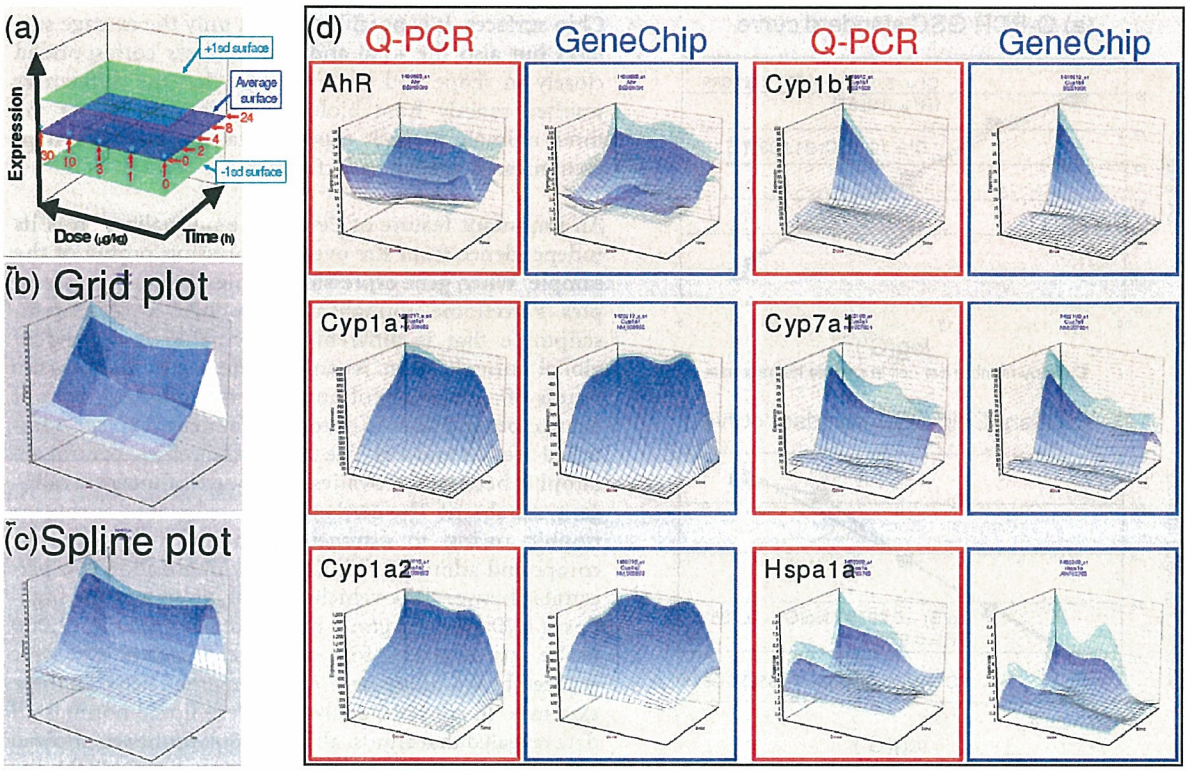

Figure 5

Correspondence between Q-PCR and GeneChip data. Sixty male C57BL/6 mice were divided into 20 groups of 3 mice each. 2,3,7,8-tetrachlorodibenzodioxin (TCDD) was administered once orally at doses of 0, 1, 3, 10 and 30μg/kg, and the liver was sampled 2, 4, 8 and 24 h after administration. The liver transcriptome was measured by the Affymetrix Mouse430-2 GeneChip. For Q-PCR, nineteen primary pairs were prepared and the Ct values of the same 60 liver samples were measured (19 genes and 5 spikes in duplicate, using a 96-well plate for 2 samples, total 30 plates). The Percellome data were plotted on to 3-dimensional graphs for average, +1sd, and – 1sd surfaces as shown in (a). The scale of expression (vertical axis) is the copy number per cell. The 0 h data (*) are copied from the 2 h/dose 0 point for better visualization of the changes after 2 h. The surfaces are demonstrated as a grid plot (b) where the grid points indicate one treatment group (n = 3), and a smoothened spline surface plot (c) for easier 3D recognition ((b), (c): Gys2 (glycogen synthase 2, 1424815_at) showing a typical circadian pattern. (d) the smoothened plots of 6 representative genes/ probe sets generated by Q-PCR (red) and GeneChip (blue). AhR (arylhydrocarbon receptor, 1450695_at) showed imperfect correspondence. Cyp1a1 (cytochrome P450, family 1, subfamily a, polypeptide 1, 1422217_a_at) and Cyp1a2 (1450715_at) showed good correlations between Q-PCR and GeneChip except for the saturation in GeneChips above c. 400 copies per cell. Cyp1b1 (1416612_at) and Cyp7a1 (1422100_at) showed good correspondence. Hspa1a (heat shock protein 1A, 1452888_at) showed fair correspondence despite low copy numbers, near the nominal detection limit of the Affymetrix GeneChip system.

The smallest sample to which we have successfully applied the direct DNA quantification method with sufficient reproducibility is the 6.75 dpc (days post coitus) mouse embryo which consists of approximately 5,000 cells. This sample size is also approximately the lower limit for double amplification protocol to obtain sufficient amount of RNA for Affymetrix GeneChip measurement (cf. http://www.affymetrix.com/Auth/support/downloads/manuals/expression_print_manual.zip.)

High-resolution technology such as laser-capture microdissection (LCM) has become popular and the average sample size analyzed is getting smaller. An alternative method for LCM samples is to count the cell number in the course of microdissection. Although we have not yet applied Percellome method to LCM samples, we have applied the alternative method to cell culture samples to gain Percellome data. Stereological and statistical calculations should become available to correct the number of partially sectioned cells in the LCM samples. Another issue for small samples is the yield of RNA. Approximately

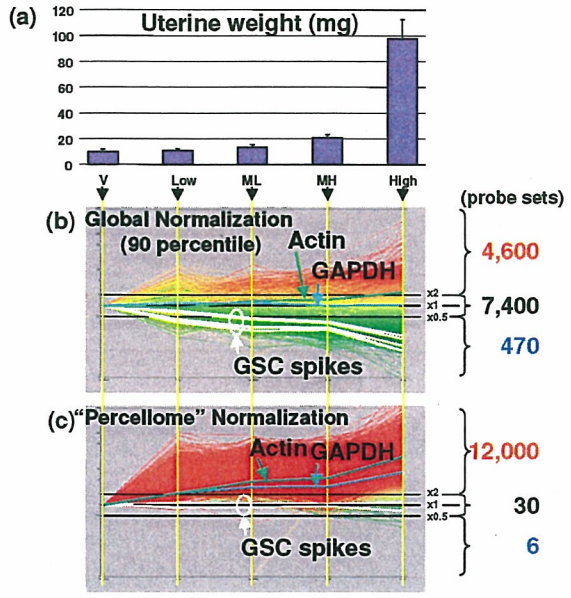

Figure 6
Uterotrophic response of ovariectomized female mice by an estrogenic test compound.(a) Shows the uterine weight, which increases in a dose-dependent manner; V, vehicle control; Low, low dose; ML, medium-low dose; MH, medium-high dose; High, high dose group. (b) Shows the line display of uterine gene expression (Affymetrix MG-U74v2 A GeneChips) normalized by global normalization (90 percentile), and (c) by the Percellome normalization. Averages of three samples per group were visualized (by K. A.). The five white lines are the GSC mRNAs. The green and blue lines are actin (AFFX-b-ActinMur/M12481_3_at) and GAPDH (glyceraldehyde-3-phosphate dehydrogenase, AFFX-GapdhMur/M32599_3_at), respectively. By global normalization, 7,400 probe sets remained unchanged and 4,600 probe sets increased more than two-fold in the H group compared to the V group, whereas almost all probe sets measured had increased. It is noted that housekeeping genes such as actin and GAPDH are significantly induced on a per cell basis.

30 ng of total RNA is retrieved from a single 6.75 dpc mouse embryo. This amount is sufficient for a double amplification protocol (DA) to prepare enough RNA for an Affymetrix GeneChip measurement. An inherent problem with the DA data is that the gene expression profile differs from that of the default single amplification protocol (SA). Consequently the DA percellome data differ from that of SA as if they were produced by a different platform. To bridge the difference, we applied the procedure that was used for data conversion between Q-PCR and GeneChip (cf. Figure 7). A set of spiked-in standard samples including the LBM sample set (of sufficient concentration) were measured by the SA protocol and diluted versions to the limit measured by the DA protocol. These data provided us with information about whether DA was successful as a whole (by comparing 5' signal to 3' signals of selected probe sets) and which probe sets were properly amplified by DA (by checking the linearity of the diluted LBM data). For those probe sets that proved to be linearly amplified, conversion functions between DA and SA were generated. These details, along with embryo expression data will be published elsewhere.

Figures 5 and 7 indicate a close correspondence between the data generated by Q-PCR and GeneChip analyses. Since each of the 60 samples was normalized individually against each GSC signal, the high similarity between the two platforms indicates the robustness and stability of this spike system (cf. Figure 7, Cyp7a1 data). Although more spikes could potentially increase the accuracy of normalization, our experience is that five spikes are practically sufficient for covering the detection range of GeneChip microarrays and Q-PCR, as long as they are used in combination with the "spike factor". The overall benefits of using a minimum number of external spikes include lower probability of cross-hybridization, a reduced number of wells and spots occupied by the spikes in the Q-PCR plates and small scale microarrays, and less effort in preparation, QC and supply.

The Percellome data can be truly absolute when all mRNA measurements including GSC spikes are strictly proportional to the original copy numbers in the sample homogenate. As noted earlier, this condition is not guaranteed by any platform despite linearity of response. Therefore, the Percellome-normalized values have some biases for each primer pair/probe set, depending on the steepness of the dose-response curves. An advantage of Percellome normalization is that, as long as such biases are consistently reproduced within a platform, the data can be compared directly among samples/studies on a common scale. Consequently, when a true value is obtained by any other measure, all the data obtained in the past can be simultaneously batch-converted to the true values.

This batch-conversion strategy can be extended to data conversion between different versions and different platforms, as long as the data are generated in copy numbers "per cell". We have shown an example between Affymetrix GeneChip and Q-PCR for limited numbers of probe sets (cf. Figure 7). Custom microarrays that accept our GSC for Percellome normalization are in preparation by Agilent Technologies (single color) and GE Healthcare (CodeLink Bioarray).

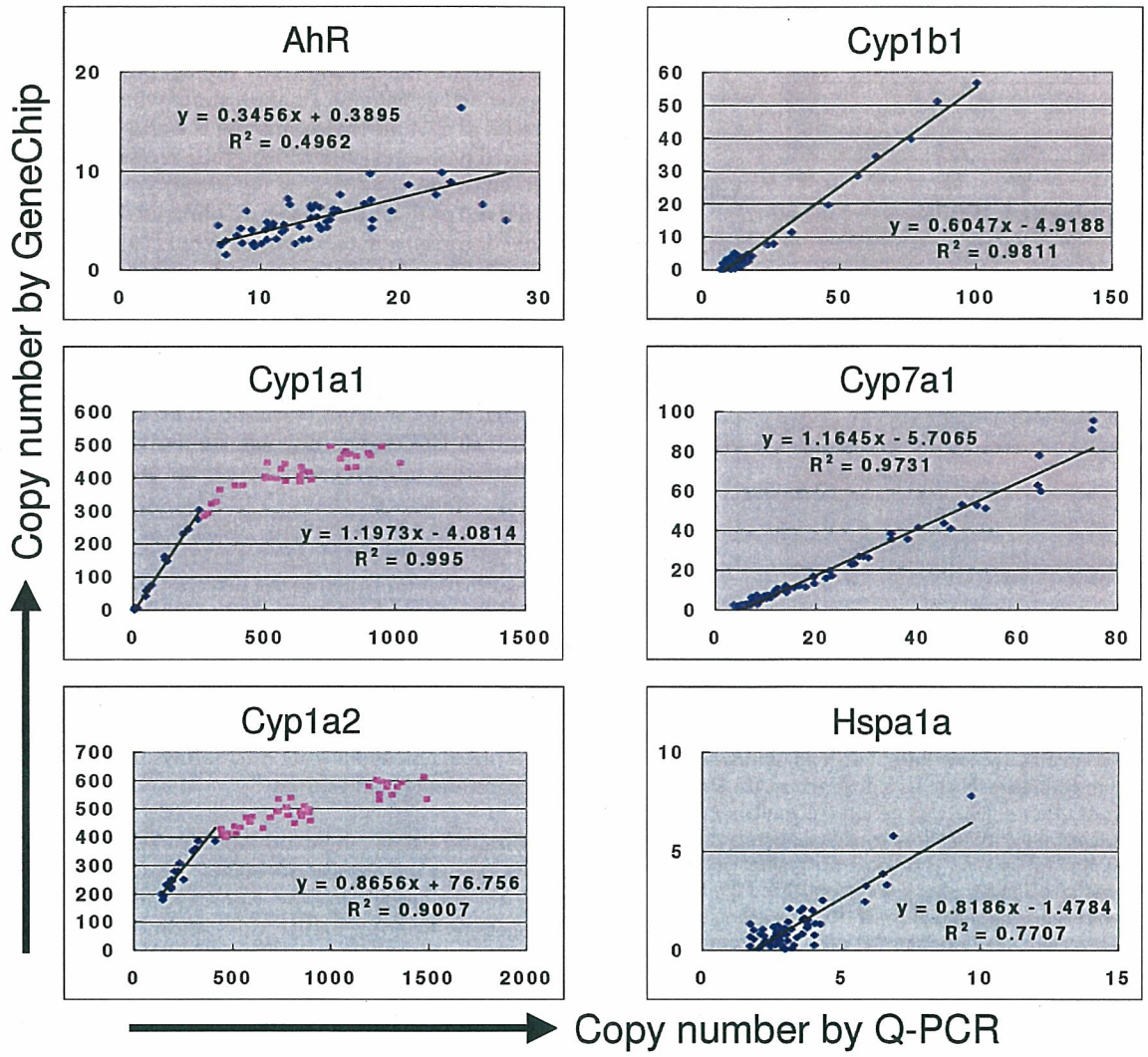

Figure 7
Conversion functions between Q-PCR and GeneChip. The data shown in Figure 5 as 3D surfaces are shown as a scatter plot (60 plots). The regression function can be used to convert Q-PCR to GeneChip and vice versa, with a level of certainty indicated by coefficient of correlation. It is noted that Cyp1a1 and Cyp1a2 became saturated above about 400 copies per cell in GeneChip system (indicated in pink plots). Cyp7a1 showed high linearity, indicating that the variation shown by the split +1sd and -1sd surfaces in Figure 5 reflected biological (animal) variation, not measurement errors.

