imental condition. Our results also support the findings of our previous works,[8,9,12] suggesting the importance of inflammatory stimuli as a promotion event after the initiation events (genetic alterations) in colon carcinogenesis. There were no differences between males and females in the effects of DSS on large and small intestinal carcinogenesis of $Apc^{Min/+}$, and the histopathology of colonic lesions including neoplasms was similar in both sexes.

As for the development of small intestinal polyps, treatment with DSS significantly increased their number and size, especially at the distal portion of the small intestine. Macrophages engulfing DSS particles were observed in the large intestine and surrounding lymph nodes of mice 1 day after DSS exposure, and then found in the jejunum and ileum 7 days after DSS treatment.[51] In our study, mild mucosal inflammation was observed in the distal portion of the small intestine of mice given 2% DSS. Thus, DSS could also influence the formation of small intestinal polyp in $Apc^{Min/+}$ mice. The Min mouse has been regarded as a human FAP model in spite of the fact that the polyps (adenomas) develop in the small intestine. Although the biological pathways in human colon and Min intestine are assumed to be similar, our model described here could be applied for investigation of the genesis, pathophysiology and chemoprevention of human FAP and/or inflammation-related colon tumorigenesis.

In our study, sequential observation on the pathological alteration in the large intestines of female $Apc^{Min/+}$ mice after 1-week exposure to 2% DSS revealed that the frequencies of dysplastic crypts and colonic neoplasms (adenoma and adenocarcinoma) gradually increased over time (Fig. 2a,b), indicating that dysplastic crypts[43] or adenomatous lesions[41] are precursor lesions for colon carcinoma and DSS treatment could promote their growth. The findings support an earlier report by Cooper *et al.*,[14] but their incidence of colonic cancer was low: 22% in Min mice exposed to 1-cycle of DSS (administration 4% DSS for 4 days and H_2O for 17 days) and 40% in Min mice exposed to 2-cycle of DSS. The discrepancy existing in these 2 studies may be due to the differences in the treatment period and the dose and molecular weight of DSS. In the present study, the incidence of colonic adenocarcinoma was 80% at week 4 and 100% at week 5 (Fig. 2a). When compared to our previous study on the effects of DSS on chemically induced colon carcinogenesis,[9] where we observed 40% and 100% incidences of colonic epithelial malignancy at week 4 and week 6, respectively, in male ICR mice, it is likely that deletion of the *Apc* gene plays an important role in colitis-associated carcinogenesis, as suggested by Cooper *et al.*[14]

In our study, we investigated the immunohistochemical expression of β-catenin, COX-2, iNOS and p53, in the colonic lesions developed in ApcMin/+ mice that received 2% DSS. The results on immunohistochemistry against these antibodies except for p53 expression in the lesions were similar to those observed in our previous studies, where the lesions were induced by AOM[8,9], HCAs[12] or DMH[11] followed by DSS in ICR mice, suggesting the similarity of histopathology and immunohistochemistry, and biological nature of the lesions observed in ICR mice given a colonic carcinogen and DSS and $Apc^{Min/+}$ mice treated with DSS. Increased immunohistochemical expression of COX-2 and iNOS in the colonic tumors of either $Apc^{Min/+}$ mice that received 2% DSS was confirmed by reverse transcription-polymerase chain reaction (data not shown). The findings of nitrotyrosine immunohistochemistry in the current study are also in accordance with those in our previous study[9] and suggest that oxidative/nitrosative stress strongly promotes the development of colonic neoplasms in $Apc^{Min/+}$ mice. iNOS has been shown to be the only isoform involved in stimulating tumor growth, probably through an increase in vascular endothelial growth factor production.[52] Moreover, NO regulates COX-2 expression.[53] Our results on the immu-

nohistochemistry of iNOS and COX-2 indicate that the inflammatory response, the interaction between NO synthase and COX pathways may stand at the center of the pathophysiological basis of inflammation-related colon carcinogenesis in $Apc^{Min/+}$ mice treated with DSS, as are the cases of inflammatory diseases,[54] and chemically induced colon carcinogenesis.[33]

In the current study, we also screened for mutations of β-catenin and K-*ras* in colon tumors developed in male $Apc^{Min/+}$ mice. In contrast with previous reports,[12,33,55] we did not detect the mutations of these genes in any of the colonic adenocarcinomas examined. However, our results are not surprising. Suzui *et al.*[56] reported that adenocarcinomas developed in $Apc^{Min/+}$ mice treated with AOM did not have β-*catenin* gene mutations. In our study, cytoplasmic and/or nuclear accumulation of β-catenin protein was detected in the colonic neoplasms, but β-*catenin* gene mutations were not present. In the FAP patients, mutations of *APC* are common, but mutations of β-*catenin* were rare.[57,58] In addition, β-*catenin* germline mutations were not found in FAP patients with germline *APC* mutations.[57] Thus, concerning the β-*catenin* mutation, the colon tumors developed in the current animal model may imitate the colon carcinogenesis as in the FAP patients, that is, by a second hit in the *APC* gene such as loss of Apc^+ allele or somatic mutations in the *Apc* gene. Immunohistochemical staining with an antibody for the C-terminal of *Apc* showed the loss of wild-type *Apc* in colonic tumors in $Apc^{Min/+}$ mice (data not shown). As for the mutation of K-*ras*, no mutations were found in the colonic adenocarcinomas examined in the current study. Our results on K-*ras* mutations are in accordance with IBD-related colon carcinogenesis[59] and suggest that activation of the K-*ras* gene is not essential for the development and growth of colonic neoplasms in our model.

p53 gene mutation occurs in the late stage of human colon carcinogenesis.[33,59] In our study, p53 immunohistochemistry revealed positive reaction in the nuclei of neoplastic cells in $Apc^{Min/+}$ mice treated with DSS, although we did not examine its mutation in our study. The accumulation of p53 shown in our study is interesting and may be important for colon cancer development in *Apc*-deficient mice, since an increased *p53* mutation load in the inflamed colon tissue from UC patients being a high-risk for colon cancer[60] and a potential mechanism link between NO and p53 in UC and sporadic colon cancer[61] were reported. In addition, COX-2, iNOS and p53 are suggested to be fundamental "play-makers" of the angiogenesis processes.[62]

Taken together, our results suggest that a novel $Apc^{Min/+}$ mouse model with DSS may provide new insight into the genesis and chemoprevention of colon cancer development in FAP patients. In our model, a single allele *Apc* gene followed by appropriate promotional stimuli is sufficient for the development and growth of colonic neoplasms in $Apc^{Min/+}$ mice, and COX-2, iNOS, p53, oxidative/nitrosative stress and interactions of these may play important roles in colon carcinogenesis in $Apc^{Min/+}$ mice given DSS. Our model can be applied for investigating the pathogenesis in carcinogenesis of IBD, since the Wnt/β-catenin signaling pathway may be involved in carcinogenesis of UC.[59,63] Our ongoing microarray analysis will provide new information of the mechanism(s) for the effects of DSS on large and small intestinal tumorigenesis in $Apc^{Min/+}$ mice.

References

1. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? Lancet 2001;357:539–45.
2. Ohshima H, Tatemichi M, Sawa T. Chemical basis of inflammation-induced carcinogenesis. Arch Biochem Biophys 2003;417:3–11.
3. Szilagyi A. Altered colonic environment, a possible predisposition to colorectal cancer and colonic inflammatory bowel disease: rationale of dietary manipulation with emphasis on disaccharides. Can J Gastroenterol 1998;12:133–46.

4. Seril DN, Liao J, Yang GY, Yang CS. Oxidative stress and ulcerative colitis-associated carcinogenesis: studies in humans and animal models. Carcinogenesis 2003;24:353–62.

5. Munkholm P. Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease. Aliment Pharmacol Ther 2003;18:1–5.

6. van der Woude CJ, Kleibeuker JH, Jansen PL, Moshage H. Chronic inflammation, apoptosis and (pre-)malignant lesions in the gastrointestinal tract. Apoptosis 2004;9:123–30.

7. Okayasu I, Hatakeyama S, Yamada M, Ohkusa T, Inagaki Y, Nakaya R. A novel method in the induction of reliable experimental acute and chronic ulcerative colitis in mice. Gastroenterology 1990;98:694–702.

8. Tanaka T, Kohno H, Suzuki R, Yamada Y, Sugie S, Mori H. A novel inflammation-related mouse colon carcinogenesis model induced by azoxymethane and dextran sodium sulfate. Cancer Sci. 2003;94:965–73.

9. Suzuki R, Kohno H, Sugie S, Tanaka T. Sequential observations on the occurrence of preneoplastic and neoplastic lesions in the mouse colon treated with azoxymethane and dextran sodium sulfate. Cancer Sci. 2004;95:721–7.

