SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
CELL GROWTH
INFERTILITY

Received
21 March 2014

Accepted
18 June 2014

Published
4 July 2014

Correspondence and requests for materials should be addressed to J.W. (jiwu@sjtu.edu.cn)

Similar morphological and molecular signatures shared by female and male germline stem cells

Wenhai Xie[1], Hu Wang[1] & Ji Wu[1,2]

[1]Key Laboratory for the Genetics of Developmental & Neuropsychiatric Disorders (Ministry of Education), Bio-X Institutes, Shanghai Jiao Tong University, Shanghai 200240, China, [2]Key Laboratory of Fertility Preservation and Maintenance of Ministry of Education, Ningxia Medical University, Yinchuan 750000, China.

The existence of mammalian female germline stem cells (FGSCs) indicates that mammalian ovaries possess germline stem cells analogous to testis, and continue to produce gametes postnatally, which provides new insights into female fertility. In this study, we compared the morphological and molecular characteristics between FGSCs and spermatogonial stem cells (SSCs) by analysis of morphology, immunofluorescence, alkaline phosphatase activity assay, reverse transcription polymerase chain reaction (RT-PCR) and microarray hybridization. The results demonstrated that the morphology and growth patterns of FGSCs are similar to those of SSCs. Microarray analysis of global gene expression profiles of FGSCs and SSCs showed similar signatures in the transcriptome level. A list of 853 co-highly expressed genes (CEG) in female and male germline stem cells may be responsible for the morphological and molecular similarity. We constructed a continuous network of the CEG based on I2D protein-protein interaction database by breadth first search. From the network, we could observe the interactions of the CEG may be responsible for maintaining the properties of germline stem cells. This study was the first attempt to compare morphological and molecular characteristics between FGSCs and SSCs. These findings would provide some clues for further research on mammalian FGSCs.

Stem cells are cells that can renewing themselves and can differentiate into mature specialized body cells. Using these characteristics, as well as morphology and surface markers, many types of adult stem cells have been identified[1,2]. Germline stem cells share two characteristics with adult stem cells, in that they can self-renew and provide daughter cells[3]; however, they are also responsible for genetic information transmission from parents to subsequent generations[4]. In mammals, male germline stem cells (spermatogonial stem cells, SSCs) were identified and confirmed very early, because the male can produce sperm throughout its entire lifetime[5]. In the conventional theory, however, female germline cells of most mammalian species enter meiosis and are arrested at the diplotene stage of meiotic prophase I before birth[6–8]. This theory has been challenged by findings suggesting that postnatal oogenesis occurs in mouse ovaries[9].

Recently, we showed that female germline stem cells (FGSCs) from neonatal and adult mouse ovaries could be successfully isolated and purified using immunomagnetic sorting for the mouse vasa homolog *(MVH)*. These FGSCs are located on the cortical surface of ovaries. Furthermore, an FGSC line was established from neonatal mice, with a normal karyotype and high telomerase activity. Adult FGSCs have also been cultured for more than 6 months. FGSCs in long-term culture maintain their capacity to produce normal oocytes and fertile offspring after transplantation into ovaries[10]. Several germ cell-specific markers have been screened to improve the efficiency of purification. Using the germline-specific protein *Fragilis*, the efficiency of FGSC purification was remarkably enhanced compared with that using *MVH*[11]. We have also successfully generated transgenic and gene knockdown mice using FGSC[12]. Pacchiarotti *et al.*[13] and Hu *et al.*[14] also demonstrated the existence of a population of germline stem cells in postnatal mouse ovaries. White *et al.* successfully isolated and purified FGSCs from adult mice and reproductive-age women using fluorescence-activated cell sorting for *Ddx4* (also called *Vasa*)[15]. Moreover, we have successfully generated transgenic rats using an FGSC line established from postnatal rats[16]. However, some recently published papers declared no mitotically active FGSCs exist in postnatal mouse ovaries[17,18]. It is necessary to describe the morphological and molecular characters of female germ line stem cells and compare similarity of FGSCs and SSCs.

In the present study, we compared the morphological characteristics of FGSCs and SSCs and performed microarray analysis of the global gene expression profiles of both cell types. Although sex differentiation occurs

at 11.5 days post-coitum (dpc)[19], there are still some similarities in the whole gene expression profiles between FGSCs and SSCs. The co-highly expressed genes (CEG) shared by FGSCs and SSCs may be responsible for those similarities. Functional annotation analysis of the CEG enriched GO terms related to cell cycle, cell division and proliferation and transcription regulation. Overlaps between the CEG and genes overexpressed in embryonic stem cell (ESCs), neural stem cells (NSCs) and hematopoietic stem cells (HSCs) validated the stemness of FGSC and SSC[20]. Motif screening identified several transcription factors related to cell cycle and stem cell maintenance. To gain further insights into the interactions of the CEG responsible for the similar properties of FGSCs and SSCs, we constructed a continuous protein-protein interaction network of the CEG based on I2D database by breadth first search. From the network we observed the interactions of the CEG responsible for maintaining the properties of germline stem cells. We obtain 35 hubs whose connectivity ranked within the top 10% most connected nodes in the maximum continuous network. According to the theory of scale-free networks, those hubs would play important roles in the continuous network that may be responsible for maintaining the similarities between FGSCs and SSCs. The functional analyses of the CEG and motif enrichment of stem special transcription factors led us to hypothesize that the CEG are responsible for the similar morphological and molecular characters shared by FGSCs and SSCs. The constructed continuous network of the CEG will provide some useful clues for studies focusing on the mechanisms of maintaining the unique properties of germline stem cells.

Results

Morphology and growth patterns of FGSCs are similar to those of SSCs. The morphology of FGSCs is very similar to that of SSCs. They are characterized by a round shape with a large bright spherical nucleus and an obvious boundary between the cytoplasm and nucleus (Fig. 1A). After culturing for 2 days, the isolated FGSCs appeared as grape-like clusters that are usually found in cultured SSCs (Fig. 1A). The nuclear boundary of the proliferated FGSCs was observed under light microscope, and was consistent with the result using Hoechst33342 staining (Supplementary Fig. S1 online).

The FGSCs possess germ cell properties similar to those of SSCs. To confirm that the FGSCs possess germline stem cell-specific properties, we performed the following experiments. First, we detected the expressions of germ cell-specific markers in the cultured FGSCs using reverse transcription polymerase chain reaction (RT-PCR). The results showed that FGSCs expressed *MVH*[21], *Dazl* (a germ cell-specific RNA-binding protein)[22], *Oct4* (a germ cell-specific transcriptional factor)[23] and *Fragilis* (a germ cell-specific expression protein)[24] (Fig. 1B). Immunofluorescence analysis showed that the FGSCs were Oct4 and MVH positive, as were SSCs (Fig. 1C). We also detected alkaline phosphatase staining of FGSCs, as has been previously described[10] (Fig. 1C). All the characteristics detected were consistent with previously identified female germline stem cells. Immunofluorescence analysis of BrdU incorporation and expression of MVH demonstrated that the FGSCs possessed proliferative ability (Fig. 2). These germ cell properties of FGSCs are similar to those of SSCs.

