedure. RT-PCR was performed by primers from a number of different RNAimage kits (GenHunter Corporation, Nashville, TN, USA). The specific primer sets that yielded differentially expressed homologs of clones for Cyp21a1-ps, Akr1c6, and SC1 were quite diverse (Table 1).

After extraction and reamplification of the differentially expressed cDNAs, standard northern analyses were performed on numerous occasions over the course of this study and the unique expression in an experimental tissue. Therefore, the DD experiments conducted during the past 12 years. Consequently, the unique cDNA fragments were individually cloned using a PCR-TRAP Cloning System (P404, Nashville, TN, USA). The specific primer sets that yielded differentially expressed homologs of clones for Cyp21a1-ps, Akr1c6, and SC1 were quite diverse (Table 1).

After extraction and reamplification of the differentially expressed cDNAs, standard northern analyses were performed on numerous occasions over the course of this study and the unique expression in an experimental tissue. Therefore, the DD experiments conducted during the past 12 years. Consequently, the unique cDNA fragments were individually cloned using a PCR-TRAP Cloning System (P404, Nashville, TN, USA). The specific primer sets that yielded differentially expressed homologs of clones for Cyp21a1-ps, Akr1c6, and SC1 were quite diverse (Table 1).

Table 1 PCR primers that yielded differentially expressed cDNAs

<table>
<thead>
<tr>
<th>Clone</th>
<th>Random primer</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyp21a1-1</td>
<td>5'-HTGTTAC-3'</td>
<td>DQ255907</td>
</tr>
<tr>
<td>Cyp21a1-2</td>
<td>5'-HTGTTGAC-3'</td>
<td>DQ266365</td>
</tr>
<tr>
<td>Cyp21a1-3</td>
<td>5'-HAGCATG-3'</td>
<td>EF189909</td>
</tr>
<tr>
<td>Cyp21a1-4</td>
<td>5'-HCAAGAC-3'</td>
<td>EF189908</td>
</tr>
<tr>
<td>Cyp21a1-5</td>
<td>5'-HCAAGAC-3'</td>
<td>DQ255932</td>
</tr>
<tr>
<td>Cyp21a1-6</td>
<td>5'-HCAAGAC-3'</td>
<td>DQ256370</td>
</tr>
<tr>
<td>Cyp21a1-7</td>
<td>5'-HAGATGC-3'</td>
<td>EF189906</td>
</tr>
<tr>
<td>Cyp21a1-8</td>
<td>5'-HCAACATC-3'</td>
<td>EF189910</td>
</tr>
<tr>
<td>Cyp21a1-9</td>
<td>5'-HGCAGC-3'</td>
<td>EF189911</td>
</tr>
<tr>
<td>Cyp21a1-10</td>
<td>5'-HTTCTAG-3'</td>
<td>DQ255906</td>
</tr>
<tr>
<td>Cyp21a1-11</td>
<td>5'-HTTCTAG-3'</td>
<td>DQ255910</td>
</tr>
<tr>
<td>Cyp21a1-12</td>
<td>5'-HTTCTAG-3'</td>
<td>EF189907</td>
</tr>
<tr>
<td>Cyp21a1-13</td>
<td>5'-HAAACCTG-3'</td>
<td>DQ255911</td>
</tr>
<tr>
<td>Cyp21a1-14</td>
<td>5'-HAGATGC-3'</td>
<td>EF189912</td>
</tr>
<tr>
<td>Cyp21a1-15</td>
<td>5'-HCAACATC-3'</td>
<td>AF159099.1</td>
</tr>
<tr>
<td>Cyp21a1-16</td>
<td>5'-HGCCTATG-3'</td>
<td>DQ255905</td>
</tr>
</tbody>
</table>

Akr1c6

<table>
<thead>
<tr>
<th>Clone</th>
<th>Random primer</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akr1c6-1</td>
<td>5'-HTGTTAC-3'</td>
<td>EF189914</td>
</tr>
<tr>
<td>Akr1c6-2</td>
<td>5'-HTGTTGAC-3'</td>
<td>EF189915</td>
</tr>
<tr>
<td>Akr1c6-3</td>
<td>5'-HCAAGAC-3'</td>
<td>DQ255932</td>
</tr>
<tr>
<td>Akr1c6-4</td>
<td>5'-HCAAGAC-3'</td>
<td>EF189913</td>
</tr>
<tr>
<td>Akr1c6-5</td>
<td>5'-HTGTTAC-3'</td>
<td>EF189916</td>
</tr>
</tbody>
</table>