10. Suzuki R, Kohno H, Sugie S, Tanaka T. Dose-dependent promoting effect of dextran sodium sulfate on mouse colon carcinogenesis initiated with azoxymethane. Histol Histopathol 2005;20:483–92.

11. Kohno H, Suzuki R, Sugie S, Tanaka T. β-Catenin mutations in a mouse model of inflammation-related colon carcinogenesis induced by 1,2-dimethylhydrazine and dextran sodium sulfate. Cancer Sci 2005; 96:69–76.

12. Tanaka T, Suzuki R, Kohno H, Sugie S, Takahashi M, Wakabayashi K. Colonic adenocarcinomas rapidly induced by the combined treatment with 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine and dextran sodium sulfate in male ICR mice possess bgr;-catenin gene mutations and increase immunorectivity for bgr;-catenin, cyclooxygenase-2, and inducible nitric oxide synthase. Carcinogenesis 2005;26: 229–38.

13. Cooper HS, Murthy S, Kido K, Yoshitake H, Flanigan A. Dysplasia and cancer in the dextran sulfate sodium mouse colitis model. Relevance to colitis associated neoplasia in the human: study of histopathology, beta-catenin and p53 expression and the role of inflammation. Carcinogenesis 2000;21:757–68.

14. Cooper HS, Everley L, Chang W-C, Pfeeiffer G, Lee B, Murthy S, Clapper ML. The role of mutant Apc in the development of dysplasia and cancer in the mouse model of dextran sulfate sodium–induced colitis. Gastroenterology 2001;121:1407–16.

15. Barbour KW, Davis T, White A, Baumann H, Berger FG. Haptoglobin, inflammation, and tumorigenesis in the MIN mouse. Redox Rep 2001;6:366–8.

16. Hansen-Petrik MB, McEntee MF, Jull B, Shi H, Zemel MB, Whelan J. Prostaglandin E2 protects intestinal tumors from nonsteroidal anti-inflammatory drug-induced regression in Apc$^{Min/-}$ mice. Cancer Res 2002;62:403–8.

17. Prescott SM, Fitzpatrick FA. Cyclooxygenase-2 and carcinogenesis. Biochim Biophys Acta 2000;1470:M69–78.

18. Eberhart CE, Coffey RJ, Radhika A, Giardiello FM, Ferrenbach S, DuBois RN. Up-regulation of cyclooxygenase 2 gene expression in human colorectal adenomas and adenocarcinomas. Gastroenterology 1994;107:1183–8.

19. DuBois RN, Radhika A, Reddy BS, Entingh AJ. Increased cyclooxygenase-2 levels in carcinogen-induced rat colonic tumors. Gastroenterology 1996;110:1295–62.

20. Hull MA, Booth JK, Tisbury A, Scott N, Bonifer C, Markham AF, Coletta PL. Cyclooxygenase 2 is up-regulated and localized to macrophages in the intestine of Min mice. Br J Cancer 1999;79:1399–405.

21. Arroyo PL, Hatch-Pigott V, Mower HF, Cooney RV. Mutagenicity of nitric oxide and its inhibition by antioxidants. Mutat Res 1992;281:193–202.

22. Wink DA, Kasprzak KS, Maragos CM, Elespuru RK, Misra M, Dunams TM, Cebula TA, Koch WH, Andrews AW, Allen JS. DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. Science (Washington, DC) 1991;254:1001–3.

23. Xiong M, Elson G, Legarda D, Leibovich SJ. Production of vascular endothelial growth factor by murine macrophages: regulation by hypoxia, lactate, and the inducible nitric oxide synthase pathway. Am J Pathol 1998;153:587–98.

24. Kojima M, Morisaki T, Tsukahara Y, Uchiyama A, Matsunari Y, Mibu R, Tanaka M. Nitric oxide synthase expression and nitric oxide production in human colon carcinoma tissue. J Surg Oncol 1999;70:222–9.

25. Bing RJ, Miyataka M, Rich KA, Hanson N, Wang X, Slosser HD, Shi SR. Nitric oxide, prostanoids, cyclooxygenase, and angiogenesis in colon and breast cancer. Clin Cancer Res 2001;7:3385–92.

26. Takahashi M, Fukuda K, Ohata T, Sugimura T, Wakabayashi K. Increased expression of inducible and endothelial constitutive nitric oxide synthases in rat colon tumors induced by azoxymethane. Cancer Res 1997;57:1233–7.

27. Rao CV, Kawamori T, Hamid R, Reddy BS. Chemoprevention of colonic aberrant crypt foci by an inducible nitric oxide synthase-selective inhibitor. Carcinogenesis 1999;20:641–4.

28. Ahn B, Ohshima H. Suppression of intestinal polyposis in Apc$^{Min/+}$ mice by inhibiting nitric oxide production. Cancer Res. 2001;61: 8357–60.

29. Scott DJ, Hull MA, Cartwright EJ, Lam WK, Tisbury A, Poulsom R, Markham AF, Bonifer C, Coletta PL. Lack of inducible nitric oxide synthase promotes intestinal tumorigenesis in the Apc$^{Min/+}$ mouse. Gastroenterology 2001;121:889–99.

30. Mabley JG, Pacher P, Bai P, Wallace R, Goonesekera S, Virag L, Southan GJ, Szabo C. Suppression of intestinal polyposis in Apc$^{Min/+}$ mice by targeting the nitric oxide or poly(ADP-ribose) pathways. Mutat Res 2004;548:107–16.

31. Lala PK, Chakraborty C. Role of nitric oxide in carcinogenesis and tumour progression. Lancet Oncol 2001;2:149–56.

32. Hofseth LJ, Hussain SP, Wogan GN, Harris CC. Nitric oxide in cancer and chemoprevention. Free Radic Biol Med 2003;34: 955–68.

33. Takahashi M, Wakabayashi K. Gene mutations and altered gene expression in azoxymethane-induced colon carcinogenesis in rodents. Cancer Sci 2004;95:475–80.

34. Hussain SP, Hofseth LJ, Harris CC. Radical causes of cancer. Nat Rev Cancer 2003;3:276–85.

35. Ames BN, Gold LS, Willett WC. The causes and prevention of cancer. Proc Natl Acad Sci U S A 1995;92:5258–65.

36. Hussain SP, Hofseth LJ, Harris CC. Tumor suppressor genes: at the crossroads of molecular carcinogenesis, molecular epidemiology and human risk assessment. Lung Cancer 2001;34:S7–S15.

37. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996;87:159–70.

38. Powell SM, Petersen GM, Krush AJ, Booker S, Jen J, Giardiello FM, Hamilton SR, Vogelstein B, Kinzler KW. Molecular diagnosis of familial adenomatous polyposis. N Engl J Med 1993; 329:1982–7.

39. Moser AR, Luongo C, Gould KA, McNeley MK, Shoemaker AR, Dove WF. ApcMin: a mouse model for intestinal and mammary tumorigenesis. Eur J Cancer 1995;31A:1061–4.

40. Paulsen JE. Modulation by dietary factors in murine FAP models. Toxicol Lett 2000;112–113:403–9.

41. Yamada Y, Hata K, Hirose Y, Hara A, Sugie S, Kuno T, Yoshimi N, Tanaka T, Mori H. Microadenomatous lesions involving loss of Apc heterozygosity in the colon of adult Apc$^{Min/+}$ mice. Cancer Res 2002; 62:6367–70.

42. Murata M, Kawanishi S. Oxidative DNA damage induced by nitrotyrosine, a biomarker of inflammation. Biochem Biophys Res Commun 2004;316:123–8.

43. Paulsen JE, Steffensen I-L, Loberg EM, Husoy T, Namork E, Alexander J. Qualitative and quantitative relationship between dysplastic aberrant crypt foci and tumorigenesis in the Min/+ mouse colon. Cancer Res 2001;61:5010–5.

44. Steffensen I-L, Paulsen JE, Eide TJ, Alexander J. 2-Amino-1-methyl-6- phenylimidazo[4,5-b]pyridine increases the numbers of tumors, cystic crypts and aberrant crypt foci in multiple intestinal neoplasia mice. Carcinogenesis (Lond.) 1997;18:1049–54.

45. Ward JM. Morphogenesis of chemically induced neoplasms of the colon and small intestine in rats. Lab Invest 1974;30:505–13.

46. Cooper HS, Murthy SN, Shah RS, Sedergran DJ. Clinicopathologic study of dextran sulfate sodium experimental murine colitis. Lab Invest 1993;69:238–49.

47. Kohno H, Suzuki R, Noguchi R, Hosokawa M, Miyashita K, Tanaka T. Dietary conjugated linolenic acid inhibits azoxymethane-induced colonic aberrant crypt foci in rats. Jpn J Cancer Res 2002;93:133–42.