The gene expression profiles in FGSCs and SSCs show similar signatures. We detected 19004 probes that were expressed in the gene expression profiles of FGSCs and SSCs. Initially we explored the microarray results of FGSCs and SSCs at the transcriptional level using unsupervised hierarchical clustering of the global gene expression profiles. A heat map for the whole gene expression profiles of FGSCs and SSCs showed obvious blocks representing similar gene expressing patterns for certain genes. As shown in Fig. 3A, the gene expression profiles of FGSCs and SSCs were similar. To identify key genes conserved in germline stem cells that

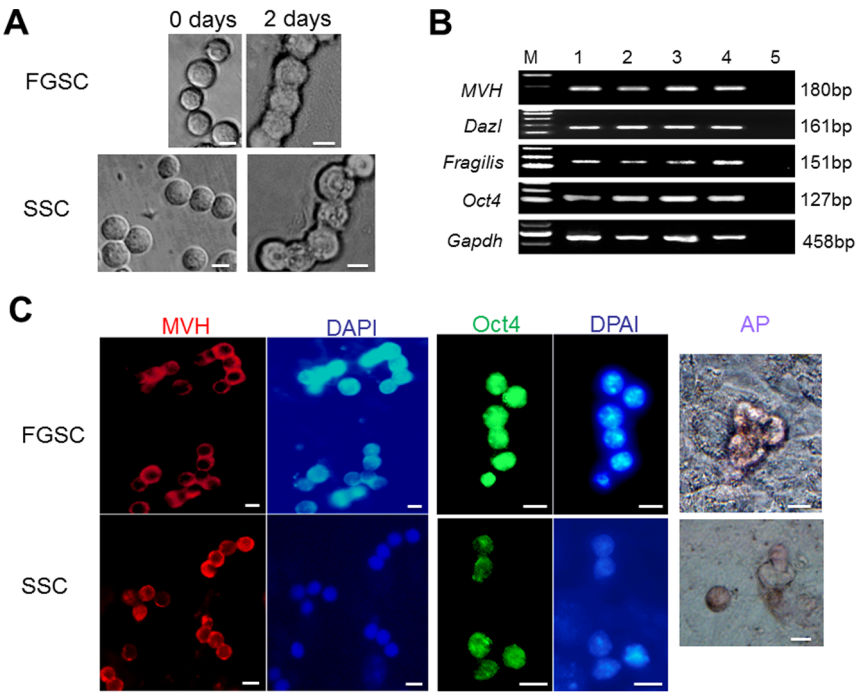

Figure 1 | Confirmation that FGSCs possess germline stem cell properties similar to SSCs. (A) Freshly isolated FGSCs and SSCs have a round shape with large, bright, spherical nuclei; there is an obvious boundary between the cytoplasm and the nucleus. Isolated FGSCs and SSCs form grape-like clusters after 2 days of culture. Scale bar = 10 μm. (B) RT-PCR analysis of germ cell markers. M, 100 bp DNA marker; lane 1, freshly isolated FGSC; lane 2, 2-day cultured FGSC; lane 3, freshly isolated SSC; lane 4, 2-day cultured SSC; lane 5, no template control. (C) Immunofluorescence analysis of MVH, Oct4 as well as alkaline phosphatase (AP) was used for characterization, and DPAI was used to stain the nucleus. Similar to SSCs, FGSCs are positive for MVH and Oct4. FGSCs are weakly stained by the AP assay, similar to SSCs. Scale bar = 10 μm.

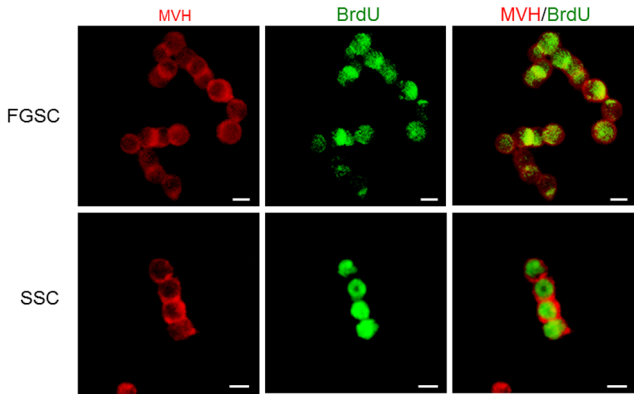

Figure 2 | FGSCs possess proliferative abilities similar to SSC. Dual immunofluorescence for MVH and BrdU showed that FGSC show mitotic activity and maintaining germline properties, similar to SSC. Scale bar = 10 μm.

might be responsible for the similar morphology of FGSCs and SSCs, we generated a list of CEG shared by both cell types based on the same normalized threshold (see methods). After removing housekeeping genes, there were 1273 and 1193 highly expressed genes in FGSCs and SSCs, respectively, of which 853 (52.88%) were CEG (Fig. 3B). The list of CEG is shown in Supplementary Table S1. The expression levels of several germ cell-specific genes in FGSCs and SSCs are shown in Fig. 3C. *MVH*, *Dazl*, *Oct4* and *Fragilis*, which are commonly used to identify germ cells, were expressed in FGSCs and SSCs at levels close to the high expression threshold. *Tdrkh*, *Akt3*, *Gm1673*, *Hba-a1*, *Mov10l1* and *Fkbp6*, which are specific to the maintenance of mouse primordial germ cells (PGC) during embryo development[25], all showed expression levels above the high expression threshold in FGSCs and SSCs (Fig. 3C).

Adult stemness genes shared by FGSCs and SSCs. In the whole gene expression profiles, we detected 216 adult stemness genes that were