SC1

<table>
<thead>
<tr>
<th>Clone</th>
<th>Random primer</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC1-1</td>
<td>5'-HTGTTAC-3'</td>
<td>EF189914</td>
</tr>
<tr>
<td>SC1-2</td>
<td>5'-HTGTTGAC-3'</td>
<td>EF189915</td>
</tr>
<tr>
<td>SC1-3</td>
<td>5'-HCAAGAC-3'</td>
<td>DQ255932</td>
</tr>
<tr>
<td>SC1-4</td>
<td>5'-HCAAGAC-3'</td>
<td>EF189913</td>
</tr>
<tr>
<td>SC1-5</td>
<td>5'-HTGTTAC-3'</td>
<td>EF189916</td>
</tr>
</tbody>
</table>

H, a HindIII restriction sequence (AAGCTT). Tg nine thymines.
Each sequence has been entered into the NCBI database and the accession numbers are available (Table 1). *In situ* hybridization was carried out as described previously (Espey et al. 2000b).

**Nucleotide sequences**

The nucleotide sequences reported in this paper have been submitted to GenBank under the accession numbers AF159099.1, DQ255905–DQ255907, DQ255909–DQ255911, DQ266363, DQ266365, DQ266370, DQ266372, and EF189906–EF189916.

**Statistical analysis**

The density of the signals from the northern blots was analyzed by the NIH Image program [http://rsb.info.nih.gov/nih-image/](http://rsb.info.nih.gov) using the download ‘NIH-image 162 Fat.hqx.’ The raw densitometric numbers at each of the six time points during the ovulatory process were used to tabulate the means± S.E.M. at each of the six time points during the ovulatory process. The 0-h value, i.e., the pre-hCG value, was designated as the control group, while the 2-, 4-, 8-, 12-, and 24-h values were the experimental groups. The significance of the differences among the means was determined by Duncan’s multiple range tests after a completely randomized one-way ANOVA of the means of the groups. The probability value used as the cutoff between ‘significant’ and ‘not significant’ was $P=0.05$. After completion of the statistical analysis, the mean of the time point that exhibited the highest signal density was arbitrarily set as 100%, and the remaining means and S.E.M. were expressed as proportions of 100%.

**Results**

**Northern analyses of mRNA expression during the periovulatory period**

The northern analyses for cDNAs that had sequences similar to *Cyp21a1-ps*, *Akr1c6*, or *SC1* showed comparable patterns of expression during the six different stages of the ovulatory process – with maximum expression occurring at 4–8 h after the animals were stimulated with an ovulatory dose of hCG (Figs 1–3). In addition, all three sets of clones hybridized with variable lengths of RNA on the northern blots, indicating the presence of multiple forms of the extracted mRNAs that generated the differentially displayed cDNAs.

**Localizations of Cyp21a1-ps, Akr1c6-ps, and SC1 mRNA expression by in situ hybridization**

*In situ* hybridization of ovarian tissues taken at 0, 8, and 24 h after hCG yielded signals that spatially localized the ovarian expression of mRNA for *Cyp21a1-ps*, *Akr1c6-ps*, and *SC1* paralogs in a temporal pattern that corresponded with the DD autoradiographs and the northern analyses. Probes prepared from clones of the gene fragments of all three classes of pseudogenes consistently showed minimal signal from the *in situ* preparations of control tissue staged at 0 h, i.e. of ovaries from animals that had not been injected with hCG (Figs 4–6). By contrast, there were substantial signals radiating from the tissue preparations at 8 h after hCG, which was the approximate time when there was maximum signal from the DDs and northern blots. Most of the expression of pseudogene mRNA was localized in the stratum granulosum of the mature follicles, but there was also diffuse signal from the interstitial tissues of the ovaries. Later, as the ovarian follicles began to transform into luteal tissue at 24 h after hCG, the *in situ* evidence showed that mRNA expression for all of the pseudogenes declined.