Another important contribution of Percellome analysis is in the area of archived data in private and public domains. Firstly, Percellome data are the result of a simple linear transformation of the raw microarray data; preserving the distribution and order of the probe set data. Therefore, parametric or non-parametric methods should be able to align the data distribution and generate estimates of mRNA copy number of the non-spiked archival samples. Any archival samples that are re-measurable by Percellome method will greatly increase the accuracy of estimation. Secondly, percellome can provide appropriate bridging information between old and new versions of Affymetrix GeneChips, such as human HU-95 and HU-133, murine MU-74v2 and MOE430 series. This should also facilitate comparisons between newly generated and archived data.

The Percellome method was developed for a large-scale toxicogenomics project [13] using the Affymetrix GeneChip system. It was intended to compile a very large-scale database of comprehensive gene expression profiles in response to various chemicals from a series of experiments conducted over an extended time period. However, the method also proved to be useful for small-scale platforms such as 96 well plate-based Q-PCRs as shown above, and probably for small-scale targeted microarrays. In both cases, highly inducible or highly transcribed genes are likely to be selected. Therefore, the expression profiles may differ significantly among samples such that profile-dependent normalization (e.g. global normalization) may not be applicable. In such cases, the profile-independent nature of the Percellome method provides a robust normalization.

To demonstrate the profile-independence of the Percellome method, we chose an extreme case – the uterotrophic response assay (cf. Figure 6). The treated uteri were composed of hypertrophic cells with abundant cytoplasm whereas the untreated uteri were composed of hypoplastic cells with scant cytoplasm. This indicates that the uteri of untreated ovariectomized mice were quiescent, and that a majority of the inducible genes were probably transcriptionally inactive. Therefore, the identification of most genes as being induced by 2-fold or greater is reasonable and expected. In most *in vivo* experiments, the gene profiles of the samples are much more similar. However, there is always a set of genes that is found to be "increased" when analyzed on a "per one cell" basis that are declared to be "decreased" by global type normalization, or vice versa. Such increase/decrease calls made by the global type normalization can differ according to the normalization parameters. In both cases, the Percellome method can inform the researcher how much the expression profiles are distorted by the treatment, such as in the case of the uterotrophic assay. We also note that *in vitro* experiments such as cell-based studies tend to generate data similar to that of uterotrophic experiment.

Conclusion

Percellome data can be compared directly among samples and among different studies, and between different platforms, without further normalization. Therefore, "percellome" normalization can serve as a standard method for exchanging and comparing data across different platforms and among different laboratories. We hope that the Percellome method will contribute to transcriptome-based studies by facilitating data exchanges among laboratories.

Methods

Animal experiments

C57BL/6 Cr Slc (SLC, Hamamatsu, Japan) mice maintained in a barrier system with a 12 h photoperiod were used in this study. For the liver transcriptome experiments, twelve week-old male mice were given a single dose of the test compound by oral gavage, and the liver was sampled at 2, 4, 8 and 24 h post-gavage. For the uterotrophic experiment, 6 week old female mice were ovariectomized 14 days prior to the 7 day repeated subcutaneous injection of a test compound [12]. Animals were euthanized by exsanguination under ether anesthesia and the target organs were excised into ice-cooled plastic dishes. Tissue blocks weighing 30 to 60 mg were placed in an RNase-free 2 ml plastic tube (Eppendorf GmbH., Germany) and soaked in RNAlater (Ambion Inc., TX) within 3 min of the beginning of anesthesia. Three animals per treatment group were used and individually subjected to transcriptome measurement.

Sample homogenate preparation

The tissue blocks soaked in RNAlater were kept overnight at 4°C or until use. RNAlater was replaced in the 2 ml plastic tube with 1.0 ml of RLT buffer (Qiagen GmbH., Germany), and the tissue was homogenized by adding a 5 mm diameter Zirconium bead (Funakoshi, Japan) and shaking with a MixerMill 300 (Qiagen GmbH., Germany) at a speed of 20 Hz for 5 min (only the outermost row of the shaker box was used).

Direct DNA quantitation

Three separate 10 μl aliquots were taken from each sample homogenate to another tube and mixed thoroughly. A final 10 μl aliquot therefrom was treated with DNAse-free RNase A (Nippon Gene Inc., Japan) for 30 min at 37°C, followed by Proteinase K (Roche Diagnostics GmbH., Germany) for 3 h at 55°C in 1.5 ml capped tubes. The aliquot was transferred to a 96-well black plate. PicoGreen fluorescent dye (Molecular Probes Inc., USA) was added to each well, shaken for 10 seconds four times and then incubated for 2 min at 30°C. The DNA concentration was measured using a 96 well fluorescence plate reader with excitation at 485 nm and emission at 538 nm. λ phage DNA (PicoGreen Kit, Molecular Probes Inc., USA) was used as standard. Measurement by this PicoGreen method and the standard phenol extraction method correlated well (coefficient of correlation = 0.97, data not shown). The smallest sample size for reproducible and reliable DNA quantitation is about 5,000 cells that corresponds to a 6.75 dpc mouse embryo.

The grade-dosed spike cocktail (GSC)

The following five *Bacillus subtilis* RNA sequences were selected from the gene list of Affymetrix GeneChip arrays (AFFX-ThrX-3_at, AFFX-LysX-3_at, AFFX-PheX-3_at, AFFX-DapX-3_at, and AFFX-TrpnX-3_at) present in the MG-U74v2, RG-U34, HG-U95, HG-U133, RAE230 and MOE430 arrays: thrC, thrB genes corresponding to nucleotides 248–2229 of X04603; lys gene for diami-

nopimelate decarboxylase corresponding to nucleotides 350–1345 of X17013; pheB, pheA genes corresponding to nucleotides 2017–3334 of M24537, dapB, jojF, jojG genes corresponding to nucleotides 1358–3197 of L38424; TrpE protein, TrpD protein, TrpC protein corresponding to nucleotides 1883–4400 of K01391. The corresponding cDNAs were purchased from ATCC, incorporated into expression vectors, amplified in *E. coli* and transcribed using the MEGAscript kit (Ambion Inc., TX). The mRNA was purified using a MACS mRNA isolation kit (Miltenyi Biotec GmbH., Germany). The concentrations of spike RNAs in the GSC were in threefold steps, from 777.6 pM for AFFX-ThrX-3_at, 259.4 pM for AFFX-LysX-3_at, 86.4 pM for AFFX-PheX-3_at, 28.8 pM for AFFX-DapX-3_at, to 9.6 pM for AFFX-TrpnX-3_at. In general, the ratio depends on the linear range of the measurement system and the available number of spikes.

Setting of the "spike factor" and addition of GSC to a sample homogenate according to its DNA concentration

The GSC was added to the sample homogenates in proportion to their DNA concentrations, assuming that all cells contain a fixed amount of genomic DNA (g/cell) across samples. The amount of GSC added to each sample G (l) was given as

$$G = C * v * f \quad (l),$$

where C is the DNA concentration (g/l), v(l) is the volume of homogenate further used for RNA extraction and f (l/g) is the "spike factor", which is an adjustment factor to ensure that the sample is properly spiked by the GSC (cf. Figure 3). Spike factors have been pre-determined for various organs/tissues to reflect differences in their total RNA/genomic DNA ratios (cf. Table 1). In this way, five spike mRNA signals can properly cover the linear dose-response range of the platform. In practice, for the Affymetrix GeneChips, the spike factor is set so that the five GSC spikes cover the range of "Present" calls given by the Affymetrix system, which corresponds to approximately 80 to 7000 in raw readouts given by the Affymetrix MAS5.0 software. A raw readout of 10 by the current Affymetrix GeneChip system corresponds to approximately one copy per cell in mouse liver (spike factor = 0.2), whereas in mouse thymus (spike factor = 0.01) it corresponds to approximately 0.05 copy per cell. For Q-PCR, the same spike factor corresponds to Ct values ranging approximately from 17 to 27, which is well within the linear range of Q-PCR (data not shown).

"Per cell" normalization (Percellome normalization)

Since murine haploid genomic DNA is made of 2.5×10^9 base pairs and one base pair is approximately 600 Daltons (Da), the haploid genomic DNA weighs 1.5×10^{12} Da, corresponding to

$d = 5 \times 10^{-12}$ (g DNA per diploid cell).

Therefore, the cell number per liter of the sample homogenate (N) is given as

$$N = C/d \text{ (cells/l)}$$

where C is the DNA concentration (g/l).

On the other hand, the copy numbers of GSC RNAs in the homogenate are given as follows:

if Sj (mole/l) (j = 1,2,3,4,5) is the mole concentration of one of the five spike RNAs in the GSC solution and G(l) is the amount of GSC added to each homogenate, the mole concentrations of the spike RNAs in the homogenate (CSj) are given as,

$$CSj = Sj*C*f \text{ (mole/l)}.$$

The GSC RNAs in moles per cell (MSj) are given as

$$MSj = CSj/N$$

$$= Sj*C*f/(C/d)$$

$$= Sj*f*d \text{ (mole/cell)}$$

The copy numbers of the GSC RNAs per cell (NSj) are given as

$$NSj = MSj*A$$

$$= Sj*f*d*A \text{ (copies per diploid cell)}$$

where A is Avogadro's number.

As a result, the GSC spikes AFFX-TrpnX-3_at, AFFX-DapX-3_at, AFFX-PheX-3_at, AFFX-LysX-3_at and AFFX-ThrX-3_at correspond approximately to 5.8, 17.3, 52.0, 156.0 and 468.1 copies per cell (per diploid DNA template) for mouse liver sample homogenates, where the spike factor = 0.2. It is our observation that the RNA/DNA ratios are virtually constant across polyploid hepatocytes (data not shown).

For each Q-PCR plate or GeneChip, the coefficients, α, β, γ and δ of functions {1} or {2} are determined from the GSC values using the least-square method. The signal values or Ct values of all the other mRNAs measured are then converted to copy numbers per cell by {3} or {4}, i.e. the inverses of functions {1} or {2}.

Table 2: Primers for Q-PCR

Gene	Forward	Reverse
AFFX-TrpnX-3_at	TTCTCAGCGTAAAGCAATCCA	GCAAATCCTTTAGTGACCGAATACC
AFFX-DapX-3_at	TCAGCTAACGCTTCCAGACC	GGCCGACAGATTCTGATGACA
AFFX-PheX-3_at	GCCAATGATATGGCAGCTTCTAC	TGCGGCAGCATGACCATTA
AFFX-LysX-3_at	CCGCTTCATGCCACTGAATAC	CCGGTTCGATCCAAATTTCC
AFFX-ThrX-3_at	CCTGCATGAGGATGACGAGA	GGCATCGGCATATGGAAAC
Ahr_1450695_at	CAGAGACCACTGACGGATGAA	AGCCTCTCCGGTAGCAAACA
Cypla1_1422217_a_at	TGCTCTTGCCACCTGCTGA	GGAGCACCCTGTTTGTTTCTATG
Cypla2_1450715_at	CCTCACTGAATGGCTTCCAC	CGATGGCCGAGTTGTTATTG
Cyplb1_1416612_at	GCCTCAGGTGTGTTTGATGGA	AGTACAGCCCTGGTGGGAATG
Cyp7a1_1422100_at	TTCTACATGCCCTTTGGATCAG	GGACACTTGGTGTGGCTCTC
Hspa1a_1452388_at	ACCATCGAGGAGGTGGATTAGA	AGGACTTGATTGCAGGACAAAC

The "LBM" ("liver-brain mix") standard sample

A pair of samples having dissimilar gene expression profiles was chosen to evaluate the linearity of the platform. The pairs chosen were brain and liver for mouse and rat, two distinct cancer cell lines for humans, and adult liver and embryo for *Xenopus laevis*. The sample pairs were processed as described above including addition of the GSC. Two final homogenates were then blended at ratios of 100:0, 75:25, 50:50, 25:75 and 0:100 (based on cell numbers) to make five samples. These five samples were measured by Q-PCR and/or GeneChips (MG-U74v2A, MEA430A, MEA430B, MG430 2.0 (shown in Figure 1), RAE230A, HG-U95A, HG-U133, and Xenopus array).

Quantitative-PCR

Duplicate homogenate samples were treated with DNaseI (amplification grade, Invitrogen Corp., Carlsbad, CA, USA) for 15 min at room temperature, followed by SuperScript II (Invitrogen) for 50 min at 42°C for reverse transcription. Quantitative real time PCR was performed with an ABI PRISM 7900 HT sequence detection system (Applied Biosystems, Foster City, CA, USA) using SYBR Premix Ex Taq (TAKARA BIO Inc., Japan), with initial denaturation at 95°C for 10 s followed by 45 cycles of 5 s at 95°C and 60 s at 60°C, and Ct values were obtained. Primers for the genes explored in this study were selected from sequences close to the areas of Affymetrix GeneChip probe sets as shown in Table 2.

Affymetrix GeneChip measurement

The sample homogenates with GSC added were processed by the Affymetrix Standard protocol. The GeneChips used were MG-U74v2A for the uterotrophic study and Mouse 430-2 for the TCDD study (singlet measurement). The efficiency of *in vitro* transcription (IVT) was monitored by comparing the values of 5' probe sets and 3' probe sets of the control RNAs (AFFX- probe sets) including the GSC (see Quality Control below). The dose-response linearity of the five GSC spikes was checked and samples showing saturation and/or high background were re-measured from either backup tissue samples, an aliquot of homogenate, or a hybridization solution, depending on the nature of the anomaly.

Quality control

Any external spiking method, including our Percellome method, is valid for high-quality RNA samples. Therefore, the quality of the sample RNA should be carefully monitored. In addition to a common checkup by RNA electrophoresis (including capillary electrophoresis if necessary), OD ratio, and cRNA yield, we monitor the performance of IVT (*in vitro* translation) or amplification. The 3' and 5' probe set data of the spiked-in RNAs and sample RNAs (actin, GAPD and other AFFX- probe sets) that are prepared in Affymetrix GeneChip are compared to monitor the extension of RNA by the IVT process. When both the spiked-in RNAs and the sample RNAs have similar levels of 5' and 3' signals respectively, it is judged that the IVT extension was normally performed. When both spiked-in and sample RNAs have significantly lower 5' signal than 3' signal, it is judged that the IVT extension was abnormal. When only the sample RNAs showed significantly lower 5' signal than 3' signal, it is judged that the IVT extension was normal but the sample RNAs were degraded. When only the spiked-in RNAs showed significantly lower 5' signal than 3' signal, it is judged that the IVT extension was normal but the spiked-in RNAs were degraded (although we have not encountered this situation). In addition, if the degraded sample was spiked-in by the non-degraded spike RNAs and measured by GeneChip, the position of spiked-in RNAs will be offset toward abnormally higher intensity. Together, this battery of checkups considerably increases the ability to detect abnormal events that will affect the reliability of the Percellome method. When any abnormality was found, each step of sample preparation was reevaluated to regain normal data for Percellome normalization.

The web site for GeneChip data

The GeneChip data are accessible at http://www.nihs.go.jp/tox/TTG_Archive.htm.