48. Mollersen L, Paulsen JE, Olstom HB, Knutsen HK, Alexander J. Dietary retinoic acid supplementation stimulates intestinal tumour formation and growth in multiple intestinal neoplasia (Min)/+ mice. Carcinogenesis 2004;25:149–53.

49. Zingarelli B, Szabo C, Salzman AL. Reduced oxidative and nitrosative damage in murine experimental colitis in the absence of inducible nitric oxide synthase. Gut 1999;45:199–209.

50. Luongo C, Moser AR, Gledhill S, Dove WF. Loss of Apc+ in intestinal adenomas from Min mice. Cancer Res 1994;54:5947–52.

51. Kitajima S, Takuma S, Morimoto M. Tissue distribution of dextran sulfate sodium (DSS) in the acute phase of murine DSS-induced colitis. J Vet Med Sci 1999;61:67–70.

52. Konopka TE, Barker JE, Bamford TL, Guida E, Anderson RL, Stewart AG. Nitric oxide synthase II gene disruption: implications for tumor growth and vascular endothelial growth factor production. Cancer Res 2001;61:3182–7.

53. Salvemini D, Misko TP, Masferrer JL, Seibert K, Currie MG, Needleman P. Nitric oxide activates cyclooxygenase enzymes. Proc Natl Acad Sci U S A 1993;90:7240–4.

54. Perez-Sala D, Lamas S. Regulation of cyclooxygenase-2 expression by nitric oxide in cells. Antioxid Redox Signal 2001;3: 231–48.

55. Takahashi M, Nakatsugi S, Sugimura T, Wakabayashi K. Frequent mutations of the β-*catenin* gene in mouse colon tumors induced by azoxymethane. Carcinogenesis 2000;21:1117–20.

56. Suzui M, Okuno M, Tanaka T, Nakagama H, Moriwaki H. Enhanced colon carcinogenesis induced by azoxymethane in min mice occurs via a mechanism independent of β-*catenin* mutation. Cancer Lett 2002;183:31–41.

57. Muller O, Nimmrich I, Finke U, Friedl W, Hoffmann I. A β-catenin mutation in a sporadic colorectal tumor of the RER phenotype and absence of β-catenin germline mutations in FAP patients. Genes Chromosomes Cancer 1998;22:37–41.

58. Miyoshi Y, Ando H, Nagase H, Nishisho I, Horii A, Miki Y, Mori T, Utsunomiya J, Baba S, Petersen G, Hamilton S, Kinzler K et al. Germ-line mutations of the *APC* gene in 53 familial adenomatous polyposis patients. Proc Natl Acad Sci. U S A. 1992;89:4452–6.

59. Itzkowitz SH, Yio X. Inflammation and cancer IV. Colorectal cancer in inflammatory bowel disease: the role of inflammation. Am J Physiol Gastrointest Liver Physiol 2004;287:G7–17.

60. Hussain SP, Amstad P, Raja K, Ambs S, Nagashima M, Bennett WP, Shields PG, Ham AJ, Swenberg JA, Marrogi AJ, Harris CC. Increased p53 mutation load in noncancerous colon tissue from ulcerative colitis: a cancer-prone chronic inflammatory disease. Cancer Res 2000; 50:3333–7.

61. Goodman JE, Hofseth LJ, Hussain SP, Harris CC. Nitric oxide and p53 in cancer-prone chronic inflammation and oxyradical overload disease. Environ Mol Mutagen 2004;44:3–9.

62. Chiarugi V, Magnelli L, Gallo O. Cox-2, iNOS and p53 as playmakers of tumor angiogenesis (review). Int J Mol Med 1998;2: 715–9.

63. Uthoff SM, Eichenberger MR, Lewis RK, Fox MP, Hamilton CJ, McAuliffe TL, Grimes HL, Galandiuk S. Identification of candidate genes in ulcerative colitis and Crohn's disease using cDNA array technology. Int J Oncol 2001;19:803–10.

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

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
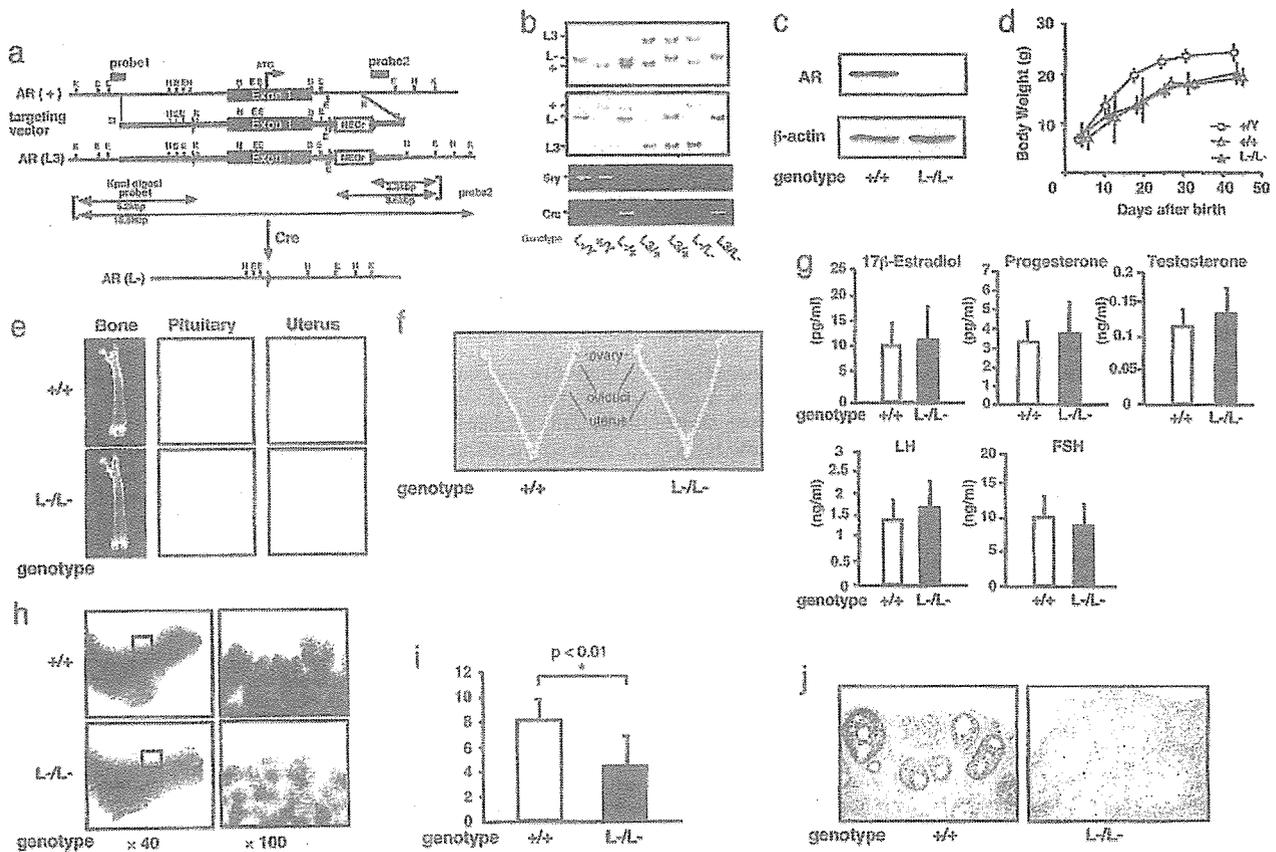