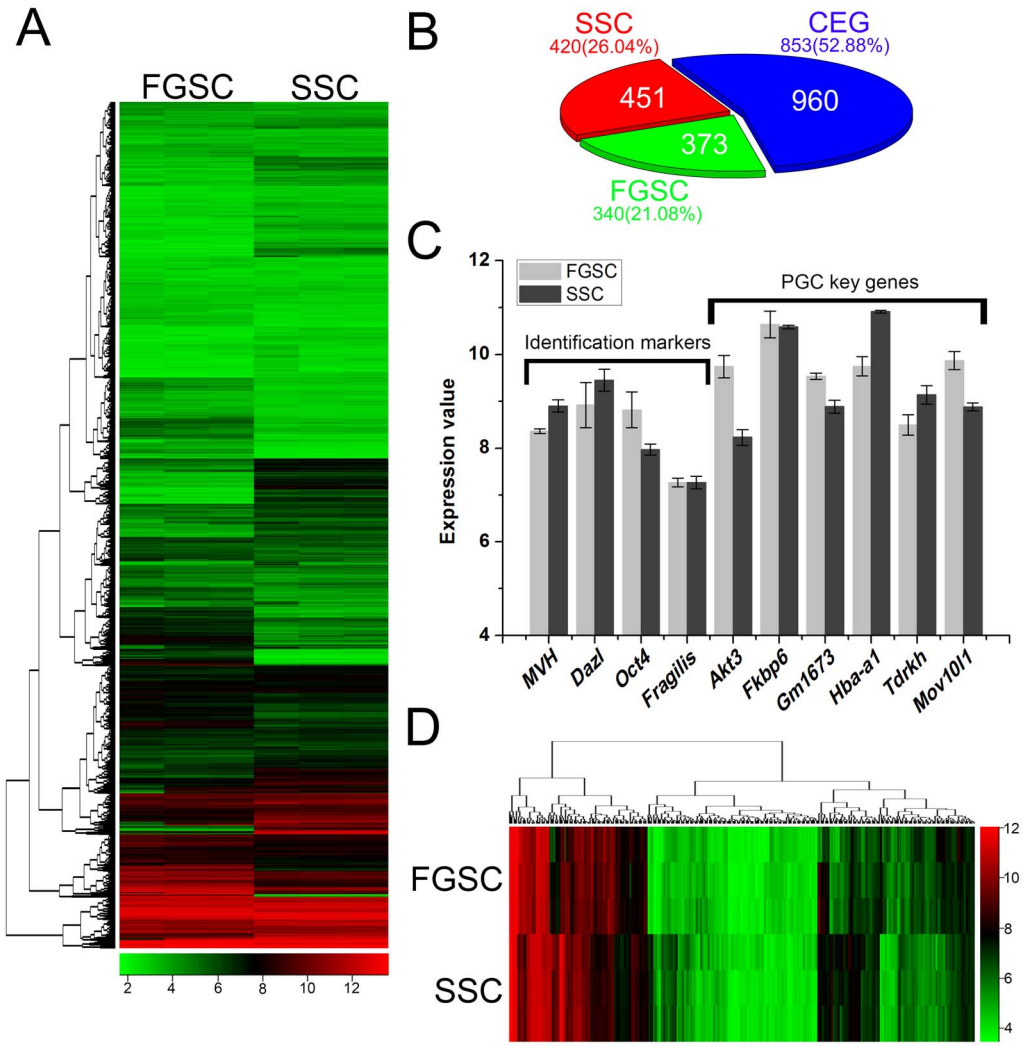

Figure 3 | Microarray analysis of gene expression profiles of FGSCs and SSCs. (A) Heatmap for the global gene expression profiles of FGSCs and SSCs. The heatmap for global gene expression profiles shows that FGSCs and SSCs share similarities at the transcriptional level. (B) A pine diagram showing the overlap of highly expressed genes of FGSCs and SSCs. The numbers of respective highly expressed genes was shown corresponding to FGSC and SSC. The CEG represented by blue portion refers to 853 genes. The proportion of respective parts is shown on the diagram. (C) Expression levels of germ cell special genes in FGSCs and SSCs detected by the microarray. Germ cell markers include *MVH*, *Dazl*, *Oct4* and *Fragilis*. PGC key genes, such as *Tdrkh*, *Akt3*, *Gm1673*, *Hba-a1*, *Mov10l1* and *Fkbp6* were expressed at similarly high levels in FGSCs and SSCs. (D) Expression heatmap for 216 stemness genes detected by the microarray show similar expression patterns in FGSCs and SSCs. In the heatmap, each row represents the expression level of one gene in the gene expression profiles of FGSCs and SSCs. Heatmap visualization was performed by gplots package in R. Red represents highly expressed; green means low expression value.

Table 1 | Comparison of the CEG with other stem cells overexpressed genes profiles

	ESC (%)	NSC (%)	HSC (%)
CEG(853)	130(15.24)	177(20.75)	146(17.12)
18 core genes	*Acadm,AW549877,Ccs,Coprs,Dtymk,Fam3c,Fhl1,Fkbp9,Gsta4,Kras,Nme1,Pvrl3,Rasa1,Ryk,Slc4a7,Zc3h14,Zfx,Zmym4*		

proposed as molecular signatures of stem cells by Melton *et al.*[20]. Unsupervised hierarchical clustering of the detected adult stemness genes showed that they had very similar expression patterns between FGSCs and SSCs (Fig. 3D). Hierarchical clustering analysis classified the stemness genes into several obvious blocks. The stemness genes in the red block may be important for the stemness of FGSCs and SSCs. The number of CEG that overlapped with the profiles of overexpressed genes in mouse embryonic, neural, and hematopoietic stem cells are shown in Table 1. The presence of hundreds overlapping genes implies that FGSCs and SSCs share stemness properties with other types of stem cells. From Table 1, it is clear that the CEG shared a larger overlap with neural stem cells than with embryo stem cells or hematopoietic stem cells, suggested that reproductive system share some genes expression patterns with nervous system.

According to the overlap between the CEG and the genes overexpressed in the other three stem cell types, we classified the CEG into two subsets: (i) $CEG^{-stemness}$ (323 genes) representing the CEG that overlapped with the overexpressed genes of the other three types of stem cells; (ii) CEG^{-germ} (530 genes) representing the CEG that did not overlap with the overexpressed genes of the other three types of stem cells. Further analysis of the $CEG^{-stemness}$ identified 18 genes that act as core genes shared by germline, embryonic, neural and hematopoietic stem cells (Supplementary Fig. S2 online). The 18 core genes shared by CEG, ESC, HSC and NSC were listed in Table 1.

The presence of adult stemness genes in the CEG was also consistent with previous studies that proposed that spermatogonial stem cells share some, but not all, phenotypic and functional characters with other stem cells[3]. Some of the CEG were not overexpressed in the other stem cells, indicating that these genes may represent germline stem cell-specific genes.

Transcription factors enriched by motif screening of cis-regulatory elements of the CEG. Motif screening of promoter *cis*-regulatory elements of the CEG would identify transcription factors that regulate the expression of certain CEG. We applied EXPANDER[26] to identify transcription factors whose binding site signatures were significantly over-represented in the CEG. Several interesting transcription factors were identified. Cell cycle-related transcription factors *c-Myc*[27] and *Egr-1*[28] were identified, which are consistent with the rapid cell division properties of FGSCs and SSCs. Motif screening also identified several transcriptional regulators that play key roles in stem cell maintenance and cell proliferation. Transcription factors *HSF* enriched several important genes, such as *Uchl1*[29] *Tex14*[30] *MII2*[31], which play important roles in maintenance of spermatogenesis. The enriched transcription factors and their target genes were shown in Supplementary Table S2. TF-binding site analysis of the CEG provided potential regulators to study mechanism of FGSC and SSC self-renewal. The list of identified transcriptional regulators will be a useful resource for scientists engaged in germline stem cell research.

Gene ontology analysis of the CEG. To assess the biological relevance of the CEG, we performed gene ontology analysis using the DAVID web server[32]. Functional analysis of the 853 CEG enriched GO terms mainly belonging to the three functional clusters (biological process, cellular component, and molecular function). The CEG in the cluster of biological processes were mainly involved in maintaining transcription, cell cycle, chromatin organization, metabolic process and protein localization (Supplementary Table S3 online). This is consistent with the rapid mitotic activity of FGSCs and SSCs. A graphical representation of biological process functional classification is shown in Supplementary Fig. S3. As FGSCs and SSCs display similar morphology and germline stem cell properties during in vitro culturing, we intended to identify genes with roles in mammalian germline stem cell maintenance; therefore, we generated lists of CEG related to cell proliferation, cell division and cell cycle (see Table 2). The list of cell proliferation related genes includes several genes that have been found to play important roles in maintaining the self-renewal of SSCs, such as *Uchl1*[29] and *Mov10l1*[33]. The adaptor protein Numb is expressed in male germ cells. Inactivation of *Numb* may influence spermatogonial differentiation[34]. *Irs2* is required for testicular development, but the exact role of *Irs2* in SSC remains unknown[35]. *Stab1* regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing commitment to differentiation[36]. *Srrt*, also known as *Ars2* has been found to play important roles in maintaining neural stem cell identity[37]. Functional analysis of those genes related to cell proliferation, cell division and cell cycle found that 17 genes involved in transcription regulation (see Table 2). The 17 transcription related genes were possibly as transcription regulators regulating expression of cell proliferation, cell division and cell cycle related genes. The lists of cell proliferation, cell division and cell cycle related genes provides several suggested starting points for direct studies on the mechanisms of germline stem cell proliferation.