**General characteristics of the cDNA fragments for Cyp21a1-ps-like, Akr1c6-like, and SC1 clones**

Based on BLASTN searches using the NCBI database, 12 clones were initially identified as having the highest sequence similarity to *Cyp21a1-ps* (Fig. 7), and 5 clones were initially identified as having the highest similarity to the *Akr1c6* gene...
Subsequently, in more recent years, it has become evident from a closer analysis of the NCBI database that all five of the Akr1c6-like clones also possess 99–100% similarity to segments of the Cyp21a1-ps reference gene (NCBI accession number NG_004071). For this reason, sequences of the Akr1c6-like clones were also compared with the Cyp21a1-ps reference gene (Fig. 8). However, the sequences of the five Akr1c6-like clones were most similar to the Cyp21a1-ps reference gene at regions upstream and downstream from the locus for the 12 principal Cyp21a1-ps-like clones. Therefore, the Akr1c6-like clones were grouped separately from the Cyp21a1-ps-like clones (Fig. 8). By contrast, the five clones for SC1 were considered as a completely separate group because they had no sequence similarity to either the Cyp21a1-ps-like clones or the Akr1c6-like clones (Fig. 9).

Loci diversity and sequence similarity of the polymorphic clones of Cyp21a1-ps, Akr1c6, and SC1

The NCBI database was used to identify the various chromosomal loci that shared the greatest similarity to each of the clones in this study. In each instance where a locus was established, the database indicated there was 99–100% similarity between the given clone and the genomic sequence at the locus identified (Table 2). For the Cyp21a1-ps-like fragments, clones 1, 2, and 13 had the greatest similarity to sequence in the rat genome at locus 1q22. Clones 3–6 were highly similar to sequence at locus 3q41. Clones 7–10 were most similar to sequence at locus 11p12. Clones 11 and 14–16 were very similar to sequence at locus 14p22. Clones 12 and 17 were essentially identical to sequences in the rat genome at loci Xq12 and Xq22 respectively (Table 2). By contrast, the five paralogs of SC1 were most similar to rat genomic sequences found at 1q21, 5q33, 13q21, 19q12, and Xq22 respectively (Table 2). Also, in this effort to identify the corresponding position of each clone within the rat genome, the specific genes that flanked each paralog locus at its 5' and 3' ends were tabulated (Table 2). Since many of the loci were on different chromosomes, the flanking genes varied substantially from paralog to paralog.

Discussion

In assessing the results of this study, it may be useful to keep in mind that the mammalian genome is endowed with an abundance of pseudogenes (Hurst 2002, Torrents et al. 2003). For example, based on a BLASTN search of the NCBI database, there are at least 8986 loci randomly distributed throughout the rat genome that have significant sequence similarity to the polymorphic forms of Cyp21a1-ps reported.
in the present study. Similarly, there are at least 2825 different loci that have significant sequence similarity to SC1.

Pseudogenes are categorized into two fundamentally different groups, based on the nature of their origin. The ‘duplicated’ pseudogenes arise from tandem duplication of segments of chromosomes or from unequal crossing over, and they usually exhibit retention of the exon–intron pattern of their paralogous parental genes (Mighell et al. 2000, Zhang & Gerstein 2004, Svensson et al. 2006). By contrast, the ‘retrotransposed’ pseudogenes arise when freshly translated mRNAs undergo reverse transcription into cDNAs, which are subsequently reintegrated into the genome at new loci (Mighell et al. 2000, Zhang & Gerstein 2004, Svensson et al. 2006).

Cytochrome P450 genes make up one of the largest members of a multigene family that originated from a common ancestral gene over 3 billion years ago (Danielson 2002). The paralogous pseudogene Cyp21a1-ps presumably arose as a tandem repeat of the functional Cyp21a1 gene (involved primarily in steroidogenesis) (Kawaguchi et al. 1992, Riepe et al. 2005), and therefore it belongs to the
Figure 7 Alignment of the 12 Cyp21a1-ps-like clones with the sequence for the Cyp21a1-ps gene at locus 20p12 of the rat genome. The portion of the reference sequence that is shown extends from bp 941,906 to 942,534 on chromosome 20. White letters on black background indicate nucleotides that are identical with the reference sequence. Black letters on gray background indicate nucleotides that are identical among a given set of clones positioned on the same loci, but different from the reference sequence. Black letters on white background indicate additional sites of point mutations in various clones, when compared with the reference gene. Empty spaces (indicated by dashes) in the reference sequence designate segments along one or more of the clones where there are nucleotide insertions. Empty spaces along segments of the clones where nucleotides are present in the reference gene indicate sites on the clones where there are nucleotide deletions. Horizontal lines separate clones according to gene locus as listed in Table 2.