Authors' contributions

JK drafted the concept of the Percellome method, led the project at a practical level, and drafted the manuscript. KA developed the algorithm for the Percellome calculation and wrote the calculation/visualization programs. KI developed the laboratory protocols for the Percellome procedures to the level of SOP for technicians. NN developed the Percellome Q-PCR protocol and performed the measurements, and helped in analyzing the Percellome data. AO helped develop the algorithm. YK led the animal studies. TN provided advice and led the toxicogenomics project using the Percellome method, to be approved by the Ministry of Health, Labour and Welfare of Japan.

Additional material

Additional File 1

Excel spreadsheet file containing 15 Affymetrix Mouse 430–2 GeneChip raw data of five LBM samples in triplicate (cf. Figure 1). The column name LBM-100-0-X_Signal indicates the component percentages, i.e. 100% liver 0% brain, and X = 1,2,3 indicates the triplicates. The LBM-100-0-X_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S1.zip]

Additional File 2

Excel spreadsheet file containing Percellome data of the same LBM samples, of which raw data is listed in Additional file 1 (cf. Figure 1).

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S2.zip]

Additional File 3

Excel spreadsheet file containing 2 Affymetrix MG-U74v2 raw data of a blank sample with the GSC (horizontal axis of Figure 2a) and blank with the five spike RNAs at a high dosage (vertical axis of Figure 2a).

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S3.zip]

Additional File 4

Excel spreadsheet file containing 2 Affymetrix MG-U74v2 raw data of a liver sample with GSC (horizontal axis of Figure 2b) and without GSC (vertical axis of Figure 2b).

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S4.zip]

Additional File 5

(first quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 2 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S5.zip]

Additional File 6

(second quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 4 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S6.zip]

Additional File 7

(third quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 8 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S7.zip]

Additional File 8

(last quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 24 hr data (15 GeneChip data) of the total of 60 Affymetrix Mouse 430-2 GeneChip raw data of the TCDD study consisting of 20 different treatment groups in triplicate (cf. Figure 5). The column name DoseXXX-TimeYY-Z_Signal indicates the dosage and sampling time after TCDD administration in hours, e.g. XXX = 001 indicates 1 microgram/kg group, YY = 02 indicates two hours after administration, and Z = 1,2,3 indicates animal triplicate. The DoseXXX-TimeYY-Z_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S8.zip]

Additional File 9

(first quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 2 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 5.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S9.zip]

Additional File 10

(second quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 4 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 6.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S10.zip]

Additional File 11

(third quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 8 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 7.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S11.zip]

Additional File 12

(last quarter of a data set consisting of 2 hr, 4 hr, 8 hr, and 24 hr data, divided because of the upload file size limitation)]: an Excel spreadsheet file containing 24 hr Percellome data (15 sample data) of the 60 samples of the TCDD study (cf. Figure 5), of which corresponding raw data is listed in Additional file 8.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S12.zip]

Additional File 13

Excel spreadsheet file containing 15 Affymetrix MG-U74v2 A GeneChip raw data of the uterotrophic response study (cf. Figure 6). The column name X-Y_Signal indicates the treatment (V = vehicle, Low = low dose, etc) and animal triplicate (Y = 1,2,3). The X-Y_Detection column indicates P for present, A for absent and M for marginal calls by Affymetrix MAS 5.0 system.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S13.zip]

Additional File 14

Excel spreadsheet file containing Percellome data of the same 15 samples of the uterotrophic response study (cf. Figure 6), of which raw data is listed in Additional file 13.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2164-7-64-S14.zip]

Acknowledgements

The authors thank Tomoko Ando, Noriko Moriyama, Yuko Kondo, Yuko Nakamura, Maki Abe, Nae Matsuda, Kenta Yoshiki, Ayako Imai, Koichi Morita, Hisako Aihara and Chiyuri Aoyagi for technical support, and Dr. Bruce Blumberg and Dr. Thomas Knudson for critical reading of the manuscript. This study was supported by Health Sciences Research Grants H13-Seikatsu-012, H13-Seikatsu-013, H14-Toxico-001 and H15-Kagaku-002 from the Ministry of Health, Labour and Welfare, Japan.

References

1. Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA: **Dissecting the regulatory circuitry of a eukaryotic genome.** *Cell* 1998, **95**:717-728.
2. Hill AA, Brown EL, Whitley MZ, Tucker-Kellogg G, Hunter CP, Slonim DK: **Evaluation of normalization procedures for oligonucleotide array data based on spiked cRNA controls.** *Genome Biol* 2001, **2**:. RESEARCH0055
3. Lee PD, Sladek R, Greenwood CM, Hudson TJ: **Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies.** *Genome Res* 2002, **12**:292-297.
4. van de Peppel J, Kemmeren P, van Bakel H, Radonjic M, van Leenen D, Holstege FC: **Monitoring global messenger RNA changes in externally controlled microarray experiments.** *EMBO Rep* 2003, **4**:387-393.
5. Yang YH, Dudoit S, Luu P, Lin DM, Peng W, Ngai J, Speed TP: **Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation.** *Nucleic Acids Res* 2002, **30**:e15.
6. Hekstra D, Taussig AR, Magnasco M, Naef F: **Absolute mRNA concentrations from sequence-specific calibration of oligonucleotide arrays.** *Nucleic Acids Res* 2003, **31**:1962-1968.
7. Sterrenburg E, Turk R, Boer JM, van Ommen GB, den Dunnen JT: **A common reference for cDNA microarray hybridizations.** *Nucleic Acids Res* 2002, **30**:e116.
8. Dudley AM, Aach J, Steffen MA, Church GM: **Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range.** *Proc Natl Acad Sci USA* 2002, **99**:7554-7559.
9. Talaat AM, Howard ST, Hale W, Lyons R, Gamer H, Johnston ST: **Genomic DNA standards for gene expression profiling in Mycobacterium tuberculosis.** *Nucleic Acids Res* 2002, **30**:e104.
10. Bolstad BM, Irizarry RA, Astrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.** *Bioinformatics* 2003, **19**:185-193.
11. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL: **Expression monitoring by hybridization to high-density oligonucleotide arrays.** *Nat-Biotechnol* 1996, **14**:1675-1680.
12. Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, Owens W: **The OECD program to validate the rat uterotrophic bioassay. Phase 2: dose-response studies.** *Environ Health Perspect* 2003, **111**:1530-1549.
13. Kanno J: **Reverse toxicology as a future predictive toxicology.** In *Toxicogenomics* Edited by: Inoue T, Pennie ED. Tokyo, Springer-Verlag; 2002:213-218.

Premature ovarian failure in androgen receptor-deficient mice

Hiroko Shiina*†‡, Takahiro Matsumoto*‡§, Takashi Sato*, Katsuhide Igarashi¶, Junko Miyamoto*, Sayuri Takemasa*, Matomo Sakari*§, Ichiro Takada*, Takashi Nakamura*§, Daniel Metzger‖, Pierre Chambon‖, Jun Kanno¶, Hiroyuki Yoshikawa†, and Shigeaki Kato*§**

*Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan; §Exploratory Research for Advanced Technology, Japan Science and Technology, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan; †Department of Obstetrics and Gynecology, Institute of Clinical Medicine, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8575, Japan; ¶Division of Cellular and Molecular Toxicology, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan; and ‖Institut de Genetique et de Biologie Moleculaire et Cellulaire, Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, Universite Louis Pasteur, College de France, 67404 Illkirch, Strasbourg, France

Edited by Bert W. O'Malley, Baylor College of Medicine, Houston, TX, and approved November 10, 2005 (received for review August 5, 2005)

Premature ovarian failure (POF) syndrome, an early decline of ovarian function in women, is frequently associated with X chromosome abnormalities ranging from various Xq deletions to complete loss of one of the X chromosomes. However, the genetic locus responsible for the POF remains unknown, and no candidate gene has been identified. Using the Cre/LoxP system, we have disrupted the mouse X chromosome androgen receptor (*Ar*) gene. Female $AR^{-/-}$ mice appeared normal but developed the POF phenotype with aberrant ovarian gene expression. Eight-week-old female $AR^{-/-}$ mice are fertile, but they have lower follicle numbers and impaired mammary development, and they produce only half of the normal number of pups per litter. Forty-week-old $AR^{-/-}$ mice are infertile because of complete loss of follicles. Genome-wide microarray analysis of mRNA from $AR^{-/-}$ ovaries revealed that a number of major regulators of folliculogenesis were under transcriptional control by AR. Our findings suggest that AR function is required for normal female reproduction, particularly folliculogenesis, and that AR is a potential therapeutic target in POF syndrome.

male hormone | nuclear receptor | female physiology | folliculogenesis | kit ligand

Premature ovarian failure (POF) is defined as an early decline of ovarian function after seemingly normal folliculogenesis (1). Genetic causes of POF have been frequently associated with X chromosome abnormalities (1, 2). Complete loss of one of the X chromosomes, as in Turner syndrome, and various Xq deletions are commonly identified as a cause of POF. However, responsible X-linked genes and their downstream targets have not been identified so far.

The androgen receptor (*Ar*) gene, which is the only sex hormone receptor gene on the X chromosome, is well known to be essential not only for the male reproductive system, but also for male physiology. In contrast, androgens are considered as male hormones; therefore, little is known about androgens' actions in female physiology, although AR expression in growing follicles has been described (3). However, because excessive androgen production in polycystic ovary syndrome causes infertility with abnormal menstrual cycles (4, 5), it is possible that AR-mediated androgen signaling also plays an important physiological role in the female reproductive system. Recently, using Cre/LoxP system, we generated an AR-null mutant mouse line (6) and demonstrated that inactivation of AR resulted in arrest of testicular development and spermatogenesis, impaired brain masculinization, high-turnover osteopenia, and late onset of obesity in males (7–9). At the same time, no overt physical or growth abnormalities were observed in female $AR^{-/-}$ mice. Therefore, to further examine potential role of AR in female physiology, we characterized female reproductive system in $AR^{-/-}$ females. Herein we show that female $AR^{-/-}$ mice develop the POF phenotype. At 3 weeks of age, $AR^{-/-}$ females had apparently normal ovaries with numbers of follicles similar to those in the wild-type females. However, thereafter the number of healthy follicles in the $AR^{-/-}$ ovary gradually declined, with a marked increase of atretic follicles, and by 40 weeks $AR^{-/-}$ mice became infertile, with no follicle detectable in the ovary. Reflecting this age-dependent progression in ovarian abnormality, several genes known to be involved in the oocyte–granulosa cell regulatory loop were identified by microarray analysis as AR downstream target genes. These findings clearly demonstrate that AR-mediated androgen signaling is indispensable for the maintenance of folliculogenesis and implicate impaired androgen signaling as a potential cause of the POF syndrome.

Materials and Methods

Generation of AR Knockout Mice. *AR* genomic clones were isolated from a TT2 embryonic stem cell genomic library by using human *AR* A/B domain cDNA as a probe (6). The targeting vector consisted of a 7.6-kb 5′ region containing exon 1, a 1.3-kb 3′ homologous region, a single loxP site, and a neo cassette with two loxP sites (10). Targeted clones (FB-18 and FC-61) were aggregated with single eight-cell embryos from CD-1 mice (11, 12). Floxed *AR* mice (C57BL/6) were then crossed with CMV-Cre transgenic mice (6). The two lines exhibited the same phenotypic abnormalities. The chromosomal sex of each pup was determined by genomic PCR amplification of the Y chromosome *Sry* gene (13).

Western Blot Analysis. To detect AR protein expression, ovarian cell lysates were separated by SDS/PAGE and transferred onto nitrocellulose membranes (14). Membranes were probed with polyclonal AR antibodies (N-20; Santa Cruz Biotechnology), and blots were visualized by using peroxidase-conjugated second antibody and an ECL detection kit (Amersham Pharmacia Biosciences).

Morphologic Classification of Growing Follicles. Sections were taken at intervals of 30 μm, and 6-μm paraffin-embedded sections were mounted on slides. Routine hematoxylin and eosin staining was performed for histologic examination by light microscopy. Follicle numbers in 12 sections per ovary were evaluated as primary follicles (oocyte surrounded by a single layer of cuboidal granulosa cells), preantral follicles (oocyte surrounded by two or

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: AR, androgen receptor; DHT, 5α-dihydrotestosterone; POF, premature ovarian failure.

†H.S. and T.M. contributed equally to this work.