Hub proteins of the maximum continuous network constructed from the CEG. We hypothesized that the CEG would be responsible for the morphological and molecular similarities between FGSCs and

Table 2 | The list of cell proliferation related genes enriched by GO analysis

Go Term	Enriched genes
Cell proliferation (20)	*Uchl1,Mov10l1,Trp53,Bcl2l2,Numb,Uhrf2,Nras,Irs2,Appl1,Gpx1,Uhrf1,Impdh2,Srrt,Naa35,Nasp,Kdm1a,Gm13841, Cep120,Satb1,Morf4l1*
Cell division (34)	*Ccnt1,Mad2l1,Lzts2,Zc3hc1,Haus8,Mcm5,Birc5,Zwint,Anln,Evi5,Pard6g,Cdk2,Nup43,Katna1,Ncaph,Mis18a,Incenp, Anapc4,Cdca3,Anxa11,Mau2,Tipin,Rnf8,Chfr,Sept2,Ncapd3,Plk1,Cdc20,Cdca8,Cdc25a,Prc1,Mapre1,Cdk1,Nek3*
Cell cycle (75)	*Mad2l1,Fbxo31,Lzts2,Zc3hc1,Gak,Dmtf1,Rbm38,Mdm2,Zwint,Anln,Evi5,Mcm6,Cdkn1c,Ppm1d,Cdk2,Katna1,Blcap, Mis18a,Nasp,Anapc4,Ppm1g,App,Chfr,Gm4799,Ncapd3,Tdrd1,Cdc20,Cdt1,Foxn3,Cdca8,Prc1,Ppp3ca,Mapre1,Tlk2, Rabgap1,Cep120,Ccnt1,Haus8,Birc5,Npm1,Zfp318,Atm,Gm9385,Pard6g,Piwil2,Nup43,Uhrf1,Ncaph,Dbf4,Incenp, Chtf18,Chaf1b,Cdca3,Anxa11,Aurka,Trp53,Mau2,Uhrf2,Tipin,Rnf8,Mfn2,Sin3a,Rpa1,Hjurp,Hjurp,Sept2,Appl1,Plk1, Ehmt2,Rassf1,Cdc25a,Clspn,E2f6,Cdk1,Nek3,Sirt7*
Transcription (17)	*Morf4l1,Dmtf1,Mcm5,Mcm6,Gm4799,Foxn3,Kdm1a,Ccnt1,Nsun4,Zfp318,Uhrf1,Chaf1b,Trp53,Sin3a,Fpr1,E2f6,Sirt7*

SSCs. Therefore, we wanted to identify key genes of the CEG that play important roles in the maintenance of germline stem cell-specific properties. According to the theory of the complex network of protein-protein interactions, we constructed a maximum continuous network from the CEG based on I2d database. The network consisted of 352 nodes connected via 862 edges. The degree index γ value of the constructed maximum continuous network was 1.7412. According to the Barabasi-Albert model, the constructed maximum continuous network was a scale-free graph. The nodes whose connectivity ranked in the top 10% of all nodes in the network act as hubs. The hubs of the network would be responsible for the similarities between FGSC and SSC. The list of the 35 hubs is shown in Table 3. The hubs were highlighted with red color in the constructed maximum continuous network (Fig. 4). Several hubs were associated with cell proliferation. For example, *Arrb2* is a regulatory protein involved in progastrin induction of colonic cell proliferation[38], and *Aurka* is involved in determining mitotic spindle orientation[39]. Scaffold protein *Shc1* temporally regulates the EGF signaling network[40]. The actual roles of these hub proteins in maintaining the properties of germline stem cells will be investigated in future studies.

The constructed network provided some clues for the regulatory mechanism of germline stem cell development. After mapping *Cdh1*, *Sall4* and *Uchl1* in the constructed network, we obtained a possible regulation network (Fig. 5). *Cdh1*, a surface marker of SSC also known as *E-cadherin*[41], *Uchl1* co-localizes with *Plzf* in mouse SSCs[29], and *Sall4* is an antagonist of *Plzf*[42]. The three germline-specific genes above bridged the *Trp53*, *Mdm2*, *Hspd1*, *Smn1* and *App* hubs in the constructed network.

Comparision of gene expression profiles between neonatal and adult FGSCs. Based on results from the microarray analysis above, we checked 18 stemness genes expression in neonatal (3 days old) and adult (6 weeks old) FGSCs. Using the RT-PCR analysis, we found that all of the 18 stemness genes are expressed in both FGSCs. Neonatal and adult FGSCs also express germline markers, such as *MVH, Dazl, Fragilis* and *Oct4*. While, pluripotency-associated genes expressed in PGCs, such as *Nanog, Ssea1* and *Sox2*, were neither expressed in neonatal FGSCs nor adult FGSCs (Supplementary Fig. S5 online).

Discussion

During the process of embryo development, sex differentiation occurs at 11.5 days postcoitum (dpc); at 12.5 dpc the phenotypes of the male and female gonads are distinguishable from each other. In this manuscript, we found that FGSCs are morphologically and molecularly similar to SSCs. Comparison of the whole gene expression profiles between FGSCs and SSCs further confirmed their similarity; the two cell types had similar signatures of globe transcription and the expression patterns of the stemness genes were similar.

Comparison of FGSCs and SSCs allowed us to identify conserved molecules related to the unique properties of germline stem cells. Eighteen genes were identified as core genes shared by germline stem cells, mouse embryonic, neural, and hematopoietic stem cells (Table 1). Several of these stemness genes have been found to play key roles in stem cell self-renewal. For example, *Zfx* (Zinc finger protein X-linked) controls the self-renewal and maintenance of embryonic stem cells[43,44]. The 18 stemness genes may play critical roles in the maintenance of stem cell properties. The detailed roles of