| Reference | CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT
<table>
<thead>
<tr>
<th>Clone #1</th>
<th>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone #2</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #3</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #4</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #5</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #6</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #7</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #8</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #9</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #10</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #11</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Clone #12</td>
<td>CTTTCTGACAGAGCTGGGCGAACGTTGGCACTCAAAGAGATATGATTTTTTAAGACATCATTTCCAGGTTTACTGTAAT</td>
</tr>
</tbody>
</table>

* Reference sequence begins at locus 941,906 on chromosome 20 of the rat genome.
Figure 8 Alignment of the five Akr1c6-like clones with two different segments of the sequence of the Cyp21a1-ps gene at locus 20p12 of the rat genome. (A) This reference sequence extends from bp 927 806 to 928 316 on chromosome 20. (B) This reference sequence extends from bp 943 078 to 943 589 on chromosome 20. See the legend of Fig. 7 for further description.
category of ‘duplicated’ pseudogenes. The scientific literature contains conflicting information about whether pseudogenes such as Cyp21a1-ps can, or cannot, produce transcripts (Mighell et al. 2000, Zhang & Gerstein 2004, Prudhomme et al. 2005). Clearly, the present data from DD RT-PCR, northern analyses, and in situ hybridization, collectively, make it evident that a number of paralogs of Cyp21a1-ps are transcribed in the ovary at the time of ovulation in a cell-specific manner. Apparently, in some cases, pseudogenes retain or acquire functional promoter sequence that allows them to be transcribed (Mighell et al. 2000, Zhang & Gerstein 2004). In other instances, pseudogene transcription can be driven by a nearby promoter that is present in an unrelated, upstream sequence (Mighell et al. 2000). However, the upstream genes that flank each of the transcribed Cyp21a1-ps-like clones in this study are not known to be ovulation related (Table 2).

The specific function(s) of gonadally expressed Cyp21-type pseudogenes is uncertain. It has been suggested that the mass of pseudogenes that now clutter the mammalian genome could have a somewhat indirect function, namely, they could serve as a ‘pool’ of DNA sequence that has the potential of generating new genes that spontaneously gain functionality during the course of evolution (Zhang & Gerstein 2004). In other words, pseudogenes might serve as a reservoir of potentially functional genes or ‘potogenes’ (Balakirev & Ayala 2003).

The sequence similarity between Cyp21a1-ps and its normal Cyp21 counterpart has resulted in the loading of sequence databases with inappropriate ‘protein-coding’ transcripts (Zhang & Gerstein 2004, Claverie 2005, van Baren & Brent 2006). For example, several years ago, we reported that 3α-hydroxysteroid dehydrogenase (3α-HSD) mRNA was transcribed in the rat ovary in response to an ovulatory dose of gonadotropin (Espey et al. 2001). (Note that, under present nomenclature, 3α-HSD is now abbreviated as Akr1c6). However, based on the growing knowledge of pseudogenes and on the completion of the rat genome project, it is now evident that our previously reported sequence for ovarian 3α-HSD is actually a Cyp21a1-ps-like...
Table 2 Genomic loci and flanking features of clones