**To whom correspondence should be addressed. E-mail: uskato@mail.ecc.u-tokyo.ac.jp.

www.pnas.org/cgi/doi/10.1073/pnas.0506736102

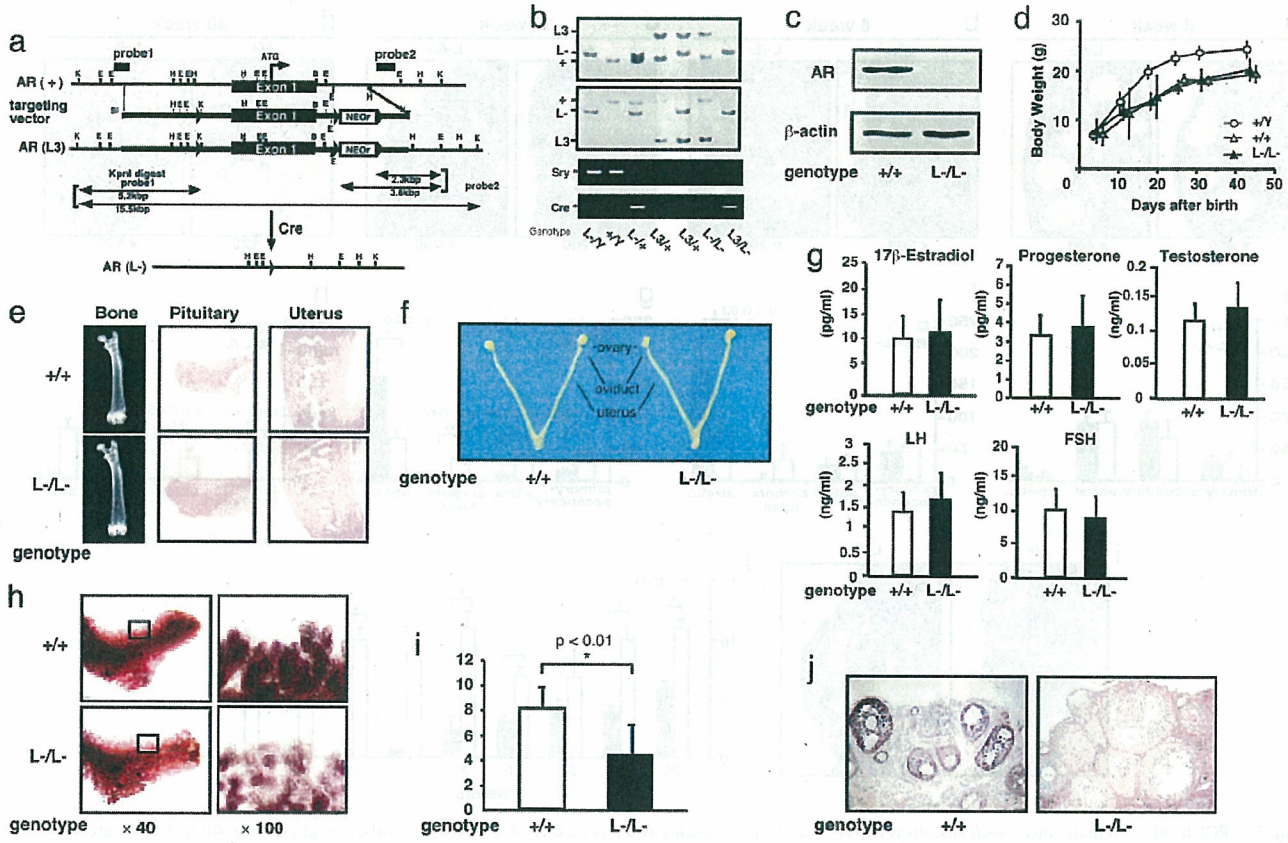

Fig. 1. Phenotypic characterization of AR knockout female mice. (*a*) Diagram of the wild-type *Ar* genomic locus (+), floxed *AR* L3 allele (L3), and *AR* allele (L-) obtained after Cre-mediated excision of exon 1. K, KpnI; E, EcoRI; H, HindIII; B, BamHI. LoxP sites are indicated by arrowheads. The targeting vector consisted of a 7.6-kb 5′ homologous region containing exon 1, a 1.3-kb 3′ homologous region, a single loxP site, and the neo cassette with two loxP sites. (*b*) Detection of the Y chromosome-specific *Sry* gene in $AR^{-/Y}$ mice by PCR. (*c*) Absence of AR protein in $AR^{-/-}$ mice ovaries by Western blot analysis using a specific C-terminal antibody. (*d*) Normal weight gain in $AR^{-/-}$ females. (*e*) Histology of pituitary, uterus, and bone tissues in $AR^{+/+}$ and $AR^{-/-}$ females at 8 weeks of age. (*f*) Female reproductive organs were macroscopically normal in $AR^{-/-}$ mice. (*g*) Serum hormone levels at the proestrus stage in $AR^{-/-}$ mice were not significantly altered. Serum 17β-estradiol, progesterone, testosterone, luteinizing hormone (LH), and follicle-stimulating hormone (FSH) levels in $AR^{+/+}$ ($n = 13$) and $AR^{-/-}$ ($n = 10$) females at 8–10 weeks of age are shown. (*h*) Lobuloalveolar development is impaired in $AR^{-/-}$ mammary glands. Whole mount of inguinal mammary glands (*Left*) and its higher magnification (*Right*) were prepared on day 3 of lactation. (*i*) Average number of pups per litter is markedly reduced in $AR^{-/-}$ mice at 8 weeks of age. Data are shown as mean ± SEM and analyzed by using Student's *t* test. (*j*) AR immunocytochemistry in $AR^{+/+}$ and $AR^{-/-}$ ovaries. Sections were counterstained with eosin.

more layers of granulosa cells with no antrum), or antral follicles (antrum within the granulosa cell layers enclosing the oocyte). Follicles were determined to be atretic if they displayed two or more of the following criteria within a single cross section: more than two pyknotic nuclei, granulosa cells within the antral cavity, granulosa cells pulling away from the basement membrane, or uneven granulosa cell layers (15).

Immunohistochemistry. Sections were subjected to a microwave antigen retrieval technique by boiling in 10 mM citrate buffer (pH 6.0) in a microwave oven for 30 min (16). The cooled sections were incubated in 1% H_2O_2 for 30 min to quench endogenous peroxidase and then incubated with 1% Triton X-100 in PBS for 10 min. To block nonspecific antibody binding, sections were incubated in normal goat serum for 1 h at 4°C. Sections were then incubated with anti-AR (1:100) or anti-cleaved caspase-3 (1:100) in 3% BSA overnight at 4°C. Negative controls were incubated in 3% BSA without primary antibody. The ABC method was used to visualize signals according to the manufacturer's instructions. Sections were incubated in biotinylated goat anti-rabbit IgG (1:200 dilution) for 2 h at room temperature, washed with PBS, and incubated in avidin–biotin–horseradish peroxidase for 1 h. After thorough washing in PBS, sections were developed with 3,3′-diaminobendizine tetrahydrochloride substrate, slightly counterstained with eosin, dehydrated through an ethanol series and xylene, and mounted.

Estrus Cycles and Fertility Test. To determine the stage of the estrus cycle (proestrus, estrus, and diestrus), vaginal smears were taken every morning and stained with Giemsa solution. For evaluation of female fertility for 15 weeks, an 8- or 24-week-old wild-type or $AR^{-/-}$ female was mated with a wild-type fertile male, replaced every 2 weeks with the other fertile male. Cages were monitored daily and for an additional 23 days, and the presence of seminal plugs and number of litters were recorded.

RNA Extraction and Quantitative Competitive RT-PCR. Total ovarian RNA was extracted by using TRIzol (Invitrogen) (16). Oligo-dT-primed cDNA was synthesized from 1 μg of ovarian RNA by using SuperScript reverse transcriptase (Gibco BRL, Gaithersburg, MD) in a 20-μl reaction volume, 1 μl of which was then diluted serially (2- to 128-fold) and used to PCR-amplify an internal control gene, *cycA*, to allow concentration estimation.

PHYSIOLOGY

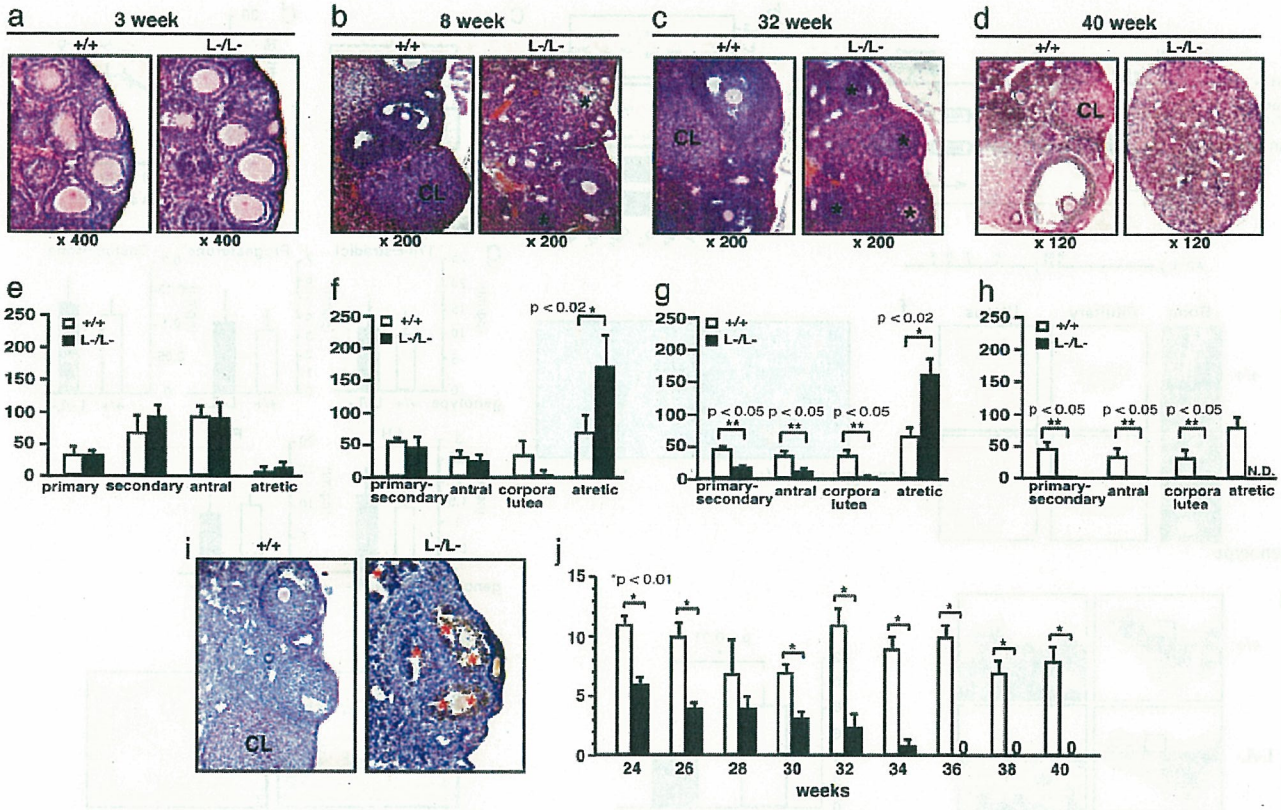

Fig. 2. POF in $AR^{-/-}$ female mice. (*a–d*) Histology of $AR^{+/+}$ and $AR^{-/-}$ ovaries at 3 weeks, 8 weeks, 32 weeks, and 40 weeks of age. All sections were stained with hematoxylin and eosin. An asterisk marks the atretic follicle. CL, corpus luteum. (*e–h*) Relative follicle counts at 3 weeks (*e*), 8 weeks (*f*), 32 weeks (*g*), and 40 weeks (*h*) of age. Numbers represent total counts of every fifth section from serially sectioned ovaries ($n = 4$ animals per genotype). (*i*) Immunohistochemical study for activated, cleaved caspase-3 revealed increased positive cells (apoptotic cells) in $AR^{-/-}$ ovaries. Sections were counterstained with hematoxylin. An asterisk marks the caspase-3-positive cell. CL, corpus luteum. (*j*) Age-dependent reduction in the number of pups per litter in $AR^{-/-}$ female mice. A continuous breeding assay was started at 24 weeks of age ($n = 6–10$ animals per genotype). For all panels, data are shown as mean ± SEM and were analyzed by using Student's *t* test.

Primers were designed from cDNA sequences of *Kitl* (M57647; nucleotides 1099–1751), Gdf9 (NM008110; nucleotides 720-1532), *Bmp15* (NM009757; nucleotides 146–973), *Ers2* (NM010157; nucleotides 1139–1921), *Pgr* (NM008829; nucleotides 1587–2425), *Cyp11a1* (NM019779; nucleotides 761-1697), *Cyp17al* (M64863; nucleotides 522–932), *Cyp19* (D00659; nucleotides 699-1049), *Fshr* (AF095642; nucleotides 625-1427), *Lhr* (M81310; nucleotides 592-1331), *Ptgs2* (AF338730; nucleotides 3–605), and *Ccnd2* (NM009829; nucleotides 150-1065) and chosen from different exons to avoid amplification from genomic DNA.

GeneChip Analysis. Ovaries were isolated and stabilized in RNA-*later* RNA Stabilization Reagent (Ambion, Austin, TX) before RNA purification (17). Total RNA was purified by using an RNeasy mini kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. First-strand cDNA was synthesized from 5 μg of RNA by using 200 units of SuperScript II reverse transcriptase (Invitrogen, Carlsbad, CA), 100 pmol T7-$(dT)_{24}$ primer [5′-GGCCAGTGAATTGTAATACGACTCAC-TATAGGGAGGCGG-$(dT)_{24}$-3′], 1× first-strand buffer, and 0.5 mM dNTPs at 42°C for 1 h. Second-strand synthesis was performed by incubating first-strand cDNA with 10 units of *Escherichia coli* ligase (Invitrogen), 40 units of DNA polymerase I (Invitrogen), 2 units of RNase H (Invitrogen), 1× reaction buffer, and 0.2 mM dNTPs at 16°C for 2 h, followed by 10 units of T4 DNA polymerase (Invitrogen) and incubation for another 5 min at 16°C. Double-stranded cDNA was purified by using GeneChip Sample Cleanup Module (Affymetrix, Santa Clara, CA) according to the manufacturer's instructions and labeled by *in vitro* transcription by using a BioArray HighYield RNA transcript labeling kit (Enzo Diagnostics, Farmingdale, NY). Briefly, dsDNA was mixed with 1× HY reaction buffer, 1× biotin-labeled ribonucleotides (NTPs with Bio-UTP and Bio-CTP), 1× DTT, 1× RNase inhibitor mix, and 1× T7 RNA polymerase and incubated at 37°C for 4 h. Labeled cRNA was then purified by using GeneChip Sample Cleanup Module and fragmented in 1× fragmentation buffer at 94°C for 35 min. For hybridization to the GeneChip Mouse Expression Array 430A or 430B or Mouse Genome 430 2.0 Array (Affymetrix), 15 μg of fragmented cRNA probe was incubated with 50 pM control oligonucleotide B2, 1× eukaryotic hybridization control, 0.1 mg/ml herring sperm DNA, 0.5 mg/ml acetylated BSA, and 1× hybridization buffer in a 45°C rotisserie oven for 16 h. Washing and staining were performed by using a GeneChip Fluidic Station (Affymetrix) according to the manufacturer's protocol. Phycoerythrin-stained arrays were scanned as digital image files and analyzed with GENECHIP OPERATING SOFTWARE (Affymetrix) (17).

Luciferase Assay. The *Kitl* promoter region (−2866 to −1 bp) was inserted into the pGL3-basic vector (Promega) for assay using the Luciferase Assay System (Promega) (14, 16). Cells at 40–50% confluence were transfected with a reference pRL-CMV

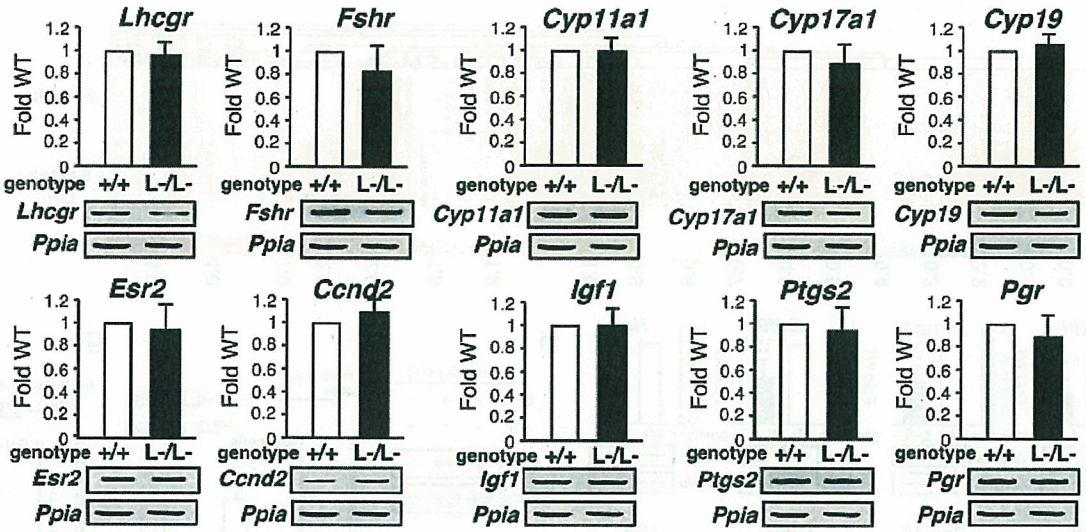

Fig. 3. No significant alterations in mRNA levels of several major regulators in folliculogenesis. Shown is semiquantitative RT-PCR of LH receptor (*Lhr*), FSH receptor (*Fshr*), p450 side chain cleavage enzyme (*Cyp11a1*), 17-α-hydroxylase (*Cyp17a1*), Aromatase (*Cyp19*), estrogen receptor-β (*Esr2*), cyclin D2 (*Ccnd2*), insulin-like growth factor 1 (*Igf1*), cyclooxygenase 2 (*Ptgs2*), or progesterone receptor (*Pgr*) gene expression in $AR^{+/+}$ and $AR^{-/-}$ ovaries. Results shown were representative (using one ovary per genotype in each experiment) of five independent experiments.

plasmid (Promega) using Lipofectamine reagent (GIBCO/BRL, Grand Island, NY) to normalize transfection. Results shown are representative of five independent experiments.