Table 3 | Hubs of the maximum continuous network constructed from the CEGs

Entrez Gene Id	Symbol	Definition	Degree
22059	*Trp53*	Transformation related protein 53	44
13627	*Eef1a1*	Eukaryotic translation elongation factor 1 alpha 1	40
22628	*Ywhag*	Tyrosine3-monooxygenase activation protein	32
17246	*Mdm2*	Transformed mouse 3T3 cell double minute 2	30
18607	*Pdpk1*	3-phosphoinositide dependent protein kinase 1	29
227700	*Sh3glb2*	SH3-domain GRB2-like endophilin B2	27
216869	*Arrb2*	Arrestin, beta 2	27
26416	*Mapk14*	Mitogen-activated protein kinase 14	24
20466	*Sin3a*	Transcriptional regulator SIN3A	24
18148	*Npm1*	Nucleophosmin 1	23
15510	*Hspd1*	Heat shock protein 1	21
18195	*Nsf*	N-ethylmaleimide sensitive fusion protein	21
15078	*H3f3a*	H3 histone, family 3A	21
67160	*Eef1g*	Eukaryotic translation elongation factor 1 gamma	20
11966	*Atp6v1b2*	ATPase, H+ transporting, lysosomal V1 subunit B2	19
17127	*Smad3*	SMAD family member 3	18
12534	*Cdk1*	Cyclin-dependent kinase 1	18
68275	*Rpa1*	Replication protein A1	17
20595	*Smn1*	Survival motor neuron 1	17
20416	*Shc1*	Src homology 2 domain-containing transforming protein C1	16
17217	*Mcm4*	Minichromosome maintenance deficient 4 homolog	16
227800	*Rabgap1*	RAB GTPase activating protein 1	15
13589	*Mapre1*	Microtubule-associated protein, RP/EB family, member 1	15
17219	*Mcm6*	Minichromosome maintenance deficient 6	15
58231	*Stk4*	Serine/threonine kinase 4	14
20878	*Aurka*	Aurora kinase A	14
12566	*Cdk2*	Cyclin-dependent kinase 2	13
18817	*Plk1*	polo-like kinase 1	13
11820	*App*	Amyloid beta precursor protein	13
319152	*Hist1h3h*	Histone cluster 1, H3h	13
21812	*Tgfbr1*	Transforming growth factor, beta receptor I	13
17218	*Mcm5*	minichromosome maintenance deficient 5	13
76788	*Klhdc10*	Kelch domain containing 10	13
13209	*Ddx6*	DEAD (Asp-Glu-Ala-Asp) box polypeptide 6	12
16202	*Ilk*	Integrin linked kinase	12

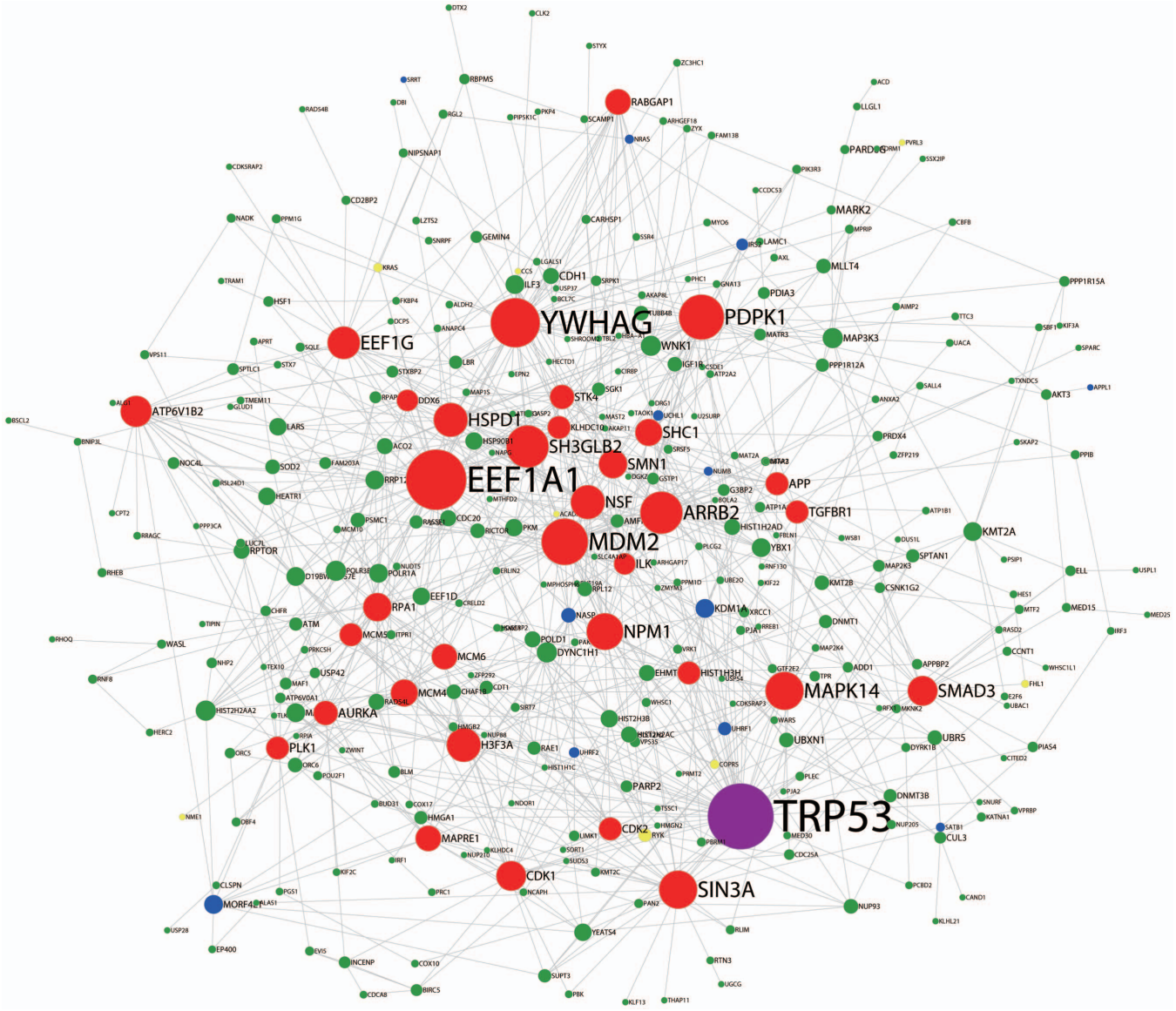

Figure 4 | An overview of the constructed maximum continuous network including 352 unique proteins of the CEG. The maximum continuous network includes 352 nodes and 862 edges. The size of nodes was corresponded to their degrees. The 35 hubs in the constructed network were highlighted with red colour. Core stemness genes (8 genes) shared by FGSCs and SSCs were labelled with yellow colour. Cell proliferation related genes (13 genes) locate in the network were labelled with blue colour. The remaining nodes in the network were painted green to discriminate from others.

these genes in germline stem cell self-renewal will be investigated in future studies.

Motif screening of the promoters of the CEG provided insights into the regulatory network of germline stem cell-specific genes. Transcription factors *HSF* enriched several important genes, such as *Uchl1*[29] *Tex14*[30] *MII2*[31], which have been found playing important roles in spermatogenesis. The functions of other *HSF* target genes in maintenance and self-renewal of germline stem cells will be revealed by future studies. Cell cycle-related transcription factor, *Egr-1* enriched 151 genes in the CEG^{-germ} set. Several target genes of *Egr-1* play important roles in stem cell self-renewal and maintenance. For example, *Dyrk1b* promotes proliferation of immature male germ cells in mice[45]. *Nme7* regulates several critical regulators in ESC self-renewal[46]. *Thap11* is essential for embryogenesis and the pluripotency of mouse embryonic stem cells[47]. *Satb1* regulates the self-renewal of hematopoietic stem cells[36].

As the limitation of confirmed interactions between proteins, the I2D database is incomplete. The protein-protein interaction database does not comprise all of the interactions between mouse genes. Although it is impossible to construct a network covering all of the CEG, the constructed network is also a useful resource for researches in germline stem cells. Although the incomplete information of protein interactions, the constructed network still provide some clues for germline stem cell research. Most of cell proliferation related genes were located in the constructed continuous network (Supplementary Fig. S4A online). Extracted those cell proliferation related genes and their first neighbors from the network would form a continuous subnetwork. The continuous subnetwork consisted of 13 cell proliferation related genes, 16 hubs and another 26 nodes suggests a possible regulation network responsible for germline stem cell proliferation (Supplementary Fig. S4A online). From the subnetwork, we can see that *Appl1* interacted with *Akt3* aparting from other cell proliferation related genes which means *Appl1* may be play roles in some other different pathways. *Appl1* acts as a critical regulator of the crosstalk between adiponectin signaling and insulin signaling pathways[48]. About the 18 core stemness genes shared by germline stem cells, embryonic stem cells, hematopoietic stem cells and neural stem cells, there are only 8 genes locate at the constructed network.