<table>
<thead>
<tr>
<th>Clone #</th>
<th>Locus</th>
<th>Similarity</th>
<th>5′-Flanking feature</th>
<th>3′-Flanking feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyp21a1</td>
<td>1q22</td>
<td>100%</td>
<td>Similar to F49E2.5d</td>
<td>Similar to zinc-finger protein 420</td>
</tr>
<tr>
<td></td>
<td>1q22</td>
<td>99%</td>
<td>Similar to F49E2.5d</td>
<td>Similar to zinc-finger protein 420</td>
</tr>
<tr>
<td></td>
<td>3q41</td>
<td>100%</td>
<td>Hypothetical protein LOC362235</td>
<td>Similar to ribosomal protein L19 and cystatin S</td>
</tr>
<tr>
<td></td>
<td>3q41</td>
<td>99%</td>
<td>Hypothetical protein LOC362235</td>
<td>Similar to ribosomal protein L19 and cystatin S</td>
</tr>
<tr>
<td></td>
<td>3q41</td>
<td>99%</td>
<td>Hypothetical protein LOC362235</td>
<td>Similar to ribosomal protein L19 and cystatin S</td>
</tr>
<tr>
<td></td>
<td>3q41</td>
<td>99%</td>
<td>Hypothetical protein LOC362235</td>
<td>Similar to ribosomal protein L19 and cystatin S</td>
</tr>
<tr>
<td></td>
<td>11p12</td>
<td>100%</td>
<td>Similar to voltage-dependent anion channel 1</td>
<td>Similar to U1 small nuclear ribonucleoprotein C</td>
</tr>
<tr>
<td></td>
<td>11p12</td>
<td>100%</td>
<td>Similar to voltage-dependent anion channel 1</td>
<td>Similar to U1 small nuclear ribonucleoprotein C</td>
</tr>
<tr>
<td></td>
<td>11p12</td>
<td>100%</td>
<td>Similar to voltage-dependent anion channel 1</td>
<td>Similar to U1 small nuclear ribonucleoprotein C</td>
</tr>
<tr>
<td></td>
<td>11p12</td>
<td>99%</td>
<td>Similar to voltage-dependent anion channel 1</td>
<td>Similar to U1 small nuclear ribonucleoprotein C</td>
</tr>
<tr>
<td></td>
<td>14p22</td>
<td>99%</td>
<td>Similar to vomeronasal 2, receptor 2</td>
<td>Similar to vomeronasal 2, receptor 1</td>
</tr>
<tr>
<td></td>
<td>14p22</td>
<td>100%</td>
<td>Similar to vomeronasal 2, receptor 2</td>
<td>Similar to vomeronasal 2, receptor 1</td>
</tr>
<tr>
<td></td>
<td>14p22</td>
<td>99%</td>
<td>Similar to vomeronasal 2, receptor 2</td>
<td>Similar to vomeronasal 2, receptor 1</td>
</tr>
<tr>
<td></td>
<td>14p22</td>
<td>100%</td>
<td>Similar to vomeronasal 2, receptor 2</td>
<td>Similar to vomeronasal 2, receptor 1</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>100%</td>
<td>Similar to mammalian retrotransposon-derived 8b</td>
<td>(No feature available)</td>
</tr>
<tr>
<td>Akr1c6</td>
<td>1q22</td>
<td>100%</td>
<td>Similar to F49E2.5d</td>
<td>Similar to zinc finger protein 507</td>
</tr>
<tr>
<td></td>
<td>14p22</td>
<td>100%</td>
<td>Similar to vomeronasal 2, receptor 2</td>
<td>Similar to vomeronasal 2, receptor 1</td>
</tr>
<tr>
<td></td>
<td>14p22</td>
<td>99%</td>
<td>Similar to vomeronasal 2, receptor 2</td>
<td>Similar to vomeronasal 2, receptor 1</td>
</tr>
<tr>
<td></td>
<td>14p22</td>
<td>100%</td>
<td>Similar to vomeronasal 2, receptor 2</td>
<td>Similar to vomeronasal 2, receptor 1</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>100%</td>
<td>Similar to mammalian retrotransposon-derived 8b</td>
<td>Similar to teashirt 3</td>
</tr>
<tr>
<td>SC1</td>
<td>1q21</td>
<td>100%</td>
<td>Similar to zinc finger protein 507</td>
<td>Hypothetical protein LOC500501</td>
</tr>
<tr>
<td></td>
<td>5q33</td>
<td>99%</td>
<td>Similar to tumor suppressor candidate 1</td>
<td>Regulator of G-protein signaling 2</td>
</tr>
<tr>
<td></td>
<td>1q321</td>
<td>100%</td>
<td>Predicted ubiquitin C-terminal hydrolase L5</td>
<td>Cadherin 13</td>
</tr>
<tr>
<td></td>
<td>19q12</td>
<td>99%</td>
<td>Cadherin 13</td>
<td>Similar to protein CXorf22</td>
</tr>
<tr>
<td></td>
<td>Xq22</td>
<td>100%</td>
<td>Hypothetical protein</td>
<td>Matrix metalloproteinase 9</td>
</tr>
</tbody>
</table>

pseudogene that is identical to clone 15 in the present paper (Fig. 8). That is to say, one of the paralogs of Cyp21a1-ps in the present study was reported previously as 3α-HSD (i.e., as Akr1c6).