Results and Discussion

Subfertility of $AR^{-/-}$ Female Mice at 8 Weeks of Age. The *Ar* gene located on the X chromosome was disrupted in mice by using the Cre/Lox P system (6) (Fig. 1 *a–c*). Female $AR^{-/-}$ mice showed normal growth compared with the wild-type littermates (Fig. 1*d*), with no detectable bone loss (Fig. 1*e*) or obesity common for male $AR^{-/Y}$ mice (8, 9). Young (8-week-old) $AR^{-/-}$ females appeared indistinguishable from the wild-type littermates, displayed normal sexual behavior (7), and produced the first offspring of normal body size at the expected age. Macroscopic appearance of their reproductive organs, including uteri, oviducts, and ovaries, also appeared normal (Fig. 1*f*). Histological analysis showed no significant abnormality in the uterus or pituitary (Fig. 1*e*), whereas mammary ductal branching and elongation were substantially reduced, as revealed by whole-mount analysis (Fig. 1*h*). Serum levels of 17β-estradiol, progesterone, testosterone, luteinizing hormone, and follicle-stimulating hormone were also within normal range in 8-week-old mutant females at the proestrus stage (Fig. 1*g*), suggesting that the two-cell two-gonadotrophin system in female reproductive and endocrine organs (18) was intact in $AR^{-/-}$ mice at 8 weeks of age. The most obvious early sign of abnormal reproductive function in the $AR^{-/-}$ females was that their average numbers of pups per litter were only about half of those of the wild-type littermates, ($AR^{+/+}$, 8.3 ± 0.4 pups per litter; $AR^{-/-}$, 4.5 ± 0.5 pups per litter) (Fig. 1*i*).

$AR^{-/-}$ Female Mice Developed POF Phenotypes. Histological analysis of 8-week-old $AR^{-/-}$ ovaries clearly showed that numbers of atretic follicles were significantly increased, with decreased numbers of corpora lutea (Fig. 2 *b* and *f*). This finding suggests that the reduced pup numbers were due to impaired folliculogenesis in AR-deficient ovaries. Indeed, AR protein expression was readily detectable in the wild-type 8-week-old ovaries (Fig. 1*j*), with AR expressed at the highest levels in growing follicle granulosa cells at all developmental stages and at relatively low levels in corpora lutea. Thus, AR appears to play a regulatory role in granulosa cells during their maturation to the luteal phase.

To investigate this possibility, we examined the ovarian phenotype of female $AR^{-/-}$ mice at different ages. At 3 weeks, ovaries contain various stages of follicles, including primary, secondary, and antral follicles in wild-type animals (Fig. 2*a*) (19). In $AR^{-/-}$ ovaries at 3 weeks of age, the folliculogenesis appeared to be unaltered, with normal numbers and localization of primary and secondary follicles (Fig. 2 *a* and *e*). However, degenerated folliculogenesis became evident with further aging. Although follicles and corpora lutea at all developmental stages were still present, corpora lutea numbers were clearly reduced in 8-week-old $AR^{-/-}$ mutants (Fig. 2 *b* and *f*), similar to that observed in another mouse line (20). Expected apoptosis was seen in atretic follicles by activated caspase-3 immunohistochemistry assays (Fig. 2*i*). But, by 32 weeks of age, defects in folliculogenesis in $AR^{-/-}$ ovaries became profound, with fewer follicles observed and increased atretic follicles (Fig. 2 *c* and *g*), and >40% (5 of 12 mice) of the $AR^{-/-}$ females were already infertile. By 40 weeks, all $AR^{-/-}$ females became infertile, with no follicles remaining (Fig. 2 *d* and *h*); at the same age, $AR^{+/+}$ females were fertile and had normal follicle numbers. Consistent with progressive deficiency in folliculogenesis, the pup number per litter steadily decreased in aging $AR^{-/-}$ females (Fig. 2*i*). These data indicate that AR plays an important physiological role at the preluteal phase of folliculogenesis.

Alteration in Gene Expressions of Several Major Regulators Involved in the Oocyte–Granulosa Cell Regulatory Loop. To explore the molecular basis underlying the impaired folliculogenesis in $AR^{-/-}$ ovaries, we analyzed expression of several major known regulators and markers of folliculogenesis (21–23). Surprisingly, no significant alterations in mRNA levels of LH receptor (*Lhr*), FSH receptor (*Fshr*), p450 side chain cleavage enzyme (*Cyp11a1*), 17-α-hydroxylase (*Cyp17a1*), aromatase (*Cyp19*), estrogen receptor-β (*Esr2*), cyclin D2 (*Ccnd2*), or insulin-like growth factor 1 (*Igf1*) of 8-week-old $AR^{-/-}$ ovaries at the proestrus stage, and further cyclooxygenase 2 (*Ptgs2*) or progesterone receptor (*Pgr*) at the estrus stage, were detected by

PHYSIOLOGY

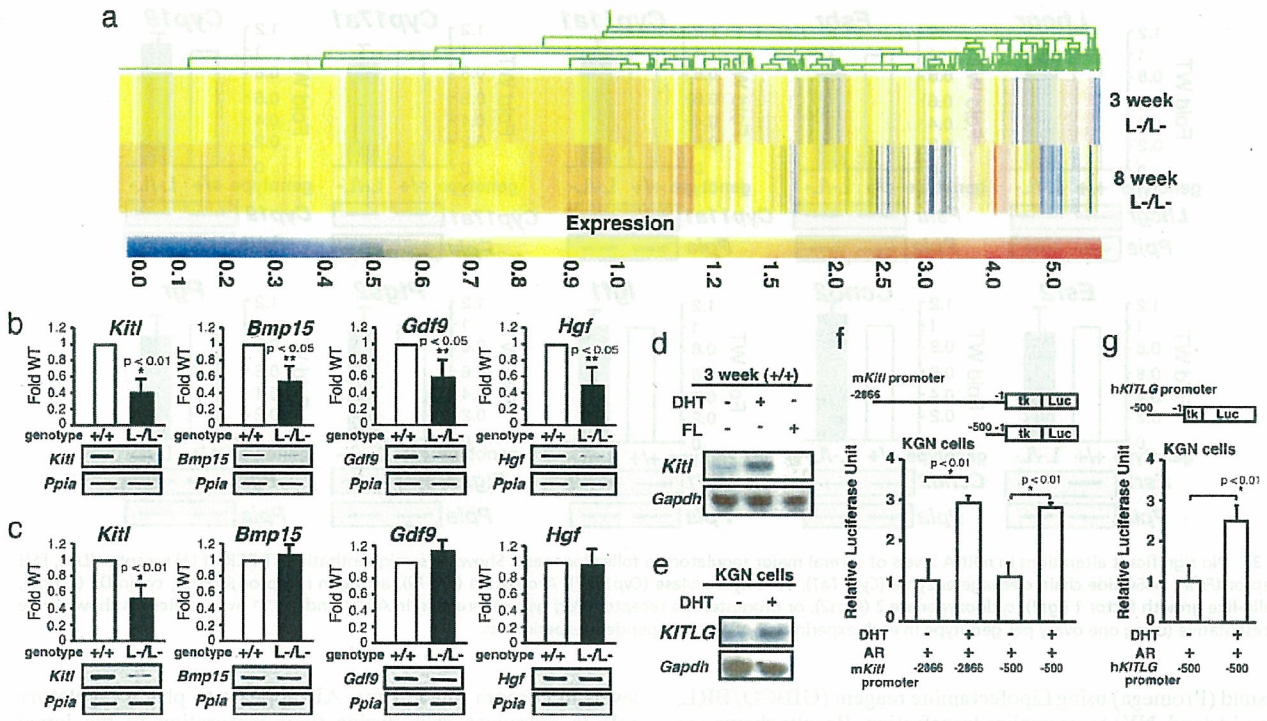

Fig. 4. Genome-wide microarray analysis and semiquantitative RT-PCR revealed that expression of the oocyte–granulosa cell regulator loop was down-regulated in $AR^{-/-}$ ovaries. (*a*) Microarray analysis of $AR^{-/-}$ compared with $AR^{+/+}$ ovaries at 3 and 8 weeks of age. Data obtained from microarray analysis as described in *Materials and Methods* were used to generate a cluster analysis. Each vertical line represents a single gene. The ratios of gene expression levels in $AR^{-/-}$ ovaries compared with wild type are presented. (*b* and *c*) Semiquantitative RT-PCR analysis of AR-regulated genes identified from the microarray study. Results shown are representative (using one ovary per genotype in each experiment) of five independent experiments. Data are shown as mean ± SEM and were analyzed by using Student's *t* test. (*d*) Comparison of *Kitl* gene expression by Northern blot analysis among placebo-, DHT-, and flutamide (FL)-treated $AR^{+/+}$ mouse ovaries. (e) Induction of *KITLG* gene expression by DHT treatment in KGN cells. (*f* and *g*) Androgen responsiveness in the mouse and human *kit ligand* promoters by a luciferase assay performed by using KGN cells. Data are shown as mean ± SEM and were analyzed by using Student's *t* test.

semiquantitative RT-PCR analysis (Fig. 3). Genome-wide microarray analysis (17) of RNA from 8-week-old $AR^{-/-}$ ovaries at the proestrus stage has been undertaken to identify AR-regulated genes. In comparison with $AR^{+/+}$ ovaries, expressions of 772 genes were down-regulated, whereas 351 genes were up-regulated in $AR^{-/-}$ ovaries (Fig. 4*a*; see also Tables 1 and 2, which are published as supporting information on the PNAS web site). Several genes known to be involved in the oocyte–granulosa cell regulatory loop (24) were identified as candidate AR target genes, including KIT ligand (*Kitl*) (25), morphogenetic protein 15 (*Bmp15*) (26), growth differentiation factor-9 (*Gdf9*) (27), and hepatocyte growth factor (*Hgf*) (28). Impaired folliculogenesis had been reported in mice deficient in each of these three regulators (26, 27, 29). To validate the microarray data, we performed semiquantitative RT-PCR analysis of 8-week-old $AR^{-/-}$ ovary RNA and confirmed that expression of these factors was down-regulated (Fig. 4*b*). To identify a regulator downstream of the AR signaling at an earlier stage of folliculogenesis, 3-week-old $AR^{-/-}$ ovaries that, as pointed out earlier, display no apparent phenotypic abnormality were examined. Fewer genes had altered expression levels (519 genes up-regulated; 326 genes down-regulated) (Fig. 4*a*; see also Tables 3 and 4, which are published as supporting information on the PNAS web site), and, of the four regulators tested by RT-PCR, only *Kitl* was found to be down-regulated at this age (Fig. 4*c*). Because *Kitl* is a granulosa cell-derived factor and stimulates oocyte growth and maturation (29–31), down-regulation of the *Kitl* expression in 3-week-old or even younger $AR^{-/-}$ ovaries may trigger impairment in folliculogenesis at a later age. To test for possible *Kitl* gene regulation by AR, 3-week-old wild-type females were treated with 5α-dihydrotestosterone (DHT). At 4 h after hormone injection, a clear induction of *Kitl* expression was observed in the ovaries, whereas a known antiandrogen flutamide attenuated the induction by DHT (Fig. 4*d*). The induction of endogenous human *kit ligand* (*KITLG*) gene by DHT was also observed in human granulosa-like tumor cells (KGN) in culture (Fig. 4*e*). Furthermore, androgen-induced transactivation of mouse and human *kit ligand* promoters (32) was observed by a luciferase reporter assay (33) in KGN (Fig. 4 *f* and *g*), 293T, and HeLa (data not shown) cells. However, no response to DHT was detected in the similar assay using promoters of the *Bmp15*, *Gdf9*, and *Hgf* genes (data not shown). Thus, we have shown that, in a regulatory cascade controlling folliculogenesis, *Kitl* represents a direct downstream target of androgen signaling.

As an upstream regulator, AR may also be indirectly involved in control of expression of other genes critical for folliculogenesis, because an age-dependent down-regulation of *Bmp15*, *Gdf9*, and *Hgf* gene expression was also observed in $AR^{-/-}$ ovaries. *Bmp15* and *Gdf9* are oocyte-derived factors that promote the development of surrounding granulosa cells in growing follicles (34, 35), whereas *Hgf* is secreted by theca cells and acts as a granulosa cell growth factor (36). Down-regulation of these factors, presumably due to decreased *Kitl* expression, may lead to impaired bidirectional communication between oocyte and granulosa cells (24) and, eventually, to early termination of folliculogenesis, as in POF syndrome.

Thus, we have identified AR as a novel regulator of follicu-

logenesis that apparently acts in the regulatory cascade upstream of the major factors controlling ovarian function, confirming the previous findings of the AR expression in granulose cells of growing follicles (3). Although not immediately relevant to the ovarian physiology, abnormal development of the mammary glands observed in our AR-deficient mice adds further strong evidence of an essential role of the AR not only in male, but also in female, reproductive function.

With increasing age of the first childbirth by women in the modern society, POF syndrome has become an important social and medical problem. Our findings suggest that POF syndrome may be caused by an impairment in androgen signaling and that X chromosomal mutations affecting the *AR* gene function may play a key role in hereditary POF. From clinical perspective, the present study provides evidence that AR can be a beneficial therapeutic target in treatment of POF syndrome patients.

We thank T. Iwamori and H. Tojo for expert advice on mammary gland anatomy, Y. Kanai for ovarian phenotypic analysis, members of the KO project team at the laboratory of Nuclear Signaling (Institute of Molecular and Cellular Biosciences) for their support, A. P. Kouzmenko for helpful suggestions, and H. Higuchi for manuscript preparation. This work was supported in part by the Program for Promotion of Basic Research Activities for Innovative Biosciences and priority areas from the Ministry of Education, Culture, Sports, Science, and Technology (to S.K.).