Fig. 1. Phenotypic characterization of AR knockout female mice. (a) Diagram of the wild-type Ar genomic locus (+), floxed AR L3 allele (L3), and AR allele (L-) obtained after Cre-mediated excision of exon 1. K, KpnI; E, EcoRI; H, HindIII; B, BamHI. LoxP sites are indicated by arrowheads. The targeting 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.
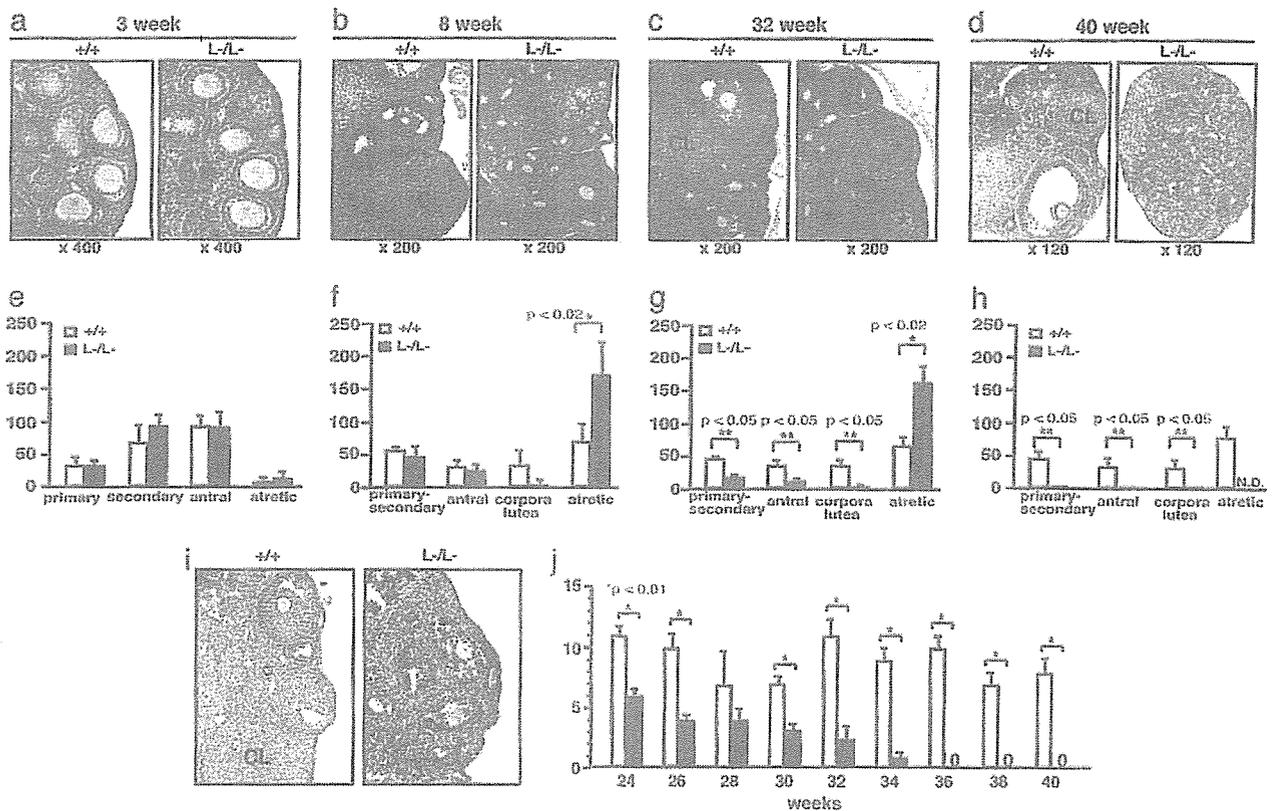
Fig. 2. POF in $AR^{-/-}$ female mice. (a–d) Histology of $AR^{+/+}$ and $AR^{-/-}$ ovaries at 3 weeks, 8 weeks, 32 weeks, and 40 weeks of age. All sections were stained with hematoxylin and eosin. An asterisk marks the atretic follicle. CL, corpus luteum. (e–h) Relative follicle counts at 3 weeks (e), 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 $\pm$ 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), *Cyp17a1* (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 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
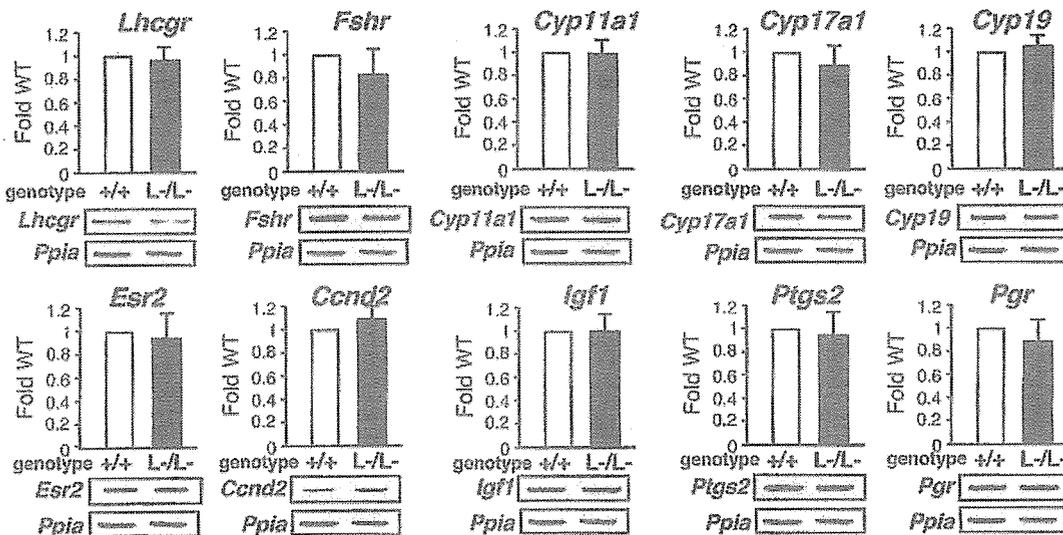
Fig. 3. No significant alterations in mRNA levels of several major regulators in folliculogenesis. Shown is semiquantitative RT-PCR of LH receptor (*Lhr*), FSH receptor (*Fshr*), p450 side chain cleavage enzyme (*Cyp11a1*), 17-α-hydroxylase (*Cyp17a1*), Aromatase (*Cyp19*), estrogen receptor-β (*Esr2*), 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
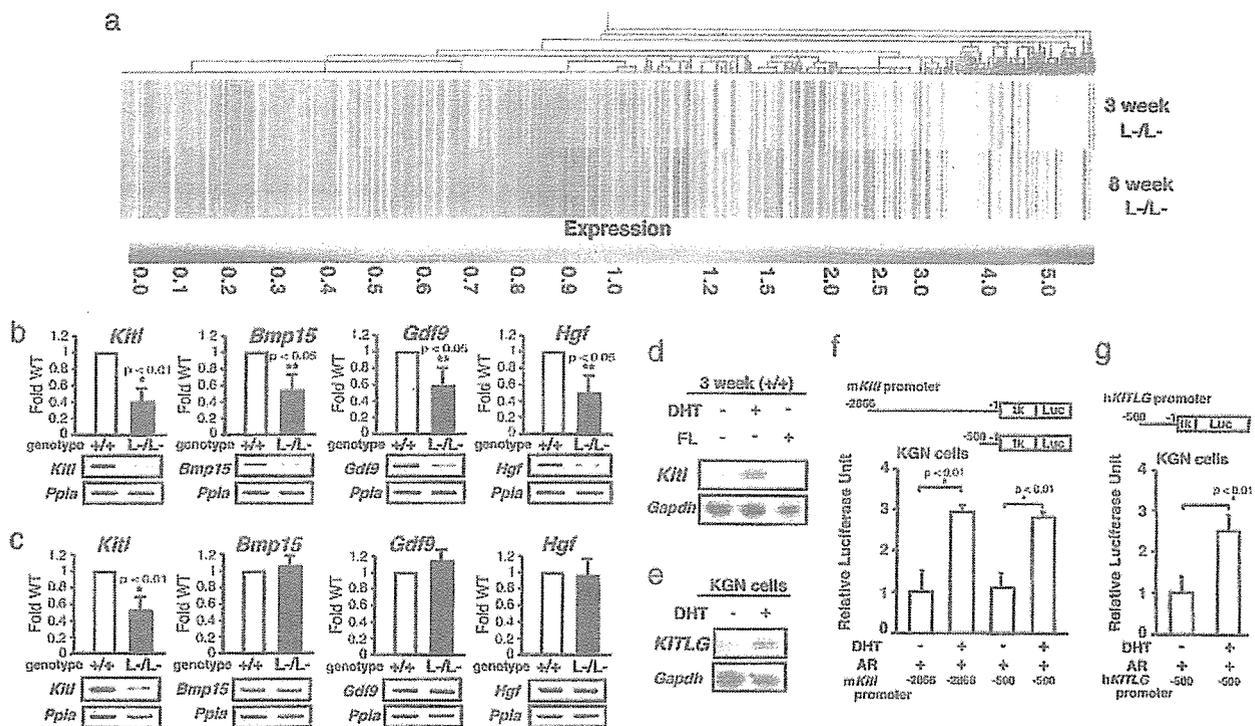
PHYSIOLOGY

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. 4a; 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. 4b). 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. 4a; 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. 4c). 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. 4d). The induction of endogenous human *kit ligand* (*KITLG*) gene by DHT was also observed in human granulosa-like tumor cells (KGN) in culture (Fig. 4e). 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.