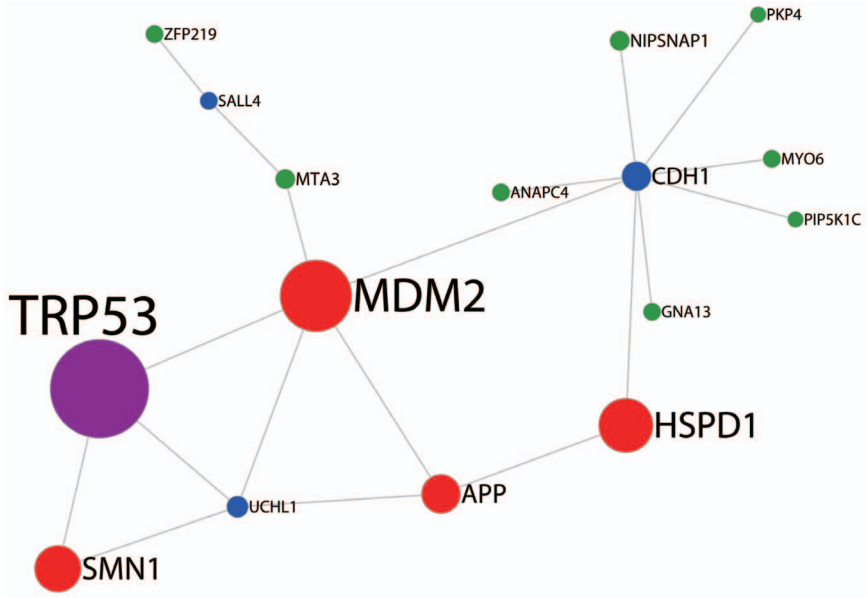

Figure 5 | A sub-network generated by key SSC genes Cdh1, Sall4 and Uchl1. The three germline specific genes *Cdh1*, *Sall4* and *Uchl1* are bridged together by hub *Mdm2* in the constructed network. There are another five hubs in the sub-network. They are *Trp53*, *Mdm2*, *Hspd1*, *Smn1* and *App*.

The 8 core stemness genes connected with each other through 9 hubs (Supplementary Fig. S4B online). Those scattered nodes connected together through hubs confirmed the importance of hubs in maintaining the stability of network. To access the function of single hub will provide new sights for germline stem cell research. For example, *Mdm2* may be bridge SSC special genes *Cdh1*, *Uchl1* and *Sall4* to maintaining germline stem cell properties. Functional analysis of the hubs indicated that most of the hubs were related to cell cycle, mitosis or acetylation, which is consistent with the mitotic activity of FGSCs and SSCs.

In our previous studies, FGSCs from adult mouse ovaries display similar properties of FGSCs from neonatal ovaries[10]. In this study, we compared the gene expression profiles between neonatal and adult FGSCs by RT-PCR analysis. The result showed that the neonatal FGSCs have same gene expression profile with adult FGSCs, while, the neonatal FGSCs do not express the marker genes which highly expressed in PGCs. This means that the FGSCs are not the same as PGCs, which also consistent with our previous studies[10].

In summary, our studies compared the morphological and molecular characteristics of FGSCs and SSCs. Microarray analysis of the globe gene expression profiles of FGSCs and SSCs showed similar signatures at the transcriptional level, which may be responsible for the similar morphologies of the two germline stem cell populations. Motif screening of the promoters of CEG and the constructed continuous network provided insights into the regulatory network of germline stem cell-specific genes. This CEG is a useful resource for scientists engaged in germline stem cell research.

Methods

Animals. CD-1 mice were used in this study. Three-day-old or 6-week-old female and 6-day-old male mice were used to isolate germline stem cells. The Institutional Animal Care and Use Committee of Shanghai approved all the procedures used in this study, which were performed in accordance with the National Research Council Guide for Care and Use of Laboratory Animals.

Isolation, purification and culture of germline stem cells. The FGSCs from 16 ovaries of 3-day-old mice or from 4 ovaries of 6-week-old (adult) mice were isolated, purified and cultured using a previously described method[10,12,49–51]. The SSCs from 12 testes of 6-day-old mice were isolated, purified and cultured using the method of Yuan *et al*[52]. After culturing the germline stem cells for 2 days, alkaline phosphatase activity assay, BrdU labeling, immunofluorescence, RNA isolation and microarray hybridization were performed (see below). For microarray hybridization, we cultured the FGSC and SSC in a feeder-free culture system. The culture medium for germline stem cells comprised minimum essential α medium (Gibco, Grand Island, NY, USA). 10% FBS (Front Biomedicals USA, Eatontown, New Jersey, USA), sodium pyruvate (1 mM, Sigma, St. Louis, MO, USA), non-essential amino acids (1 mM, Gibco), L-glutamine (2 mM, Sigma), β-mercaptoethanol (0.1 mM, Sigma), LIF (10 ng/ml, Santa Cruz, Dallas, TX, USA), mouse epidermal growth factor (10 ng/ml, Peprotech, Rock Hill, NJ, USA), glial cell line-derived neurotrophic factor (10 ng/ml, Peprotech), basic fibroblast growth factor (10 ng/ml, Peprotech) and penicillin (15 μg/ml, Sigma). After 2 days of culture, primary oocytes were easily discernible from proliferated FGSCs; the oocytes could be picked out using micropipette under a stereomicroscope. The remaining cells comprised the FGSCs.

Alkaline phosphatase activity assay. The cultured germline stem cells were fixed with 4% paraformaldehyde for no more than 2 min at room temperature. The alkaline phosphatase activity of the germline stem cells was measured using an alkaline phosphatase activity detection kit, according to the manufacturer's instructions (Millipore, Billerica, MA, USA).

Hoechst 33342 staining. After incubated with Hoechst 33342 (1 μg/ml, Sigma) at room temperature for 20 min, the FGSCs were rinsed twice in PBS and then ready for analysis.

BrdU labeling. BrdU (50 μg/ml, Sigma) was added to germline stem cell cultures for 5 h. The germline stem cells were then ready for immunofluorescence of BrdU staining.

Immunofluorescence. Cultured FGSCs and SSCs were washed twice with PBS (phosphate-buffered saline) and then fixed for 15 min with 4% paraformaldehyde at room temperature. After fixation, the cells were rinsed twice in PBS and then incubated with PBS containing 0.1% (v/v) Triton X-100 for 15 min, at room temperature. After rinsing, the fixed cells were blocked with 10% normal goat serum in PBS for 10 min at room temperature. The cells were then incubated at 4°C overnight with primary antibodies: rabbit polyclonal anti-MVH (1 : 180 dilution; Abcam, Cambridge, MA, USA), rabbit polyclonal anti-Oct4 (1 : 150 dilution; Chemicon, Temecula, CA, USA), and mouse monoclonal anti-BrdU (1 : 150 dilution; Lab Vision Corporation, Fremont, CA, USA). After incubation with primary antibodies, the cells were rinsed twice in PBS, and then incubated in darkness with a fluorescein isothiocyanate-conjugated secondary antibody either goat anti-rabbit IgG or goat anti-mouse IgG (1 : 180 dilution; Proteintech, Chicago, USA) for 60 min at 37°C, and then rinsed and the nuclei were stained by 4′, 6-diamidino-2-phenylindole (DAPI) (1 : 1000 dilution; Sigma)-containing PBS for 10 min at room temperature. After washing twice with PBS, the cells were examined in fresh PBS under an inverted fluorescence microscope (Olympus, Tokyo, Japan).