1. Laml, T., Preyer, O., Umek, W., Hengstschlager, M. & Hanzal, H. (2002) *Hum. Reprod. Update* **8,** 483–491.
2. Davison, R. M., Davis, C. J. & Conway, G. S. (1999) *Clin. Endocrinol. (Oxford)* **51,** 673–679.
3. Tetsuka, M., Whitelaw, P. F., Bremner, W. J., Millar, M. R., Smyth, C. D. & Hillier, S. G. (1995) *J. Endocrinol.* **145,** 535–543.
4. Ehrmann, D. A., Barnes, R. B. & Rosenfield, R. L. (1995) *Endocr. Rev.* **16,** 322–353.
5. Norman, R. J. (2002) *Mol. Cell. Endocrinol.* **191,** 113–119.
6. Kato, S. (2002) *Clin. Pediatr. Endocrinol.* **11,** 1–7.
7. Sato, T., Matsumoto, T., Kawano, H., Watanabe, T., Uematsu, Y., Sekine, K., Fukuda, T., Aihara, K., Krust, A., Yamada, T., *et al.* (2004) *Proc. Natl. Acad. Sci. USA* **101,** 1673–1678.
8. Sato, T., Matsumoto, T., Yamada, T., Watanabe, T., Kawano, H. & Kato, S. (2003) *Biochem. Biophys. Res. Commun.* **300,** 167–171.
9. Kawano, H., Sato, T., Yamada, T., Matsumoto, T., Sekine, K., Watanabe, T., Nakamura, T., Fukuda, T., Yoshimura, K., Yoshizawa, T., *et al.* (2003) *Proc. Natl. Acad. Sci. USA* **100,** 9416–9421.
10. Li, M., Indra, A. K., Warot, X., Brocard, J., Messaddeq, N., Kato, S., Metzger, D. & Chambon, P. (2000) *Nature* **407,** 633–636.
11. Sekine, K., Ohuchi, H., Fujiwara, M., Yamasaki, M., Yoshizawa, T., Sato, T., Yagishita, N., Matsui, D., Koga, Y., Itoh, N. & Kato, S. (1999) *Nat. Genet.* **21,** 138–141.
12. Yoshizawa, T., Handa, Y., Uematsu, Y., Takeda, S., Sekine, K., Yoshihara, Y., Kawakami, T., Arioka, K., Sato, H., Uchiyama, Y., *et al.* (1997) *Nat. Genet.* **16,** 391–396.
13. Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P. & Lovell-Badge, R. (1990) *Nature* **346,** 245–250.
14. Yanagisawa, J., Yanagi, Y., Masuhiro, Y., Suzawa, M., Watanabe, M., Kashiwagi, K., Toriyabe, T., Kawabata, M., Miyazono, K. & Kato, S. (1999) *Science* **283,** 1317–1321.
15. Britt, K. L., Drummond, A. E., Cox, V. A., Dyson, M., Wreford, N. G., Jones, M. E., Simpson, E. R. & Findlay, J. K. (2000) *Endocrinology* **141,** 2614–2623.
16. Ohtake, F., Takeyama, K., Matsumoto, T., Kitagawa, H., Yamamoto, Y., Nohara, K., Tohyama, C., Krust, A., Mimura, J., Chambon, P., *et al.* (2003) *Nature* **423,** 545–550.
17. Fujimoto, N., Igarashi, K., Kanno, J., Honda, H. & Inoue, T. (2004) *J. Steroid Biochem. Mol. Biol.* **91,** 121–129.
18. Couse, J. F. & Korach, K. S. (1999) *Endocr. Rev.* **20,** 358–417.
19. Elvin, J. A. & Matzuk, M. M. (1998) *Rev. Reprod.* **3,** 183–195.
20. Hu, Y. C., Wang, P. H., Yeh, S., Wang, R. S., Xie, C., Xu, Q., Zhou, X., Chao, H. T., Tsai, M. Y. & Chang, C. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 11209–11214.
21. Elvin, J. A., Yan, C., Wang, P., Nishimori, K. & Matzuk, M. M. (1999) *Mol. Endocrinol.* **13,** 1018–1034.
22. Zhou, J., Kumar, T. R., Matzuk, M. M. & Bondy, C. (1997) *Mol. Endocrinol.* **11,** 1924–1933.
23. Burns, K. H., Yan, C., Kumar, T. R. & Matzuk, M. M. (2001) *Endocrinology* **142,** 2742–2751.
24. Matzuk, M. M., Burns, K. H., Viveiros, M. M. & Eppig, J. J. (2002) *Science* **296,** 2178–2180.
25. Joyce, I. M., Pendola, F. L., Wigglesworth, K. & Eppig, J. J. (1999) *Dev. Biol.* **214,** 342–353.
26. Yan, C., Wang, P., DeMayo, J., DeMayo, F. J., Elvin, J. A., Carino, C., Prasad, S. V., Skinner, S. S., Dunbar, B. S., Dube, J. L., *et al.* (2001) *Mol. Endocrinol.* **15,** 854–866.
27. Dong, J., Albertini, D. F., Nishimori, K., Kumar, T. R., Lu, N. & Matzuk, M. M. (1996) *Nature* **383,** 531–535.
28. Parrott, J. A., Vigne, J. L., Chu, B. Z. & Skinner, M. K. (1994) *Endocrinology* **135,** 569–575.
29. Driancourt, M. A., Reynaud, K., Cortvrindt, R. & Smitz, J. (2000) *Rev. Reprod.* **5,** 143–152.
30. Huang, E. J., Manova, K., Packer, A. I., Sanchez, S., Bachvarova, R. F. & Besmer, P. (1993) *Dev. Biol.* **157,** 100–109.
31. Packer, A. I., Hsu, Y. C., Besmer, P. & Bachvarova, R. F. (1994) *Dev. Biol.* **161,** 194–205.
32. Grimaldi, P., Capolunghi, F., Geremia, R. & Rossi, P. (2003) *Biol. Reprod.* **69,** 1979–1988.
33. Kitagawa, H., Fujiki, R., Yoshimura, K., Mezaki, Y., Uematsu, Y., Matsui, D., Ogawa, S., Unno, K., Okubo, M., Tokita, A., *et al.* (2003) *Cell* **113,** 905–917.
34. Otsuka, F. & Shimasaki, S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 8060–8065.
35. Joyce, I. M., Clark, A. T., Pendola, F. L. & Eppig, J. J. (2000) *Biol. Reprod.* **63,** 1669–1675.
36. Parrott, J. A. & Skinner, M. K. (1998) *Endocrinology* **139,** 2240–2245.

Antitumor Activity of ZSTK474, a New Phosphatidylinositol 3-Kinase Inhibitor

Shin-ichi Yaguchi, Yasuhisa Fukui, Ichiro Koshimizu, Hisashi Yoshimi, Toshiyuki Matsuno, Hiroaki Gouda, Shuichi Hirono, Kanami Yamazaki, Takao Yamori

Background: **We previously synthesized a novel *s*-triazine derivative, ZSTK474 [2-(2-difluoromethylbenzimidazol-1-yl)-4,6-dimorpholino-1,3,5-triazine], that strongly inhibited the growth of tumor cells. We identified its molecular target, investigated its effects on cellular signaling pathways, and examined its antitumor efficacy and toxicity in vivo.** ***Methods:*** **We used COMPARE analysis of chemosensitivity measurements from 39 human cancer cell lines and identified phosphatidylinositol 3-kinase (PI3K) as a molecular target for ZSTK474. PI3K was immunoprecipitated from A549 cell lysates, and its activity was measured by assessing the incorporation of ^{32}P into phosphatidylinositol. We used the crystal structure of the PI3K–LY294002 complex to model the binding of ZSTK474 to PI3K (where LY294002 is a known PI3K inhibitor). PI3K downstream activity was analyzed by immunoblotting. Antitumor activity of ZSTK474 was examined against A549, PC-3, and WiDr xenografts in nude mice. Phosphorylation of Akt, a serine/threonine protein kinase and a major signaling component downstream of PI3K, was assessed in vivo by immunohistochemistry.** ***Results:*** **PI3K was identified as a molecular target for ZSTK474 by COMPARE analysis. We confirmed that ZSTK474 directly inhibited PI3K activity more efficiently than the PI3K inhibitor LY294002. At concentrations of 1 μM, ZSTK474 and LY2194002 reduced PI3K activity to 4.7% (95% confidence interval [CI] = 3.2% to 6.1%) and 44.6% (95% CI = 38.9% to 50.3%), respectively, of the untreated control level. Molecular modeling of the PI3K–ZSTK474 complex indicated that ZSTK474 could bind to the ATP-binding pocket of PI3K. ZSTK474 inhibited phosphorylation of signaling components downstream from PI3K, such as Akt and glycogen synthase kinase 3β, and mediated a decrease in cyclin D1 levels. ZSTK474 administered orally to mice had strong antitumor activity against human cancer xenografts without toxic effects in critical organs. Akt phosphorylation was reduced in xenograft tumors after oral administration of ZSTK474.** ***Conclusion:*** **ZSTK474 is a new PI3K inhibitor with strong antitumor activity against human cancer xenografts without toxic effects in critical organs. ZSTK474 merits further investigation as an anticancer drug. [J Natl Cancer Inst 2006;98:545–56]**

Phosphatidylinositol 3-kinase (PI3K), which is active in signal transduction, generates phosphatidylinositol-3,4,5-trisphosphate (PIP_3) by phosphorylating phosphatidylinositol-4,5-bisphosphate *(1,2)*. PI3K is involved in various cellular processes, including cell survival, vesicle transport, and cytoskeletal rearrangement *(3,4)*. The importance of PI3K in tumorigenesis is supported by the following evidence. The chicken retrovirus ASV16 carries an oncogenic PI3K gene *(5)*, and mutation and/or amplification of PI3K genes *(6–10)* has been reported in human and other mammalian cancer cells. Constitutive activation of PI3K may contribute to the malignant phenotype of signet ring carcinomas *(11)*, and PTEN, a PIP_3 phosphatase, has been identified as a tumor suppressor gene *(12)*.

Many clinical studies also indicate that deregulation of PI3K pathway plays a role in various cancers *(3,7)* and that PI3K may thus be a good target for anticancer drug development *(3)*. Validation of PI3K as a good target for drug development began with the use of dominant-negative mutants and RNA interference studies *(13,14)*. Indeed, a number of compounds have been identified that inhibit PI3K, including wortmannin *(15,16)*, which was originally isolated from soil bacteria and is toxic to fungi; the closely related viridin analogues *(17,18)*; and LY294002, a morpholino derivative of the broad-spectrum kinase inhibitor quercetin *(19)*. Because of inherent difficulties with the stability, solubility, and toxicity of these compounds, efforts are under way to develop new inhibitors of the PI3K pathway *(20,21)*.

In previous work, we synthesized a chemical library of *s*-triazine derivatives and screened these compounds for their ability to inhibit tumor cell growth *(22,23)*. Among more than 1500 *s*-triazine derivatives, we found a new compound, ZSTK474 [2-(2-difluoromethylbenzimidazol-1-yl)-4,6-dimorpholino-1,3,5-triazine], that showed strong antiproliferative activity. However, its molecular target or its potential for a novel anticancer drug was unknown.

We previously established a panel of 39 human cancer cell lines (termed JFCR39) coupled to a drug activity database *(24–26)* that is comparable to the panel developed by the National

Affiliations of authors: Division of Molecular Pharmacology, Cancer Chemotherapy Center, Japanese Foundation for Cancer Research, Ariake, Koto-ku, Tokyo, Japan (SY, KY, TY); Research Laboratory, Zenyaku Kogyo Co., Ltd., Ohizumi-machi, Nerima-ku, Tokyo, Japan (SY, IK, HY, TM); Laboratory of Biological Chemistry, Department of Applied Biological Chemistry, Faculty of Agricultural and Life Science, Graduate School of Agricultural and Life Science, University of Tokyo, Yayoi, Bunkyo-ku, Tokyo, Japan (YF); The School of Pharmaceutical Sciences, Kitasato University, Shirokane, Minato-ku, Tokyo, Japan (HG, SH).

Correspondence to: Takao Yamori, PhD, Division of Molecular Pharmacology, Cancer Chemotherapy Center, Japanese Foundation for Cancer Research, 3-10-6 Ariake, Koto-ku, Tokyo 135-8550, Japan (e-mail: yamori@jfcr.or.jp).

See "Notes" following "References."

DOI: 10.1093/jnci/djj133

Cancer Institute *(27,28)*. We compared cell growth inhibition profiles against the JFCR39 (herein termed fingerprints) of more than 60 standard drugs including doxorubicin, fluorouracil, cytarabine, methotrexate, vincristine, cisplatin, paclitaxel, and irinotecan, by using COMPARE analysis *(26)*, and we showed that COMPARE analysis is an information-intensive approach to identifying the molecular target of a new compound, as described by Paull et al. *(28)*. This system can be used to predict the molecular target or the mode of action of test compounds by assessing the correlation coefficient between the fingerprints mediated by such test compounds and by various reference compounds with known modes of action, including PI3K inhibitors.

The purpose of this study was to identify the molecular target of ZSTK474 with the aid of COMPARE analysis, to investigate the effects of ZSTK474 on cellular signaling pathways, and to examine the antitumor efficacy and toxicity of ZSTK474 in vivo.

Materials and Methods

Chemicals

ZSTK116, ZSTK474, ZSTK781, and ZSTK1178 were synthesized in the Research Laboratory, Zenyaku Kogyo Co., Ltd. (Tokyo, Japan). For in vitro studies, these compounds were dissolved in dimethyl sulfoxide. ZSTK474 was suspended in 5% hydroxypropylcellulose in water as a solid dispersion form for animal experiments. Other anticancer drugs and chemicals were purchased as follows: doxorubicin and 5-fluorouracil from Kyowa Hakko Kogyo (Tokyo, Japan); cisplatin from Nippon Kayaku (Tokyo, Japan); irinotecan from Yakult Honsha (Tokyo, Japan); and wortmannin, LY294002, tetramethylrhodamine isothiocyanate-conjugated phalloidin, Igepal CA-630, and platelet-derived growth factor from Sigma (St. Louis, MO).

Cell Lines

A panel of 39 human cancer cell lines (termed JFCR39) *(25,26)* and B16F10 melanoma cells *(29)* were previously described. All cell lines were cultured in RPMI 1640 medium supplemented with 5% fetal bovine serum, penicillin (100 U/mL), and streptomycin (100 μg/mL) at 37 °C in humidified air containing 5% CO_2. Cell lines OVCAR3, A549, PC-3, and WiDr, which were used for in vitro detailed analysis or for in vivo study, originated from an ovarian cancer, non–small-cell lung cancer, prostate cancer, and colon cancer, respectively. For in vivo study, A549, PC-3, and WiDr were grown as subcutaneous tumors in nude mice.