1. Laml, T., Preyer, O., Umek, W., Hengstschläger, 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.

PHYSIOLOGY

— 668 —

Methodology article

"Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays

Jun Kanno*[†1], Ken-ichi Aisaki[†1], Katsuhide Igarashi[1], Noriyuki Nakatsu[1], Atsushi Ono[1], Yukio Kodama[1] and Taku Nagao[2]

Address: [1]Division of Cellular and Molecular Toxicology, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan and [2]President, National Institute of Health Sciences, 1-18-1, Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

Email: Jun Kanno* - kanno@nihs.go.jp; Ken-ichi Aisaki - aisaki@nihs.go.jp; Katsuhide Igarashi - igarashi@nihs.go.jp; Noriyuki Nakatsu - n-nakatsu@nihs.go.jp; Atsushi Ono - Atsushi@nibio.go.jp; Yukio Kodama - kodama@nihs.go.jp; Taku Nagao - nagao@nihs.go.jp

* Corresponding author   †Equal contributors

Abstract

Background: Transcriptome data from quantitative PCR (Q-PCR) and DNA microarrays are typically obtained from a fixed amount of RNA collected per sample. Therefore, variations in tissue cellularity and RNA yield across samples in an experimental series compromise accurate determination of the absolute level of each mRNA species per cell in any sample. Since mRNAs are copied from genomic DNA, the simplest way to express mRNA level would be as copy number per template DNA, or more practically, as copy number per cell.

Results: Here we report a method (designated the "Percellome" method) for normalizing the expression of mRNA values in biological samples. It provides a "per cell" readout in mRNA copy number and is applicable to both quantitative PCR (Q-PCR) and DNA microarray studies. The genomic DNA content of each sample homogenate was measured from a small aliquot to derive the number of cells in the sample. A cocktail of five external spike RNAs admixed in a dose-graded manner (dose-graded spike cocktail; GSC) was prepared and added to each homogenate in proportion to its DNA content. In this way, the spike mRNAs represented absolute copy numbers per cell in the sample. The signals from the five spike mRNAs were used as a dose-response standard curve for each sample, enabling us to convert all the signals measured to copy numbers per cell in an expression profile-independent manner. A series of samples was measured by Q-PCR and Affymetrix GeneChip microarrays using this Percellome method, and the results showed up to 90 % concordance.

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.

Background

Normalization of gene expression data between different samples generated in the same laboratory using a single platform, and/or generated in different geographical
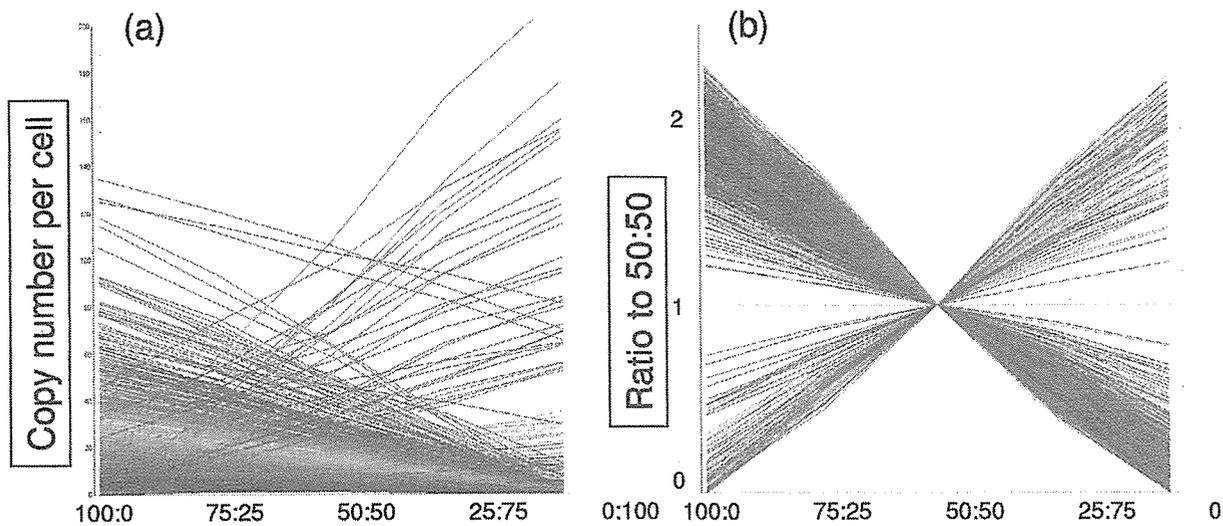
Figure 1
Dose-response linearity check by LBM. Dose-response linearity of the Affymetrix GeneChip by the LBM (liver-brain mix) sample set. Five samples, i.e. mixtures of mouse liver and brain at ratios of 100:0, 75:25, 50:50, 25:75 and 0:100, were spiked with GSC and measured by Affymetrix GeneChips Mouse430-2. Signals were normalized by the Percellome method as described in the text. Line graphs are in (a) copy numbers and (b) ratio to 50:50 sample for the top 1,000 probe sets with coefficient of correlation (R^2) closest to 1 among those having 1 copy or more per cell in the 50:50 sample (19,979 probe sets out of 45,101). The number of probe sets with $R^2 > 0.950$ was 8,655, and $R^2 > 0.900$ was 11,719.

regions using multiple platforms, is central to the establishment of a reliable reference database for toxicogenomics and pharmacogenomics. Transforming expression data into a "per cell" database is an effective way of normalizing expression data across samples and platforms. However, transcriptome data from the quantitative PCR (Q-PCR) and DNA microarray analyses currently deposited in the database are related to a fixed amount of RNA collected per sample. Variations in RNA yield across samples in an experimental series compromise accurate determination of the absolute level of each mRNA species per cell in any sample. Normalization against housekeeping genes for PCRs, and global normalization of ratiometric data for microarrays, is typically performed to account for this informational loss. Additional methods, such as the use of external mRNA spikes, reportedly improve the quality of data from microarray systems. For example, Holstege et al. [1] described a spike method against total RNA, based on their finding that the yields of total RNA from wild type and mutant cells were very similar. Hill et al. [2] reported a spike method against total RNA for normalizing hybridization data such that the sensitivities of individual arrays could be compared. Lee et al. [3] demonstrated that "housekeeping genes" cannot be used as a reference control, and van de Peppel et al. [4] described a normalization method of mRNA against total RNA using an external spike mixture. To achieve satisfactory performance they used multiple graded doses of external spikes, covering a wide range of expression, in order to align the ratiometric data by Lowess normalization [5]. Hekstra et al. [6] presented a method for calculating the final cRNA concentration in a hybridization solution. Sterrenburg et al. [7] and Dudley et al. [8] reported the use of common reference control samples for two-color microarray analyses of the human and yeast genomes, respectively. These are pools of antisense oligo sequences against all sense oligos present on the microarray. Instead of antisense oligos, Talaat et al. [9] used genomic DNA as a common reference control in studies of *E. coli*. Statistical approaches have been proposed for ratiometric data to improve inter-microarray variations, especially of non-linear relations [10]. However, because control samples may differ among studies, ratiometric data cannot easily be compared across multiple studies unless a common reference, such as a mixture of all antisense counterparts of spotted sense sequences is used [7-9]. Nevertheless, as long as the normalization is calibrated to total RNA, variations in total RNA profile cannot be effectively cancelled out. Although
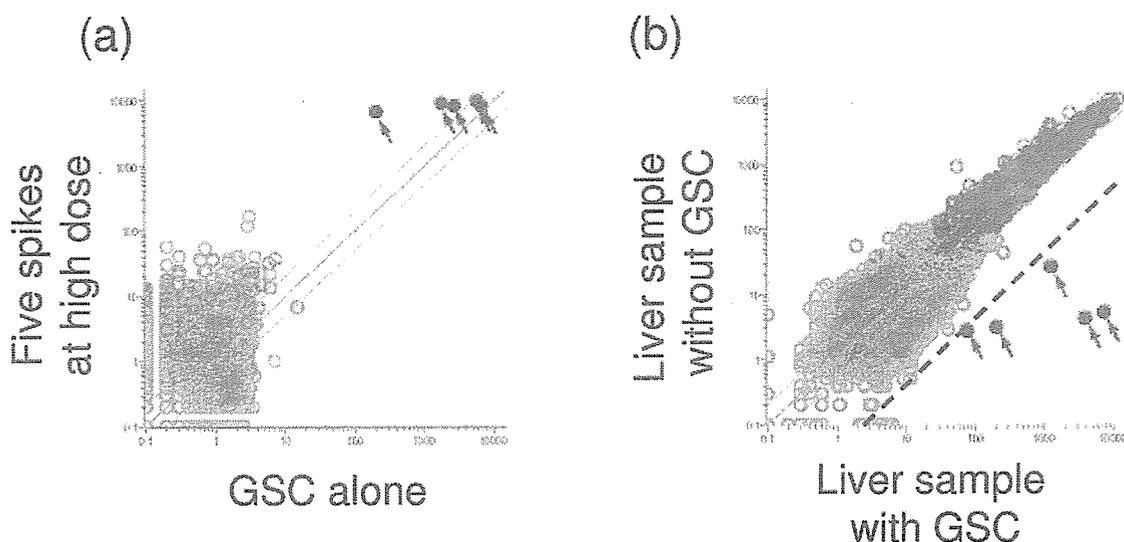
Figure 2
Cross-hybridization of GSC. Cross-hybridization of the GSC spike mRNAs to Affymetrix GeneChip. (a) A scatter plot of a blank sample with the GSC (horizontal axis) and a blank with the five spike RNAs at a high dosage (vertical axis) measured by MG-U74v2A GeneChips (raw values generated by Affymetrix MAS 5.0 software). The five spikes are indicated by black dots with arrows. Signals of the murine probe sets were below 20 on the horizontal axis, indicating negligible cross-hybridization of GSC spike mRNAs to the murine probe sets. (b) A scatter plot of a liver sample with GSC (horizontal axis) and without GSC (vertical axis) measured by MG-U74v2A GeneChips. The five spikes are again indicated by black dots with arrows. The dotted line is the 1/25 fold (4%) line. Cross-hybridization of mouse liver mRNAs to the GSC signals was considered negligible (less than 4%).