RT-PCR. Total RNA from germline stem cells was extracted using the Trizol reagent (Qiagen Ltd, Crawley, UK), according to the manufacturer's instructions. After extraction, 1 μg of total RNA was used to synthesize cDNA with Primescript Reverse Transcriptase (TaKaRa, Dalian, China). Primers sequences were listed in Supplmentary Table S5.

RNA isolation and microarray hybridization. RNA from germline stem cells cultured for 2 days (see above) was pooled in equal concentrations. Three respective replicates of FGSC and SSC RNA that passed quality check were hybridized to the

Mouse WG-6 v2.0 microarray, according to the manufacturer's instructions (Illumina, San Diego, CA). The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through the GEO Series accession number GSE51313.

Data analysis and gene ontology analysis. Data analysis was conducted in the R software environment for statistical computing and graphics. The Lumi package was used to process the raw data exported from GenomeStudio (Illumina)[53]. We detected 19004 probes expressed based on the cut-off p-value ($p < 0.01$), representing the whole gene expression profiles of FGSCs and SSCs. To compare the global similarity of the whole gene expression profiles between FGSCs and SSCs, gplots package in R was used for unsupervised hierarchical clustering. We set the normalized gene expression value at 8.0 as a threshold to obtain highly expressed genes. Housekeeping genes were removed from the highly expressed gene list. These comprised well-known housekeeping genes and homologs of identified human housekeeping genes[54]. The housekeeping gene list is shown in Supplementary Table S4. The profiles of overexpressed genes in mouse embryonic stem cells, neural stem cells, and hematopoietic stem cells used to compared with CEG were referenced to Miguel's work[20]. The CEG were uploaded to the database for annotation, visualization and integrated discovery (DAVID) for gene ontology (GO) enrichment[32]. Enriched functional categories ($p < 0.01$) were identified in the GO analyses of CEG.

Computational analysis of promoter cis-regulatory elements of the CEG. For promoter analysis we applied EXPANDER[26] to detect cis-regulatory promoter elements that control the expressions of the CEG shared by FGSCs and SSCs. The whole set of expressed genes was used as a background set of promoters. Bonferroni correction in EXPANDER was used to identify significant transcription factors (TFs). The promoter sequences were used for screening spanned 1000 bp upstream and 200 bp downstream of the transcriptional start sites.

Construction of the maximum continuous protein-protein interaction network of the CEG by breadth first search. According to the theory of complex networks of protein-protein interactions, we constructed the interaction network of the CEG based on I2D database. I2D version 2.0 includes 150,230 protein-protein interactions for mouse[55,56]. After converting the Entrez IDs of the CEG into the corresponding Swiss-Prot Id, we mapped the CEG to the 150,230 protein-protein interactions. We then compiled a PERL script to construct maximum continuous protein-protein interaction network of the mapped the CEG by breadth first search. We defined the hub proteins as those whose connectivity was in the top 10% of the connected nodes in the maximum continuous network. Cytoscape 3.0.2 was used to visualize and analysis the maximum continuous protein-protein interaction network[57].