Analysis of Cell Proliferation Inhibition

The inhibition of cell proliferation was assessed by measuring changes in total cellular protein in a culture of each cell line in the JFCR panel of cell lines after 48 hours of drug treatment by use of a sulforhodamine B assay *(30)*. The 50% growth inhibition (GI_{50}) value of the drug was calculated as described previously *(26,27)*. The graphic representation of a drug's mean differential growth inhibition for the cell line panel was based on a calculation that used a set of GI_{50} values, as described previously *(28,31)*. To analyze the correlation between the mean graphs of drug A and drug B, the COMPARE computer algorithm was developed as previously described by Paull et al. *(28)*. The Pearson correlation coefficient between the mean graphs of drug A and drug B was calculated (n = 39).

PI3K Assay in Tumor Cell Extracts

The PI3K assay has been previously described by Fukui et al. *(32)*. In brief, A549 cells were lysed in a buffer containing 20 mM Tris–HCl (pH 7.5), 150 mM NaCl, 5 mM EDTA, and 1% Igepal CA-630 (Sigma), the lysates were centrifuged at 20000*g* and 4 °C for 10 minutes, and the supernatants were used as cell lysate (protein = 2–4 mg/mL). To immunoprecipitate PI3K, we incubated 200 μL of cell lysate with anti-p85 polyclonal antibody (1 : 200 dilution; Upstate, Charlottesville, VA) and protein G–agarose (5 μL; Upstate). The anti-p85 polyclonal antibody recognizes the p85 regulatory subunit of PI3K. Class Ia catalytic subunits of PI3K (p110α, p110β, and p110δ) constitutively associate with the p85 subunit. Therefore, PI3Kα, PI3Kβ, and PI3Kδ can be immunoprecipitated by the anti-p85 polyclonal antibody. Agarose beads containing immunoprecipitates were washed twice with buffer A (20 mM Tris–HCl at pH 7.5, 150 mM NaCl, 5 mM EDTA, and 1% Igepal CA-630), once with buffer B (500 mM LiCl and 100 mM Tris–HCl at pH 7.5), once with distilled water, and once with buffer C (100 mM NaCl and 20 mM Tris–HCl at pH 7.5). Immunoprecipitates were suspended in 20 μL of buffer C containing phosphatidylinositol at 200 μg/mL. The mixture was preincubated with or without ZSTK474, LY294002, or wortmannin, as indicated, at 25 °C for 5 minutes. [γ-^{32}P]ATP (2 μCi per assay mixture; final concentration, 20 μM) and $MgCl_2$ (final concentration, 20 mM) were added to start the reaction. The reaction mixture was incubated at 25 °C for 20 minutes. During this incubation, formation of phosphatidylinositol-3-phosphate was linear (data not shown). Phosphorylated products of phosphatidylinositol were separated by thin-layer chromatography and visualized by autoradiography. The phosphatidylinositol-3-phosphate region was scraped from the plate, and radioactivity was also measured with liquid scintillation spectroscopy (LS6500 Scintillation System; Beckman Instruments, Fullerton, CA). The level of inhibition for each compound was determined as the percentage of ^{32}P counts per minute obtained without the test compound. To asses the reversibility of PI3K inhibition, the PI3K immunoprecipitates were treated with 1 μM ZSTK474, 10 μM LY294002, or 0.1 μM wortmannin for 5 minutes, and the reaction mixture was divided into two aliquots. One aliquot was washed four times with a buffer containing 100 mM NaCl and 20 mM Tris–HCl (pH 7.5) to remove free ZSTK474, LY294002, or wortmannin from the PI3K, and the PI3K activity of the washed precipitate was measured. The PI3K activity of the other aliquot was measured without washing.

PI3K Assay for Recombinant Catalytic Subunits of PI3K

PI3K assay for recombinant p110, the catalytic subunit of PI3K, was described previously by Gray et al. *(33)*. In a reaction volume of 20 μL, human p110 isoforms β, -γ, or -δ (Upstate) were incubated with 10 μM phosphatidylinositol-4,5-bisphosphate and 10 μM ATP in assay buffer for 30 minutes at room temperature. Stop buffer (5 μL) containing EDTA and biotinylated PIP_3 was added, followed by 5 μL of detection buffer containing europium-labeled anti–glutathione *S*-transferase (GST) monoclonal antibody, GST-tagged general receptor for phosphoinositides (GRP1) PH domain, and streptavidin allophycocyanin. Plus

and minus kinase control wells were also included. The plate was read in the time-resolved fluorescence mode, and the homogenous time-resolved fluorescence (HTRF) signal was determined according to the following formula: HTRF = 10 000 × (emission at 665 nm/emission at 620 nm).

Molecular Modeling of the PI3K–ZSTK474 Complex

Conformational analysis of ZSTK474 was performed first with the FlexX docking program *(34)*, which was used to construct a model of the PI3K–ZSTK474 complex structure. This program treats ring systems, such as the morpholino groups in ZSTK474, as rigid structures. The conformational analysis was done with the CAMDAS, version 2.1, program (conformational analyzer with molecular dynamics and sampling) *(35)*. Conditions used for CAMDAS calculation were similar to those described previously *(36)*. Twelve dihedral angles with respect to two morpholino rings were used to cluster similar conformations. As a result of the CAMDAS calculation, 64 different conformers were obtained.

The crystal structure of the PI3K–LY294002 complex (Protein Data Bank Identification Code 1E7V) was used to obtain the PI3K template structure for docking. The docking calculation of each ZSTK474 conformation with PI3K was performed with the FlexX program. A total of 10 482 different docking models were generated. Each docking model was then evaluated with the following five different score functions: the F-score *(34)*, the D-score *(37)*, the G-score *(38)*, the ChemScore *(39)*, and the PMF-score *(40)*. These functions were implemented in the SYBYL C-Score module. Finally, a new index, termed the average of auto-scaled score, was calculated to rank all docking models. This index has been recently defined in our laboratory to more adequately rank docking models *(41)*. In the top 20 docking models that were ranked according to the average of auto-scaled scores, all ZSTK474 molecules were found to similarly interact with PI3K. The only difference among the 20 models was with respect to the conformation of the two morpholino rings of ZSTK474. Therefore, we present one of these 20 models, in which two morpholino rings of ZSTK474 take a chair conformation, because this model has the lowest internal energy.

Assay for Chromatin Condensation

OVCAR3 cells were cultured in RPMI 1640 medium supplemented with 5% fetal bovine serum for 24 hours and then cultured for 48 hours with or without 10 μM ZSTK474. Cells were collected, fixed in 1% glutaraldehyde in phosphate-buffered saline (PBS) at 4 °C for 24 hours, stained with 167 μM Hoechst 33342 (Sigma), and examined by fluorescence microscopy to identify apoptotic cells.

Flow Cytometry Analysis

Cells were harvested, washed with ice-cold PBS, and fixed in 70% ethanol. Cells were then washed twice with ice-cold PBS again, treated with RNase A (500 μg/mL; Sigma) at 37 °C for 1 hour, and stained with propidium iodide (25 μg/mL; Sigma). The DNA content of the cells was analyzed with a flow cytometer (FACScalibur, Becton Dickinson, Franklin Lakes, NJ).

Membrane Ruffling

Murine embryonic fibroblasts were treated with or without 1 μM ZSTK474 for 15 minutes and then stimulated with platelet-derived growth factor (10 ng/mL) for 5 minutes. Cells were fixed with 3% formaldehyde at room temperature for 30 minutes and permeabilized with 0.1% Triton X-100 in PBS for 5 minutes. Filamentous actin in cells was stained with tetramethylrhodamine isothiocyanate-conjugated phalloidin at 0.1 μg/mL, and cells were examined by fluorescence microscopy.

Analysis of PIP_3 Production

PIP_3 production in intact cells was measured as described previously by Kimura et al. *(42)*. In brief, to label murine embryonic fibroblasts, medium was replaced with prewarmed phosphate-free Dulbecco's modified Eagle medium with 25 mM HEPES (pH 7.4) containing [^{32}P]orthophosphate (0.1 mCi/mL), and cultures were incubated at 37 °C for 4 hours. Lipids were extracted with chloroform and separated by thin-layer chromatography. PIP_3 spots were visualized by autoradiography and identified by standards synthesized by PI3K in vitro (data not shown). Radioactivity in the spots was quantitated with a PhosphorImager SI system (Molecular Dynamics, Sunnyvale, CA).

Western Blot Analysis

A549 cells were lysed in a buffer containing 10 mM Tris–HCl at pH 7.4, 50 mM NaCl, 50 mM NaF, 30 mM sodium pyrophosphate, 50 mM Na_3VO_4, 5 mM EDTA, aprotinin at 100 Kal U/mL, 1 mM phenylmethylsulfonyl fluoride, 0.5% Nonidet P-40, and 0.1% sodium dodecyl sulfate. Proteins in cell lysates were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis followed by electroblotting onto a polyvinylidene difluoride membrane (Amersham Biosciences, Piscataway, NJ). Rabbit polyclonal antibodies for Akt, phosphorylated Akt (phosphorylated residue Ser-473), phosphorylated glycogen synthase kinase 3β (GSK-3β; phosphorylated residue Ser-9), phosphorylated FKHR (phosphorylated residue Thr-24)/phosphorylated FKHRL1 (phosphorylated residue Thr-32), phosphorylated TSC-2 (phosphorylated residue Ser-1462), phosphorylated mTOR (phosphorylated residue Ser-2448), phosphorylated p70S6K (phosphorylated residue Thr-389), phosphorylated MEK1/2 (phosphorylated residues Ser-217/221), and phosphorylated ERK1/2 (phosphorylated residues Thr-202/Tyr-204) were purchased from Cell Signaling Technology (Beverly, MA) and were diluted 1 : 1000. Rabbit anti–cyclin D1 polyclonal antibody (1 : 200 dilution) was purchased from BioSource International (Camarillo, CA). Horseradish peroxidase–conjugated goat anti–rabbit immunoglobulin G (Cell Signaling Technology) was used as a secondary antibody. Immunoreactive bands were identified with the ECL-plus Western Blotting Detection System (Amersham Biosciences).

Animal Experiments

Animal care and treatment was performed in accordance with the guidelines of the animal use and care committee of the Research Laboratory, Zenyaku Kogyo, and conformed to the NIH *Guide for the Care and Use of Laboratory Animals*. All procedures were approved by the above committee. Male BDF_1 mice and female nude mice with BALB/c genetic backgrounds were purchased from Charles River Japan, Inc. (Yokohama, Japan). Mice were maintained under specific pathogen–free conditions and provided with sterile food and water ad libitum. We used 56 BDF_1 mice. Each BDF_1 mouse was subcutaneously

injected with 5×10^5 B16F10 melanoma cells. Seven days after inoculation, animals were divided randomly into test groups (each with seven mice), and the administration of the drugs, as indicated, began (day 0). ZSTK474 at 100, 200, or 400 mg/kg of body weight was orally administered daily from days 0 to 13. Reference drugs were irinotecan, cisplatin, doxorubicin, and 5-fluorouracil. Each reference drug was administered intravenously from day 0 at the maximum tolerable dose by use of the following optimal schedules: irinotecan at 100 mg/kg administered on days 0, 3, and 7; cisplatin at 10 mg/kg on day 0; doxorubicin at 12 mg/kg on day 0; and 5-fluorouracil at 50 mg/kg on days 0, 3, and 7. Human tumor xenografts were generated with A549 lung cancer cells, PC-3 prostate cancer cells, and WiDr colon cancer cells and were grown as subcutaneous tumors in nude mice. Forty-five nude mice were used in these experiments. Each nude mouse was subcutaneously inoculated with a tumor fragment of 3 mm × 3 mm × 3 mm. When tumors reached a volume of 100–300 mm^3, animals were divided randomly into test groups (each with five mice) (day 0). ZSTK474 (400 mg/kg) was administered orally from day 0 until day 13 except for days 3 and 10. For longer administration periods, ZSTK474 at 400 mg/kg was orally administered daily from days 0 to 26, except for days 6, 13, and 20. Then, the mice were photographed on day 28. In another experiment, ZSTK474 at a higher dosage of 800 mg/kg was orally administered in the same schedule, and the body weight changes of the mice were monitored to observe toxicity.

Measurement of Tumor Volume

The length (L) and width (W) of the subcutaneous tumor mass were measured by calipers in live mice, and the tumor volume (TV) was calculated as: $TV = (L \times W^2)/2$. The tumor volume at day n was expressed as relative tumor volume (RTV), according to the formula $RTV = TV_n/TV_0$, where TV_n is the tumor volume at day n, and TV_0 is the tumor volume at day 0. Tumor regression ($T/C\%$) on day 14 was determined by use of the RTV values as follows: $T/C\% = 100 \times$ (mean RTV of treated group)/(mean RTV of control group). To assess toxicity, we measured the body weight of the tumor-bearing mice. After the chronic administration of ZSTK474, mice were subjected to necropsy examination, including the histopathologic evaluation of bone marrow.

Bone Marrow in Mice Femur

Right-side femurs were fixed with 10% neutral formalin and decalcified with Plank–Rychlo fluid *(43)*. After embedding in paraffin, 3-μm sections of bone marrow were stained with hematoxylin and eosin.

Immunohistochemistry

When the volume of tumors reached 100–300 mm^3, ZSTK474 at 400 mg/kg was orally administered once. Mice were killed 4 hours after ZSTK474 administration, and the tumors were excised. Tumor tissue was fixed in 10% neutral formalin and embedded in paraffin. The 6-μm sections were deparaffinized in xylene and then in a 100% to 50% ethanol series. Immunohistochemistry-specific phosphorylated Akt (phosphorylated on Ser-473) antibody and phosphorylated Akt (phosphorylated on Ser-473) blocking peptide were purchased from Cell Signaling Technology. The tissue sections were analyzed by immunohistochemistry for the expression of phosphorylated Akt (Ser-473). For control experiment, phosphorylated Akt (Ser-473) blocking peptide was incubated with phosphorylated Akt (Ser-473) antibody at 4 °C for 2 hours before incubating with the tissue section. The ABC method (Vectastain ABC system kit; Vector Laboratories, Burlingame, CA) was used for immunohistochemical staining, with the sections also being counterstained with hematoxylin.

Statistical Analysis

Pearson correlation coefficients were calculated for the COMPARE analysis and statistical correlation. The two-sided Mann–Whitney U test was used to test the statistical significance of the antitumor efficacy of ZSTK474 in relative tumor growth ratio on days 4, 7, 11, and 14. The number of samples is indicated in the description of each experiment. All statistical tests were two-sided.

Results

Measurement of the Growth-Inhibitory Activity of ZSTK474 and Identification of its Molecular Target

We synthesized and screened a chemical library of *s*-triazine derivatives based on their ability to mediate growth inhibition of tumor cells. We found that ZSTK474 showed potent antiproliferative activity. The chemical structure of ZSTK474 is shown in Fig. 1.