some of these reports share the idea that "absolute expression" and "transcripts per cell" should entail robust normalization, further practical development to enable universal application has been awaited.

Here, we report a method for normalizing expression data across samples and methods to the cell number of each sample, using the DNA content as indicator. This normalization method is independent of the gene expression profile of the sample, and may contribute to transcriptome studies as a common standard for data comparison and interchange.

Results
Dose-response linearity of the measurement system as a basis for the Percellome method
The fidelity of transcript detection is the key to this "per cell" based normalization method, which generates transcriptome data in "mRNA copy numbers per cell". The Q-PCR system was tested by serially diluting samples to confirm the linear relationship between Ct values and the log

of sample mRNA concentration (data not shown). High density oligonucleotide microarrays from Affymetrix [11] were used in our experiments. We tested the linearity of the Affymetrix GeneChips using a set of five samples made of mixtures of liver and brain in ratios of 100:0, 75:25, 50:50, 25:75, and 0:100 (designated "LBM" for liver-brain mix). The results showed a linear relationship ($R^2 > 0.90$) between fluorescence intensity and input for a sufficient proportion of probe sets, i.e. about 37% of the probe sets in the older MG-U74v2 and 70% in the newest Mouse Genome 430 2.0 GeneChip were above the detection level (approximately one copy per cell) in the 50:50 sample (Figure 1) [see Additional files 1 and 2].

Dose-response linearity alone is not sufficient to generate true mRNA copy numbers. An important additional requirement is that the ratio of signal intensity to mRNA copy number should be equal among all GeneChip probe sets of mRNAs and PCR primers. The Q-PCR primer sets were designed to perform at similar amplification rates to minimize differences between amplicons. The melting

Table 1: The spike factors for various organs/tissues

Species	Organ/Tissue (adult, unless otherwise noted)	Spike Factor	total RNA/genomic DNA	SD
Mouse	Liver	0.2	211	46
Mouse	Lung	0.02	22	4
Mouse	Heart	0.05	-	-
Mouse	Thymus	0.01	8	2
Mouse	Colon Epitherium	0.05	105	30
Mouse	Kidney	0.1	-	-
Mouse	Brain	0.1	-	-
Mouse	Suprachiasmatic nucleus (SCN)	0.1	-	-
Mouse	Hypothalamus	0.1	63	4
Mouse	Pituitary	0.1	52	8
Mouse	Ovary	0.02	35	4
Mouse	Uterus	0.02	42	12
Mouse	Vagina	0.02	81	38
Mouse	Testis	0.15	56	7
Mouse	Epididymis	0.07	53	16
Mouse	Bone marrow	0.02	14	3
Mouse	Spleen	0.02	-	-
Mouse	Whole Embryo	0.15	97	36
Mouse	Fetal Telencephalon E10.5–16.5	0.1	48	9
Mouse	Neurosphere (E11.5–14.5)	0.03	42	10
Mouse	E9.5 embryo heart	0.15	58	15
Mouse	cell lines	0.2	-	-
Rat	Liver	0.2	-	-
Rat	Kidney	0.2	-	-
Rat	Uterus	0.04	56	5
Rat	Ovary	0.04	56	9
Human	Cancer Cell Lines	0.2	116	26
Xenopus	liver	0.03	-	-
Xenopus	embryo	0.15	-	-

temperature was set between 60° and 65°C with a product size of approximately 100 base pairs using an algorithm (nearest neighbor method, TAKARA BIO Inc., Japan), and the amplification co-efficiency (E) was set within the range 0.9 ± 0.1 ($E = 2^{\{-(1/slope)\}}-1$ on a plot of log2 (template) against Ct value). For the GeneChip system, the signal/copy performance of each probe set depended on the strategy of designing the probes to keep the hybridization constant/melting temperature within a narrow range, ensuring that the dose-response performances of the probe sets were similar (cf. http://www.affymetrix.com/technology/design/index.affx). Failing this, any differences should at least be kept constant within the same make/version of the GeneChip. Taking into consideration the biases that lead to imperfections in estimating absolute copy numbers in each gene/probe set, we developed normalization methods to set up a common scale for Q-PCR and Affymetrix GeneChip systems.

The grade-dosed spike cocktail (GSC) and the "spike factor" for the Percellome method

A set of external spike mRNAs was used to transfer the measurement of cell number in the sample (as reflected by its DNA content) to transcriptome analysis. For the spikes, we utilized five *Bacillus subtilis* mRNAs that were left open for users in the Affymetrix GeneChip series. The extent to which the *Bacillus* RNAs cross-hybridized with other probe sets was checked for the Affymetrix GeneChip system. The GSC was applied to Murine Genome U74Av2 Array (MG-U74v2) GeneChips with or without a liver sample. As shown in Figure 2, cross-hybridization between *Bacillus* RNAs and the murine gene probe sets was negligible [see Additional files 3 and 4]. Mouse Genome 430 2.0 Array (Mouse430-2), Mouse Expression Arrays 430A (MOE430A) and B (MOE430B), Rat Expression Array 230A (RAE230A), *Xenopus laevis* Genome Array and Human Genome U95Av2 (HG-U95Av2) and U133A (HG-U133A) Arrays sharing the same probe sets for these spike mRNAs showed no sign of cross-hybridization with the *Bacillus* probes (data not shown).

We prepared a cocktail containing in vitro transcribed *Bacillus* mRNAs in threefold concentration steps, i.e. 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) and 9.6 pM (for AFFX-TrpnX-3_at). By referring to the amount of DNA in a diploid cell and employing a "spike factor" determined by the ratio of
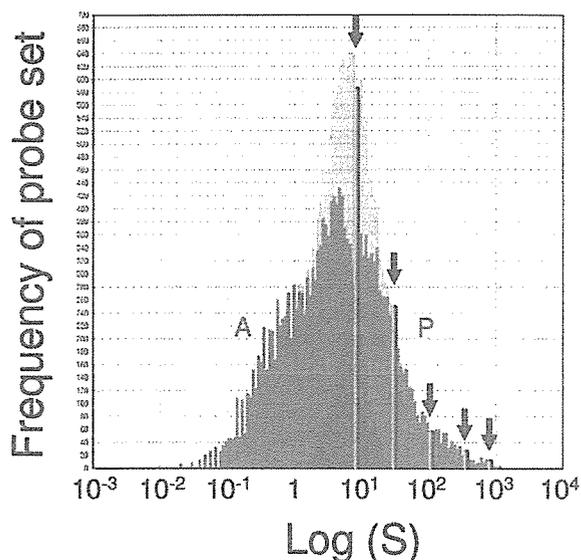
Figure 3
Positioning of GSC spike mRNAs in Affymetrix GeneChip dose-response range. A frequency histogram of the probe sets of Affymetrix GeneChip Mouse430-2 is shown. The histogram for all probe sets (gray) shows near-normal distribution. Blue columns are the "present" calls (P), red columns "absent" calls (A) and green "marginal" calls. The five yellow lines indicate the positions of the GSC spike mRNAs that are chosen to cover the "present" call range by a proper "spike factor".

total RNA to genomic DNA in a tissue type (Table 1), the spike mRNAs were calculated to correspond to 468.1, 156.0, 52.0, 17.3 and 5.8 copies per cell (diploid), respectively, for the mouse liver samples (spike factor = 0.2). The ratio of mRNAs in the cocktail is empirically chosen depending on the linear range of the measurement system and the available number of spikes. Here, we set the ratio to three to cover the "present" call probe sets of the Affymetrix GeneChip system (Figure 3).