1. Van der Kooy, D. & Weiss Samuel. Why Stem Cells? *Science* **287**, 1439–1441 (2000).
2. Spradling, A., Drummond-Barbosa, D. & Kai, T. Stem cells find their niche. *Nature* **414**, 98–104 (2001).
3. Kubota, H., Avarbock, M. R. & Brinster, R. L. Spermatogonial stem cells share some, but not all, phenotypic and functional characteristics with other stem cells. *Proc. Natl. Acad. Sci. USA* **100**, 6487–6492 (2003).
4. De Rooij, D. G. & Russell, L. D. All You Wanted to Know About Spermatogonia but Were Afraid to Ask. *J Androl* **21**, 776–798 (2000).
5. Rooij, D. G. & Kramer, M. F. Spermatogonial stemcell renewal in rats and mice. *Z Zellforsch Mikrosk Anat* **85**, 206–209 (1968).
6. Anderson, L. D. & Hirshfield, A. N. An overview of follicular development in the ovary: from embryo to the fertilized ovum in vitro. *Md Med J* **41**, 614–620 (1992).
7. McLaren, A. Meiosis and differentiation of mouse germ cells. *Symp Soc Exp Biol* **38**, 7–23 (1984).
8. Peters, H. Migration of Gonocytes Into the Mammalian Gonad and Their Differentiation. *Philos Trans R Soc Lond B Biol Sci* **259**, 91–101 (1970).
9. Johnson, J., Canning, J., Kaneko, T., Pru, J. K. & Tilly, J. L. Germline stem cells and follicular renewal in the postnatal mammalian ovary. *Nature* **428**, 145–150 (2004).
10. Zou, K. *et al.* Production of offspring from a germline stem cell line derived from neonatal ovaries. *Nat Cell Biol* **11**, 631–636 (2009).
11. Zou, K., Hou, L., Sun, K., Xie, W. & Wu, J. Improved Efficiency of Female Germline Stem Cell Purification Using Fragilis-Based Magnetic Bead Sorting. *Stem Cells Dev* **20**, 2197–2204 (2011).
12. Zhang, Y. *et al.* Production of transgenic mice by random recombination of targeted genes in female germline stem cells. *J Mol Cell Biol* **3**, 132–141 (2011).
13. Pacchiarotti, J. *et al.* Differentiation potential of germ line stem cells derived from the postnatal mouse ovary. *Differentiation* **79**, 159–170 (2010).
14. Hu, Y. *et al.* GSK3 inhibitor-BIO regulates proliferation of female germline stem cells from the postnatal mouse ovary. *Cell Prolif* **45**, 287–298 (2012).
15. White, Y. A. R. *et al.* Oocyte formation by mitotically active germ cells purified from ovaries of reproductive-age women. *Nat Med* **18**, 413–421 (2012).
16. Zhou, L. *et al.* Production of fat-1 transgenic rats using a germline stem cell line from a postnatal female. *Mol Hum Reprod* **20**, 271–281 (2014).
17. Zhang, H. *et al.* Experimental evidence showing that no mitotically active female germline progenitors exist in postnatal mouse ovaries. *Proc. Natl. Acad. Sci. USA* **109**, 12580–12585 (2012).
18. Lei, L. & Spradling, A. C. Female mice lack adult germ-line stem cells but sustain oogenesis using stable primordial follicles. *Proc. Natl. Acad. Sci. USA* **110**, 8585–8590 (2013).
19. Tilmann, C. & Capel, B. Cellular and Molecular Pathways Regulating Mammalian Sex Determination. *Recent Prog Horm Res* **57**, 1–18 (2002).
20. Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. & Melton, D. A. "Stemness": Transcriptional Profiling of Embryonic and Adult Stem Cells. *Science* **298**, 597–600 (2002).
21. Toyooka, Y. *et al.* Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development. *Mech Dev* **93**, 139–149 (2000).
22. Ruggiu, M. *et al.* The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. *Nature* **389**, 73–77 (1997).
23. Scholer, H. R., Ruppert, S., Suzuki, N., Chowdhury, K. & Gruss, P. New type of POU domain in germ line-specific protein Oct-4. *Nature* **344**, 435–439 (1990).
24. Tanaka, S. S., Yamaguchi, Y. L., Tsoi, B., Lickert, H. & Tam, P. P. L. IFITM/Mil/Fragilis Family Proteins IFITM1 and IFITM3 Play Distinct Roles in Mouse Primordial Germ Cell Homing and Repulsion. *Dev Cell* **9**, 745–756 (2005).
25. Sabour, D. *et al.* Identification of genes specific to mouse primordial germ cells through dynamic global gene expression. *Hum Mol Genet* **20**, 115–125 (2011).
26. Shamir, R. *et al.* EXPANDER - an integrative program suite for microarray data analysis. *BMC Bioinformatics* **6**, 232 (2005).
27. Wey, A., Cerdeno, V., Pleasure, D. & Knoepfler, P. c- and N-myc Regulate Neural Precursor Cell Fate, Cell Cycle, and Metabolism to Direct Cerebellar Development. *Cerebellum* **9**, 537–547 (2010).
28. Hofer, G., Grimmer, C., Sukhatme, V. P., Sterzel, R. B. & Rupprecht, H. D. Transcription Factor Egr-1 Regulates Glomerular Mesangial Cell Proliferation. *J Biol Chem* **271**, 28306–28310 (1996).
29. Luo, J., Megee, S. & Dobrinski, I. Asymmetric Distribution of UCH-L1 in Spermatogonia Is Associated With Maintenance and Differentiation of Spermatogonial Stem Cells. *J Cell Physiol* **220**, 460–468 (2009).
30. Greenbaum, M. P. *et al.* TEX14 is essential for intercellular bridges and fertility in male mice. *Proc. Natl. Acad. Sci. USA* **103**, 4982–4987 (2006).
31. Glaser, S. *et al.* The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. *Epigenetics Chromatin* **2**, 5 (2009).
32. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57 (2008).
33. Frost, R. J. A. *et al.* MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc. Natl. Acad. Sci. USA* **107**, 11847–11852 (2010).
34. Grisanti, L. *et al.* Inactivation of Numb and Numblike in spermatogonial stem cells by cell-permeant Cre recombinase. *Differentiation* **78**, 131–136 (2009).
35. Griffeth, R. J., Carretero, J. & Burks, D. J. Insulin Receptor Substrate 2 Is Required for Testicular Development. *PLoS ONE* **8**, e62103 (2013).
36. Will, B. *et al.* Satb1 regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. *Nat Immunol* **14**, 437–445 (2013).
37. Andreu-Agullo, C., Maurin, T., Thompson, C. B. & Lai, E. C. Ars2 maintains neural stem-cell identity through direct transcriptional activation of Sox2. *Nature* **481**, 195–198 (2012).
38. Jin, G. *et al.* Progastrin Stimulates Colonic Cell Proliferation via CCKBR- and β-Arrestin–Dependent Suppression of BMP2. *Gastroenterology* **145**, 820–830 (2013).
39. Regan Joseph, L. *et al.* Aurora A Kinase Regulates Mammary Epithelial Cell Fate by Determining Mitotic Spindle Orientation in a Notch-Dependent Manner. *Cell Rep* **4**, 110–123 (2013).
40. Zheng, Y. *et al.* Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature* **499**, 166–171 (2013).
41. Tokuda, M., Kadokawa, Y., Kurahashi, H. & Marunouchi, T. CDH1 Is a Specific Marker for Undifferentiated Spermatogonia in Mouse Testes. *Biol Reprod* **76**, 130–141 (2007).
42. Hobbs Robin, M. *et al.* Functional Antagonism between Sall4 and Plzf Defines Germline Progenitors. *Cell stem cell* **10**, 284–298 (2012).
43. Galan-Caridad, J. M. *et al.* Zfx Controls the Self-Renewal of Embryonic and Hematopoietic Stem Cells. *Cell* **129**, 345–357 (2007).
44. Harel, S. *et al.* ZFX Controls the Self-Renewal of Human Embryonic Stem Cells. *PLoS ONE* **7**, e42302 (2012).
45. Masuda, T. *et al.* Cold-inducible RNA-binding protein (Cirp) interacts with Dyrk1b/Mirk and promotes proliferation of immature male germ cells in mice. *Proc. Natl. Acad. Sci. USA* **109**, 10885–10890 (2012).
46. Wang, C.-H. *et al.* A shRNA Functional Screen Reveals Nme6 and Nme7 Are Crucial for Embryonic Stem Cell Renewal. *Stem Cells* **30**, 2199–2211 (2012).
47. Dejosez, M. *et al.* Ronin Is Essential for Embryogenesis and the Pluripotency of Mouse Embryonic Stem Cells. *Cell* **133**, 1162–1174 (2008).
48. Mao, X. *et al.* APPL1 binds to adiponectin receptors and mediates adiponectin signalling and function. *Nat Cell Biol* **8**, 516–523 (2006).
49. Wu, J. *et al.* Short-type PB-cadherin promotes self-renewal of spermatogonial stem cells via multiple signaling pathways. *Cell Signal* **20**, 1052–1060 (2008).
50. Wang, H, S. L., Xiang, J., Ding, X. B., Luo, H. C., Wang, S. Z. & Wu, J. Isolation, culture and transplantation of female germline stem cells from neonatal and prepubertal mice. *Protocol Exchange* (2013).

51. Wang, H. *et al.* Conversion of female germline stem cells from neonatal and prepubertal mice into pluripotent stem cells. *J Mol Cell Biol* **6**, 164–171 (2014).
52. Yuan, Z., Hou, R. & Wu, J. Generation of mice by transplantation of an adult spermatogonial cell line after cryopreservation. *Cell Prolif* **42**, 123–131 (2009).
53. Du, P., Kibbe, W. A. & Lin, S. M. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* **24**, 1547–1548 (2008).
54. Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends Genet* **29**, 569–574 (2013).
55. Brown, K. & Jurisica, I. Unequal evolutionary conservation of human protein interactions in interologous networks. *Genome Biol* **8**, R95 (2007).
56. Brown, K. R. & Jurisica, I. Online Predicted Human Interaction Database. *Bioinformatics* **21**, 2076–2082 (2005).
57. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* **13**, 2498–2504 (2003).

Acknowledgments

This work was supported by National Basic Research Program of China (2013CB967401 and 2010CB945001), the National Nature Science Foundation of China (81370675, 81200472 and 81121001), and Shanghai Jiao Tong University Medicine-Engineering Fund (YG2013ZD04).

Author contributions

W.X. wrote the main manuscript text, prepared figures 1–5, supplementary figures S1–4 and supplementary tables S1–5. W.X. and H.W. prepared supplementary figure S5. J.W. conceived and designed the experiments. All authors reviewed the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xie, W., Wang, H. & Wu, J. Similar morphological and molecular signatures shared by female and male germline stem cells. *Sci. Rep.* **4**, 5580; DOI:10.1038/srep05580 (2014).