To identify potential molecular targets for ZSTK474, we used COMPARE analysis, an approach that is based on chemosensitivity measurements from the JFCR39 panel of cell lines. We treated cells in the JFCR39 panel with ZSTK474 and examined its effect on cell proliferation. In this way, we obtained the fingerprint for ZSTK474 (Fig. 2). The COMPARE analysis revealed that the fingerprint of ZSTK474 correlated with the fingerprints of noanticancer drug currently in use. However, the fingerprint of ZSTK474 did correlate statistically significantly with that of LY294002, a PI3K inhibitor ($r = .766$ and $P < .001$), as shown in Fig. 2, and with that of wortmannin, another well-characterized PI3K inhibitor ($r = .519$ and $P = .001$). Thus ZSTK474 appears to act as a PI3K inhibitor, although its structure is different from that of LY294002 or of wortmannin (Fig. 1).

We also compared the potency of ZSTK474 in cell growth inhibition with those of LY294002 and wortmannin. The mean logarithm of GI_{50} for ZSTK474 was −6.49 (at 0.32 μM ZSTK474), which was substantially lower than that observed with LY294002 (−5.13 at 7.4 μM) or wortmannin (−5.00 at 10 μM), indicating that ZSTK474 had stronger cell growth inhibitory activity than either of these two known PI3K inhibitors.

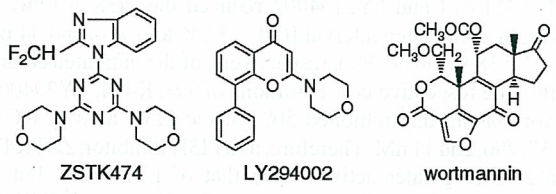

Fig. 1. Chemical structures of the phosphatidylinositol 3-kinase inhibitors ZSTK474, LY294002, and wortmannin. ZSTK474 has one 1-benzimidazolyl and two morpholino groups as substituents on the 1,3,5-triazine ring.

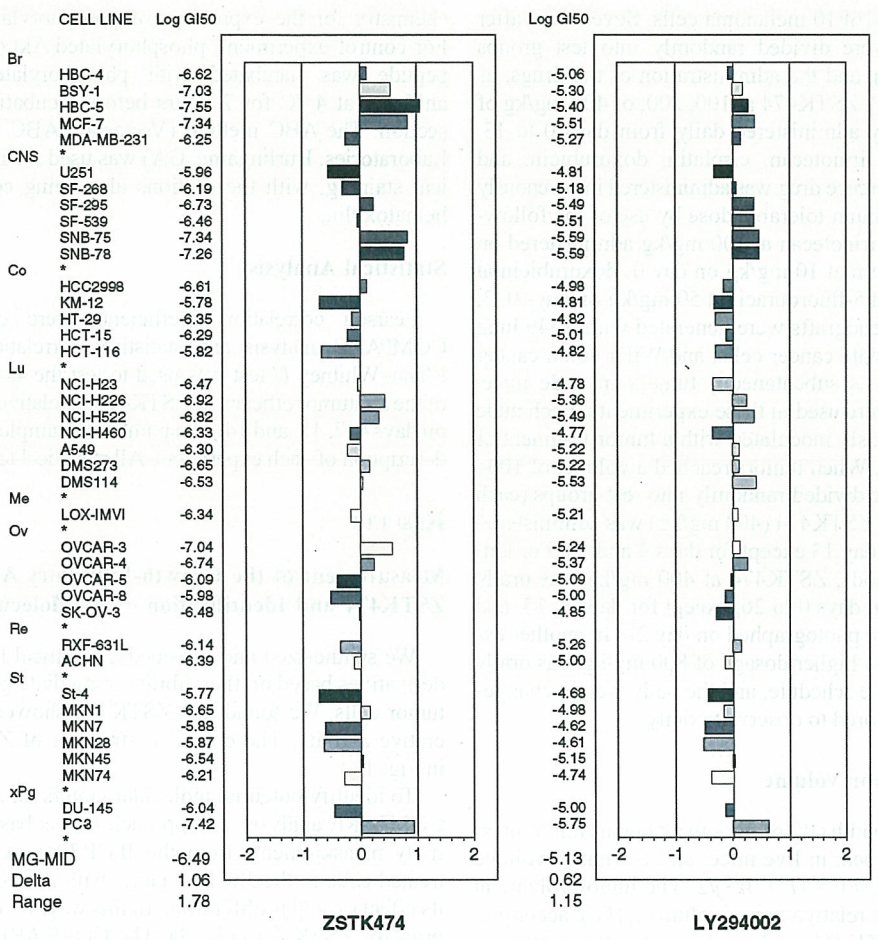

Fig. 2. Growth inhibition activity of ZSTK474 and LY294002 against a panel of 39 human cancer cell lines in the JFCR39 panel. Mean graph was produced by computer processing of the 50% growth inhibition (GI_{50}) values. Logarithm of the GI_{50} value for each cell line is indicated. In the plot, **columns to the right of zero** indicate that the sensitivity of the cell line to the compound, and **columns to the left** indicate resistance to the compound. The *x*-axis represents logarithm of difference between the mean of GI_{50} values for 39 cell lines and the GI_{50} value for each cell line in the JFCR39 panel. One scale represents one logarithm difference. MG-MID = mean of logarithm of GI_{50} values for 39 cell lines in the JFCR39 panel; Delta = logarithm of difference between the MG-MID and the logarithm of the GI_{50} value for the most sensitive cell line; Range = logarithm of difference between the logarithm of the GI_{50} value for the most resistant cell line and the logarithm of the GI_{50} value for the most sensitive cell line. Br = breast; CNS = central nervous system; Co = colon ; Lu = lung; Me = melanoma; Ov = ovarian; Re = renal; St = stomach; xPg = prostate. Each **hatch mark** corresponds to the cell line indicated to the left.

Inhibition of PI3K Activity by ZSTK474

We next examined the ability of ZSTK474 to inhibit PI3K activity. We selected A549 cells as the source of PI3K because these cells had an average level of sensitivity to ZSTK474 among cells in the JFCR39 panel. For these experiments, we used PI3K that had been immunoprecipitated from A549 cell lysates with an anti-p85 polyclonal antibody. ZSTK474 inhibited PI3K activity in a dose-dependent manner (Fig. 3, A). At concentrations of 1 μM, ZSTK474 and LY2194002 reduced the PI3K activity to 4.7% (95% confidence interval [CI] = 3.2% to 6.1%) and 44.6% (95% CI = 38.9% to 50.3%), respectively, of the untreated control activity. The respective concentrations of ZSTK474, LY294002, and wortmannin that inhibited 50% of the PI3K activity (IC_{50}) were 37, 790, and 11 nM. Therefore, as a PI3K inhibitor, ZSTK474 had a 20-fold greater activity than that of LY294002, but its activity was similar to that of wortmannin. However, ZSTK474 reversibly inhibited PI3K activity, as does LY294002, whereas wortmannin irreversibly inhibited PI3K activity (Fig. 3, B). ZSTK474 did not substantially inhibit the activity of 139 other protein kinases (Supplementary Table 1, available at http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue8), indicating that ZSTK474 is highly specific to PI3K.

There are four subtypes of p110, the catalytic subunit of PI3K—α, β, γ, and δ. The specificity of ZSTK474 to these subtypes was of interest because a high frequency of mutations in the PIK3CA gene, encoding p110α, in human cancers has been recently reported *(7,8)*. With the PI3K assay described above, it was not possible to determine subtype specificity because the precipitates obtained with the anti-p85 polyclonal antibody could contain a mixture of p110α, -β, and -δ. To determine the subtype specificity of ZSTK474, we examined ZSTK474 for its activity against recombinant p110β, -γ, and -δ. We found that ZSTK474 inhibited the activities of p110β, -γ, and -δ, with respective IC_{50} values of 17, 53, and 6 nM (Supplementary Fig. 1 available at http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue8), indicating that ZSTK474 was apparently not subtype specific.

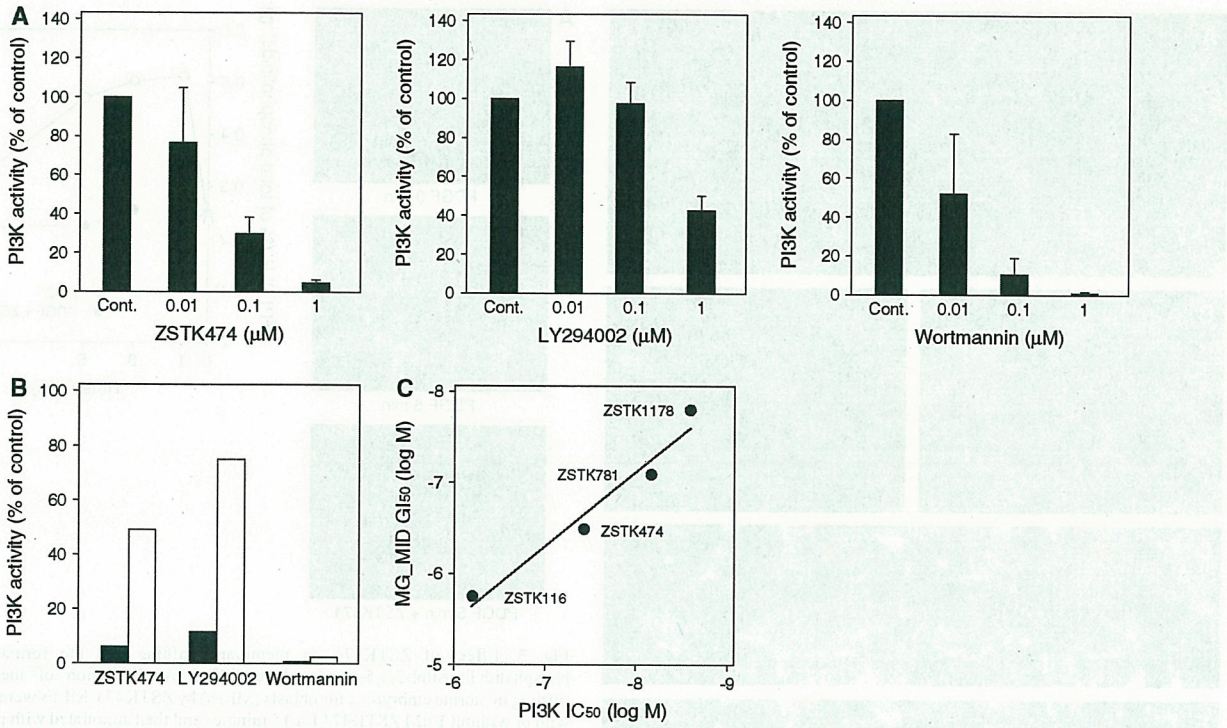

Fig. 3. Effect of ZSTK474 on phosphatidylinositol 3-kinase (PI3K) activity. A) Inhibition of PI3K activity. PI3K was immunoprecipitated from A549 cell lysates with anti-p85 polyclonal antibody, and immunoprecipitates were thoroughly washed sequentially with buffer A, buffer B, distilled water, and buffer C. After treatment of the PI3K immunoprecipitates with ZSTK474, LY294002, or wortmannin, as indicated, for 5 minutes, activity was measured by incubating PI3K immunoprecipitates with phosphatidylinositol and [γ-^{32}P]ATP. Phosphorylated derivatives of phosphatidylinositol were separated by thin-layer chromatography. The radioactivity in all derivatives was measured, and the results are expressed as the percentage of control untreated PI3K activity. Data are the mean of the results of two experiments performed in triplicate. **Error bars** = upper 95% confidence intervals. B) Reversibility of PI3K inhibition. The PI3K immunoprecipitates were treated with 1 μM ZSTK474, 10 μM LY294002, or 0.1 μM wortmannin for 5 minutes and divided into two aliquots. One aliquot was washed four times with a buffer containing 100 mM NaCl and 20 mM Tris–HCl (pH 7.5) to remove free ZSTK474, LY294002, or wortmannin from the PI3K, and PI3K activity of the washed precipitate was measured. The PI3K activity of the other aliquot was measured without washing. The PI3K activity of the immunoprecipitates with **(solid columns)** or without **(open columns)** washing is shown as the percentage of control untreated PI3K activity. Data are the mean of the results of two or three experiments. C) Activity of the *s*-triazine derivatives ZSTK116, ZSTK474, ZSTK781, and ZSTK1178. Scattered plot of concentrations of the various derivatives that inhibited 50% of PI3K activity (IC_{50}) versus the mean of logarithm of 50% growth inhibition (GI_{50}) (MG-MID) values of *s*-triazine derivatives. A high and statistically significant correlation (r = .98, Pearson correlation coefficient, and P = .023) was observed between the two activities (n = 4). All statistical tests were two-sided.

Correlation of PI3K Inhibition with Growth-Inhibitory Activity of *s*-Triazine Derivatives

ZSTK474 is an *s*-triazine derivative, as are ZSTK1178, ZSTK781, and ZSTK116, and all have different PI3K inhibitory activities. We investigated whether the PI3K inhibitory activities (IC_{50}) of these four *s*-triazine derivatives were related to their growth inhibitory activities (assessed as the mean logarithm of GI_{50}). A statistically significantly high correlation (r = .98 and P = .023) was observed between these two inhibitory activities (Fig. 3, C), suggesting that ZSTK474 inhibited cell proliferation by inhibiting PI3K activity.

Molecular Modeling of the PI3K–ZSTK474 Complex

Because all previously described PI3K inhibitors apparently bind to the ATP-binding site of p110, the catalytic subunit of PI3K *(44)*, we investigated whether ZSTK474 also binds to this site by using a model structure, the crystal structure of the PI3Kγ–LY294002 complex, to generate a PI3K template structure for docking analysis *(44)*. Analysis of this model structure indicated that there are three hydrogen-bonding interactions between PI3K and ZSTK474 (i.e., K802 NZ, one of the morpholino oxygens; S806 OG, the other morpholino oxygen; and V882 NH, the benzimidazole nitrogen [Fig. 4, A]).

We next compared the binding modes of ZSTK474 with those of ATP in the active site of PI3K. Two of the three hydrogen bonds between ZSTK474 and PI3K, V882 NH and S806 OG, mimicked those in the ATP–PI3K complex; therefore, the overall arrangements of ZSTK474 and ATP in the PI3K active site are very close (Fig. 4, B). The benzimidazole nitrogen of ZSTK474 forms a hydrogen bond with V882 NH. This hydrogen bond mimics the hydrogen bond between ATP N1 and V882 NH. As discussed previously *(45)*, all kinase inhibitors appear to have a hydrogen bond acceptor in a position equivalent to ATP N1. These features of the ZSTK474–PI3K complex support the hypothesis that ZSTK474 binds to the ATP-binding site of PI3K. In contrast, ZSTK474 and LY294002 appear to bind to different sites in the ATP-binding pocket (Fig. 4, C), with LY294002 forming two hydrogen bonds between PI3K at V882 NH and K833 NZ.

As shown in Fig. 4, D and E, the amino acid residues Met-804 (M804), Trp-812 (W812), and Met-953 (M953) of PI3K create a unique space in the ATP-binding pocket that is larger than the