We tested this grade-dosed spike cocktail (GSC) by Q-PCR and confirmed that the Ct values of the spike mRNAs were linearly related to the log concentrations (cf. Figure 4a), i.e. could be expressed as

$$Ct = \alpha \log C + \beta \quad \{1\}$$

The GSC was also tested by the GeneChip system and it was confirmed that the log of the spike mRNA signal intensities was linearly related to the log of their concentrations (cf. Figure 4b),

$$\log S = \gamma \log C + \delta \quad \{2\}$$

The linear relationship between the Ct values (Ct) and the log of RNA concentration (log C) was reasonable given the definition of Ct values (derived from the number of PCR cycles, i.e. doubling processes). The linear relationship between the log of GeneChip signal intensity (log S) and the log of RNA concentration (log C) was rationalized by the near-normal distribution of log S over all transcripts (cf. Figure 3).

Calculation of copy numbers of all genes/probe sets per cell
As described above, using a combination of DNA content and the spike factor of the sample, the GSC spike mRNAs become direct indicators of the copy numbers (C') per cell. When the samples were measured by Q-PCR or GeneChip analysis, the five GSC spike signals in each sample should obey function {1} for Q-PCR and function {2} for GeneChip with a good linearity. If the observed linearity was poor, a series of quality controls was performed and the measurement repeated. The coefficients of the functions were determined for each sample by the least squares method. Under the assumption that all genes/probe sets share the same signal/copy relationship, signal data for all genes/probe sets were fitted to the functions {1'} or {2'}, which are the individualized functions of {1} and {2} for each sample measurement (i).

$$Ct = \alpha i \log(C') + \beta i \quad \{1'\}$$

$$\log (S) = \gamma i \log(C') + \delta i \quad \{2'\}$$

(i = sample measurement no.)

The Q-PCR Ct values (Ct) and microarray signal values (S) of all mRNA species in the sample (i) are converted to copy numbers per cell (C') by the inverses of functions {1'} and {2'}, i.e. {3} and {4} below:

$$C' = B^{\wedge}((Ct-\beta i)/\alpha i) \quad \{3\}$$

for Q-PCR (Figure 4a).;

$$C' = B^{\wedge}((\log S - \gamma i)/\delta i) \quad \{4\}$$

for GeneChips (Figure 4b),

where B is the logarithmic base used in {1} and {2} (see Materials and Methods for details).

Real world performance of the Percellome method
The correspondence between Q-PCR and GeneChip was tested using a sample set from 2,3,7,8-tetrachlorodibenzodioxin (TCDD)-treated mice. Sixty male C57BL/6 mice
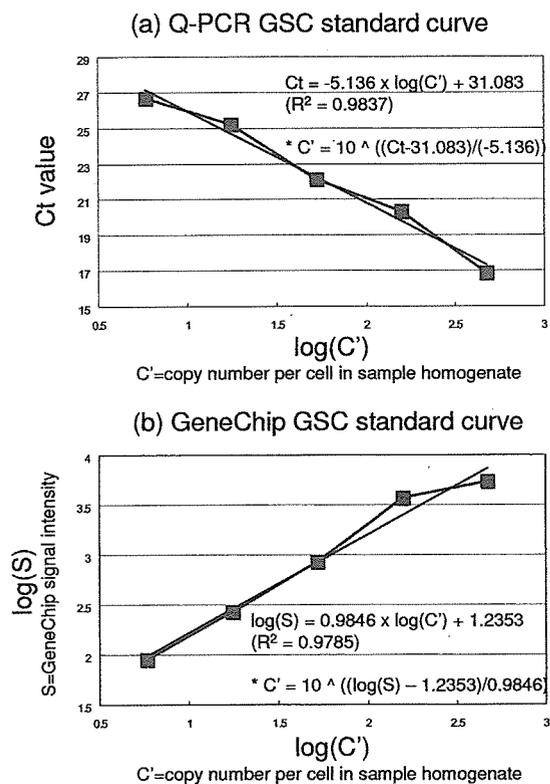
(a) Q-PCR GSC standard curve

$Ct = -5.136 \times \log(C') + 31.083$
$(R^2 = 0.9837)$

$* \; C' = 10 \wedge ((Ct-31.083)/(-5.136))$

Ct value

log(C')

C'=copy number per cell in sample homogenate

(b) GeneChip GSC standard curve

S=GeneChip signal intensity

log(S)

$\log(S) = 0.9846 \times \log(C') + 1.2353$
$(R^2 = 0.9785)$

$* \; C' = 10 \wedge ((\log(S) - 1.2353)/0.9846)$

log(C')

C'=copy number per cell in sample homogenate

Figure 4
The dose-response linearity of the GSC spikes in Q-PCR and the Affymetrix GeneChip array system. Linear relationships are shown between (a) the Q-PCR Ct values and log of copy number (log (C')), and (b) the GeneChip log signal intensity (log(S)) and log of copy number (log (C')) of the GSC mRNAs. The regression functions were obtained by the least squares method. The inverse functions (*) were further used to generate the copy numbers of all other genes/probe sets for Percellome normalization.

were divided into 20 groups of 3 mice each. TCDD was administered once orally at doses of 0, 1, 3, 10 and 30 μg/kg, and the livers were sampled 2, 4, 8 and 24 h after administration. Nineteen primer pairs were prepared for Q-PCR and the Ct values of the liver transcriptome were measured. The same 60 liver samples were measured using the Affymetrix Mouse430-2 GeneChip [see Additional files 5 through 8 and 9 through 12]. Q-PCR and GeneChip data were normalized against cell number by functions {3} and {4}, respectively. The averages and standard deviations (sd) of each group (n = 3) were calculated and plotted as three layers of isoborograms on to 5 × 4 matrix three-dimensional graphs (Figure 5). Together with another sample set (data not shown), a total of thirty-six primer pairs were compared, and there was a

correlation of up to 90% between the Q-PCR and Gene-Chip surfaces. It is notable that not only the average surfaces but also the +1sd and -1sd surfaces corresponded closely in shape and size. We infer that the differences resulted mainly from biological variations among the three animals in each experimental group rather than from measurement error (cf. Figure 7).

An important feature of Percellome normalization is its independence from the overall expression profile of the sample. When gene expression profiles differ among samples, Percellome normalization produces a robust transcriptome that is different from total-RNA dependent global normalization. As an example, Figure 6 shows the results of an experiment on the uterotrophic response of ovariectomized mice to estrogen treatment [12] [see Additional files 13 and 14]. The uteri of the vehicle control are atrophic because the ovaries, the source of intrinsic estrogens, are absent. The uteri of the treated groups are hypertrophic owing to estrogenic stimulus from the test compound administered. Global normalization (90 percentile) between the vehicle control group and the high-dose (1,000 mg/kg) group indicated that 4,600 of 12,000 probe sets showed 2-fold or greater increase, 470 were reduced by 0.5 or less, and 7,400 remained between these extremes. In contrast, analysis of Percellome-normalized data revealed that almost all the 12,000 probe sets showed a 2-fold or greater increase, including actin, GAPDH and other housekeeping genes. The hypertrophic tissues, consisting of cells with abundant cytoplasm, provide convincing evidence for the increases in various cellular components including housekeeping gene products.

Another important feature of Percellome normalization is the commonality of the expression scale across platforms. Batch conversion can be performed between results obtained from different platforms when the data are generated by the Percellome method. A practical strategy for such normalization is to prepare a set of samples from a target organ of interest with differences in gene expression, and measure them once by each platform. Data conversion functions with good linear dose-response relationships can be obtained individually for those genes/probe sets that are measured by both platforms (Figure 7).