Over the course of millions of years, there has been a huge proliferation of genomic elements that are derived from only a few initial germ line invasions by exogenous retroviruses such as SC1 that become endogenous (Anway et al. 2001, Belshaw et al. 2004). These ERVs now appear to exist as rather permanent ‘symbiotic’ or ‘parasitic’ DNA sequences that use their hosts’ replicative machinery for vertical transmission of their DNA along the evolutionary tree over the course of time (Prudhomme et al. 2005). Thus, ERVs can persist within any given individual as ‘proviruses’ that possess the same genetic machinery as a viable virus (Anway et al. 2001, Belshaw et al. 2004, Prudhomme et al. 2005). During the course of their replication (by retrotransposition or transposition) within a given individual, they retain transcription initiation and termination elements in their long-terminal repeat segments, and there is evidence that a significant number of these pseudogenes actually are transcribed (Prudhomme et al. 2005, Svensson et al. 2006). For example, the retrovirus SC1 that is clearly expressed during the ovulatory process contains CAAT, TATA, and GATA ‘box promoter elements and many putative transcription factor binding site regulatory elements’ (Anway et al. 2001) or see NCBI accession number AY009092.

The first significant report of SC1 expression in mammalian ovarian tissue was published in 2001 (Anway et al. 2001). This study found that maximum transcription of SC1 in immature rat ovaries occurred at 48 h after the animals were primed with 10 units of eCG – a time point that corresponds with what most investigators considered to be 0 h into the ovulatory process in the immature rat model. Furthermore, this earlier study found that retroviral transcription in the rat ovary was relatively low at 14 h after the rats had been injected with hCG to initiate the ovulatory process. These earlier findings are quite different from the results of the present study. Clearly, the results in Fig. 3 shows that all five of the different clones for SC1 transcripts were at the lowest expression levels at 0 h into the ovulatory process (i.e. at 48 h after the administration of eCG to prime the immature animals), while the greatest expression was at 4–8 h into the ovulatory process. The reason(s) for these distinctly opposite findings in the present study are unclear.

It has been suggested that replication-competent ERVs that are comparable with SC1 can undergo persistent reinfection of a host organism (Belshaw et al. 2004). Such a chronic condition can result in ‘extremely long periods of smoldering infection’ (Belshaw et al. 2004). Therefore, since there is ample evidence that an ovulatory surge in gonadotropin hormones induces an acute inflammatory reaction in mammalian ovaries (Espey 1980, 1994, 2001, Hernandez-Gonzalez et al. 2006), it is difficult not to consider the possibility that the hormonally induced expression of SC1 in the rat ovary might contribute in some way to the inflammatory degradation and rupture of the follicle wall. In any event, the present data confirm that expression of transcripts of SC1 is associated with a specific inflammatory reaction in the ovary. Since recent studies also document that ovulating follicles express genes associated with innate immune responses that detect ‘self’ from ‘non-self’, or ‘altered-self’, it is possible that the SC1 mRNAs may impact these signaling pathways as well (Shimada et al. 2006). In future
studies on the molecular events of ovulation, it would be helpful to learn whether SC1 elements induce the expression of proinflammatory cytokines or vice versa. If the former scenario is the case, then it would also be interesting to know whether antibodies to epitopes on the gag, pol, and/or env proteins from SC1 can affect the efficiency of mammalian ovulation.

In conclusion, it is evident that transcripts of pseudogenes of the types Cyp21a1-ps and retrovirus SC1 are a simultaneous event in the ovulatory process of immature rats. While there is no conspicuous role for the Cyp21a1-ps-like elements in the mechanism of ovulation, it is possible that translated products of endogenous SC1, such as the envelop protein for SC1, might contribute in some significant way to the ovarian inflammatory reaction that is a fundamental component of the ovulatory process.

Declaration of Interest

The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

Funding

Supported in part by NSF Grant 0234358 (L L E) and NIH Grants HD-16229 and HD-07495 (J S R).

Acknowledgements

We appreciate the assistance of Mrs Claire Lo in preparing the in situ hybridization data and the assistance of Ms Karla Moncada in generating some of the clones in this study.

References


Hurst LD 2002 The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends in Genetics* 18 486.


www.endocrinology-journals.org


Received in final form 24 April 2008
Accepted 1 May 2008
Made available online as an Accepted Preprint 2 May 2008