Discussion

We have developed a novel method for normalizing mRNA expression values to sample cell numbers by adding external spike mRNAs to the sample in proportion to the genomic DNA concentration. For non-diploid or aneuploid samples, an average DNA content per cell should be determined beforehand for accurate adjustment. When there is significant DNA synthesis, a similar adjustment should be considered.
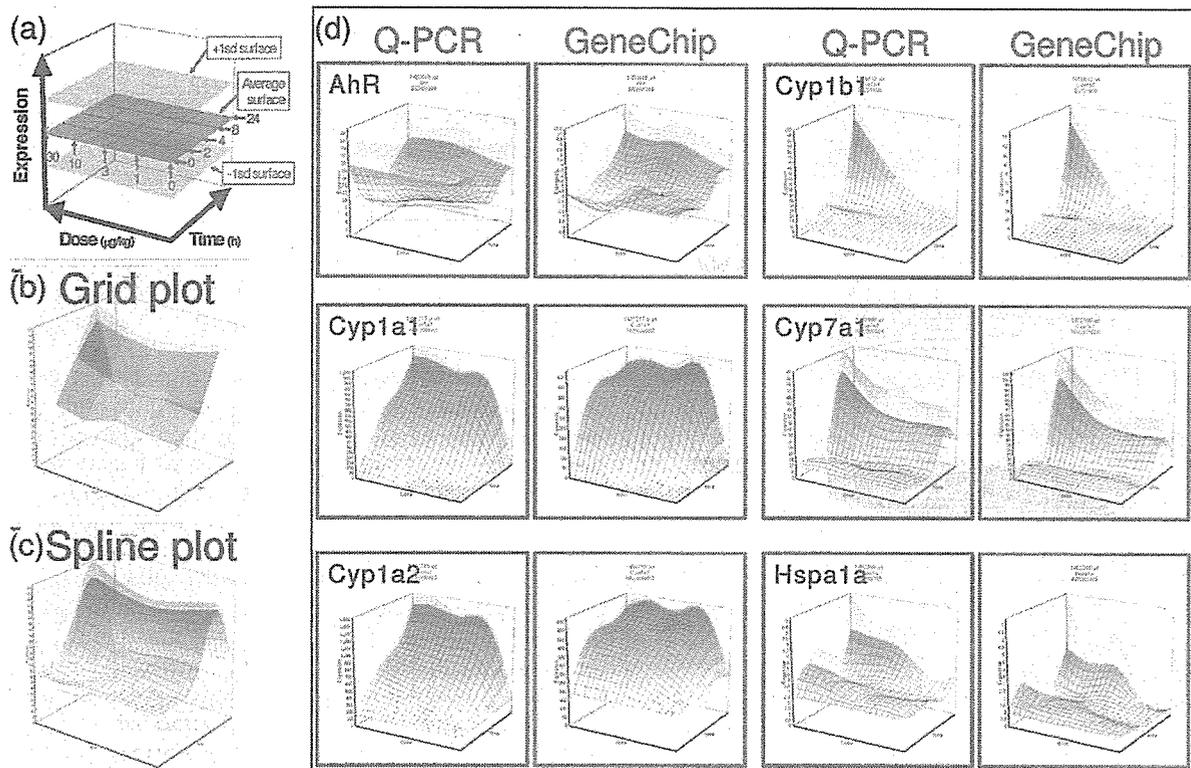
Figure 5
Correspondence between Q-PCR and GeneChip data. Sixty male C57BL/6 mice were divided into 20 groups of 3 mice each. 2,3,7,8-tetrachlorodibenzodioxin (TCDD) was administered once orally at doses of 0, 1, 3, 10 and 30μg/kg, and the liver was sampled 2, 4, 8 and 24 h after administration. The liver transcriptome was measured by the Affymetrix Mouse430-2 Gene-Chip. 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 (arylhy-drocarbon 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 micro-

dissection (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
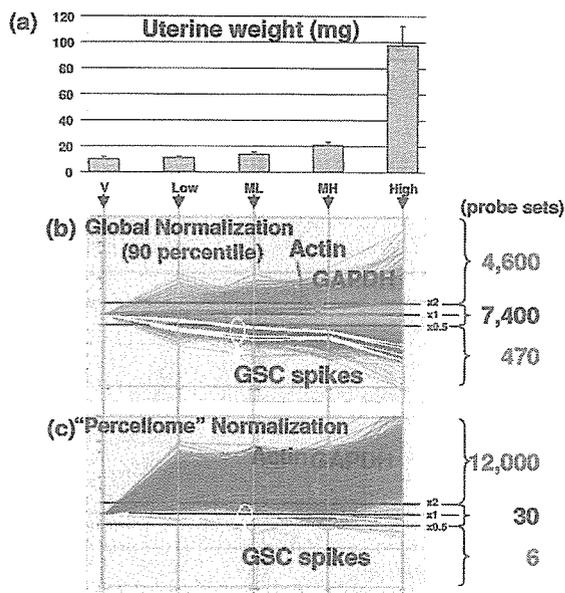
Figure 6
Uterotrophic response of ovariectomized female mice by an estrogenic test compound.(a) Shows the uterine weight, which increases in a dose-dependent manner; V, vehicle control; Low, low dose; ML, medium-low dose; MH, medium-high dose; High, high dose group. (b) Shows the line display of uterine 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).
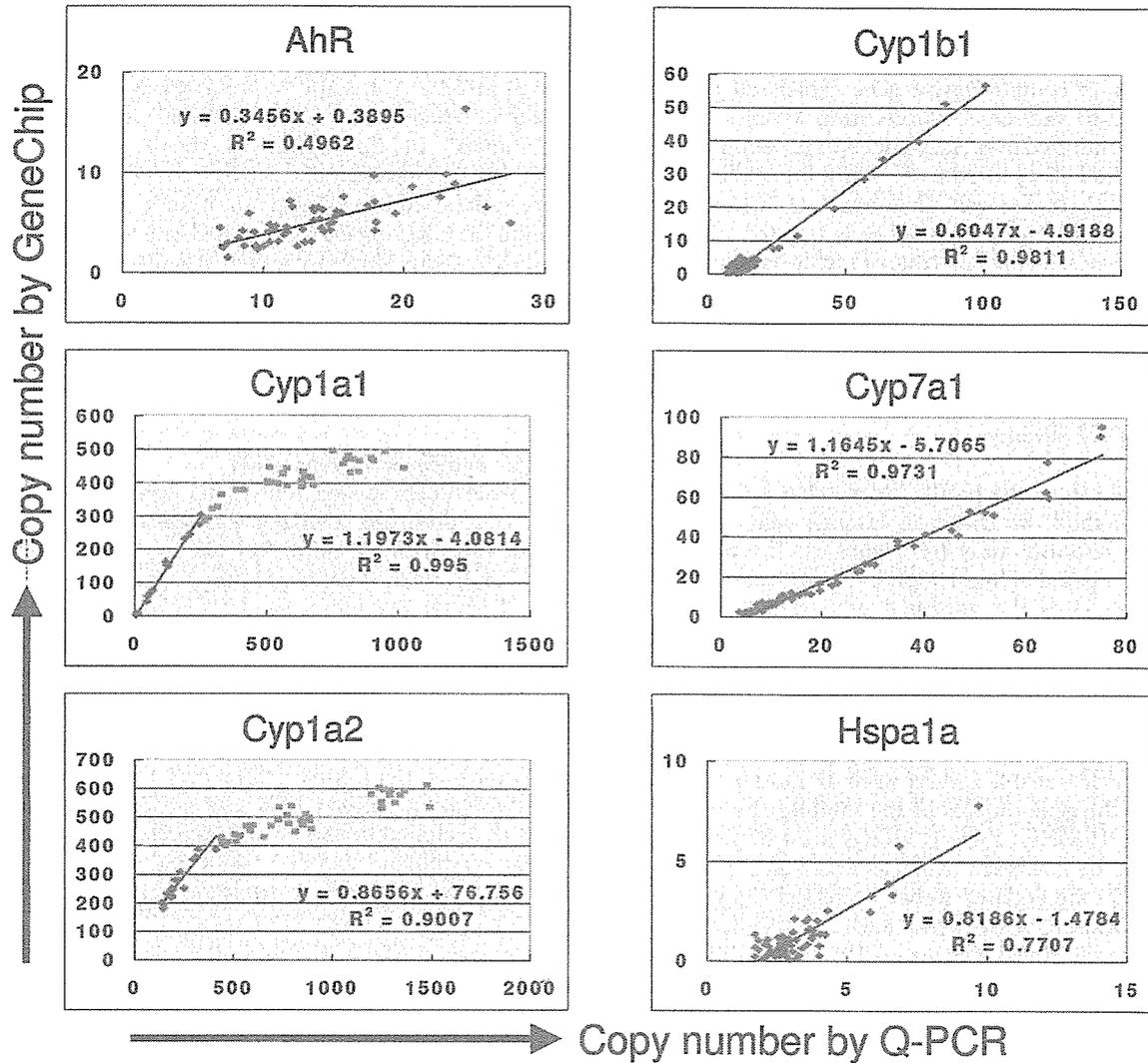
Figure 7
Conversion functions between Q-PCR and GeneChip. The data shown in Figure 5 as 3D surfaces are shown as a scatter plot (60 plots). The regression function can be used to convert Q-PCR to GeneChip and vice versa, with a level of certainty indicated by coefficient of correlation. It is noted that 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 Gene-Chip 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

$$G = C * v * f \quad (1),$$

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

$N = C/d$ (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,

$CSj = Sj*C*f$ (mole/l).

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

$MSj = CSj/N$

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

$= Sj*f*d$ (mole/cell)

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

$NSj = MSj*A$

$= Sj*f*d*A$ (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}.