library, in combination with next-generation sequencing technology as the detection platform (Sims et al, 2011), to identify key subtype-specific regulators of cancer cell proliferation and/or survival. The relevance of such subtype-specific targets has been exemplified by *ESR1* (estrogen receptor α) for luminal-subtype breast cancers; these cancers share not only clinical features such as prognosis and the response to chemotherapy, but also the pattern of gene expression. *ESR1* has been used not only for diagnosis but also as a molecular target to treat cancer patients with this subtype (Howell, 2013; Sorlie et al, 2001). Importantly, in this study, specific growth determinants were distinguished amongst the ovarian cancer subtypes at the genome-wide as well as gene level. This observation supports the potential for subtype-specific therapeutic options in treating ovarian carcinoma and reinforces the clinical importance of the classification scheme proposed in this study.

Although the molecular mechanisms linking *TUBGCP4* or *NAT10* with Stem-A growth remains to be elucidated, susceptibility to vincristine and vinorelbine underscores the importance of tubulin polymerization in Stem-A cells. Both drugs are well-established chemotherapeutic agents that block cell proliferation by inhibiting microtubule assembly through its interaction with tubulin heterodimers (Lobert et al, 1996); however, they are not standard chemotherapeutic reagents for the treatment of EOC, unlike paclitaxel (Armstrong et al, 2006; McGuire et al, 1996). The molecules implicated in the tubulin polymerization pathway may provide us with a potential platform to more effectively target Stem-A ovarian cancer. As such, the survival of patients with ovarian cancer could be improved by the stratification and targeting strategy described in this study.

MATERIALS AND METHODS

Eligibility criteria and quality control of expression data

In order for our study to make broader generalizations and attain a larger sample size, reduced eligibility criteria were adopted (George, 1996). Female adult (age ≥20 years) patients with a clinical diagnosis of primary or metastatic ovarian cancer were included in our analysis. We imposed no limit on patient race, pre-treatment history or medical conditions, or on the stages, grades, and histology of the disease.

To control for the quality of expression data, we checked the quality of the Affymetrix chips (Affymetrix, Santa Clara, CA) using Bioconductor AffyQCReport package (Gautier et al, 2004) and the following criteria: average perfect-match (Neve et al, 2006) intensity, kernel density plot, GAPDH 3′:5′ ratio, β-actin 3′:5′ ratio, and centre of intensity for positive and negative controls. All chips passed at least one of the criteria, and hence, none of the samples was discarded.

Data preprocessing of Affymetrix expression data

Ovarian cancer datasets were downloaded from multiple data repositories: Gene Expression Omnibus (GEO), Array Express, Expression Project for Oncology (ExpO), and The Cancel Genome Atlas (TCGA). Microarray data on Affymetrix U133A or U133Plus2 platforms were utilized for the analysis. Robust Multichip Average (RMA) normalization was performed on each dataset. ComBat (Johnson et al, 2007), a high precision and accurate technique for removing batch effect while conserving meaningful variation (Chen et al, 2011), was applied for batch adjustment on the compiled, normalized data. Removal of ovarian cancer cell lines, normal tissues and primary cultured normal cells from the batch-adjusted data yielded a dataset of 1538 ovarian tumour samples, predominantly composed by EOCs (Supporting Information Table 15A). Probes (1185) corresponding to 941 genes (Supporting Information Table 2) were retained by applying a threshold of standard deviation across samples >1.05. Expression values of selected genes were normalized and centred with Cluster 3.0 and further processed for subtype identification. An additional validation dataset of 418 samples were similarly collected and subjected to the same preprocessing procedure. Clinical information of the validation dataset is given in Supporting Information Table 15B.

Consensus clustering

CC (Monti et al, 2005) using Gene Pattern software (Reich et al, 2006) was employed to identify robust clusters corresponding to the distinct subgroups in EOC. We chose hierarchical clustering with agglomerative average linkage, with Euclidean distance and a sub-sampling ratio of 0.8 for 1000 iterations. The condition of $K_{max} = 18$ was employed, as it gave a reasonable Gini index and purity of ~0.8. "Other" was used to indicate the unclassified samples not grouped in any of the five subtypes in the initial CC analysis shown in Fig 1A. They were not included in following statistical analyses for characterization of the molecular subtypes.

Univariate and multivariate Cox regression analysis

From 845 samples with overall survival information, we extracted 537 samples from three institutions (GSE3149: 5, GSE9891: 241 and TCGA: 291) with clinical variables (Table 1). This information was transformed to binary information (presence/absence of a phenotype) prior to assessment of their prognostic association with overall survival by Cox proportional hazards regression analysis (Therneau & Grambsch, 2000). The same procedure was applied for Cox proportional hazards regression analysis for progression-free survival. We extracted 518 samples (GSE9891: 199 and TCGA: 319) from 596 samples with progression-free survival information. Univariate and multivariate Cox regression were performed using R (http://www.R-project.org). Multivariate analyses with clinical variables were conducted independently for each subtype.

Statistical analysis for clinical parameters

GraphPad Prism was used to examine statistical significance of clinical stage, primary or metastatic tumours, histological subtypes, or the malignant potential of each subtype by Fisher's exact test. For Kaplan–Meier analyses, the statistical significance was calculated by log-rank test.

Subtype-specific gene set enrichment

A total of 6898 gene sets were collected (Supporting Information Table 6). The ss-GSEA score (Verhaak et al, 2010) was computed to estimate the pathway activity for all 1538 ovarian cancer samples or 142 cell lines for each gene set. Based on the computed ss-GSEA

score, a binary comparison was conducted for each subtype to identify subtype-specific pathway enrichment. Gene sets specifically and significantly enriched in a subtype were selected using SAM (FDR $q = 0$) and ROC (ROC > 0.85 as overexpressed gene sets).

Predictive modelling and validation by BinReg

Expression data analysis, based on a binary regression model using the BinReg ver. 2.0, was described previously (Gatza et al, 2010). BinReg uses a Bayesian statistical analysis to fit a binary probit regression model on training data given a set of genes that are most correlated with the binary response/phenotype of interest (*e.g.* Epi-A *vs.* Non-Epi-A). The regression coefficients of these genes indicate the discriminating power of the genes and are weights for the overall meta-gene profile. The overall meta-gene profile is used for comparison and predicts the status of the phenotype of the new sample or dataset. In this study, we built a binary regression model for each subtype that singled out a subtype from the rest (*i.e.* Epi-A *vs.* Non-Epi-A) and adopted a divide-and-conquer approach for generating signatures for each of the different subtypes (Supporting Information Figs 7A and B). Briefly, the top 50 core samples were selected by their highest SW of all five subtypes, and subdivided into two sets of data: training set A and training set B. These training sets were utilized to determine appropriate parameters (Supporting Information Table 17; Supporting Information Materials and Methods) for the binary regression model. Subsequently, the condition was used to predict the remaining samples by training set A. To predict the status of the phenotype on a dataset, a Bayesian probit regression model was fit to assign the probability that the sample exhibited evidence of a phenotype, based on the concordance of its gene expression values with the signature (Gatza et al, 2010).

Expression microarrays of cultured cell lines

Most of Duke, Kyoto and Singapore cell lines were derived from an original collection assembled in a Duke laboratory (Supporting Information Table 11) (Matsumura et al, 2011). Therefore, expression data for these 28 cell lines from the collection could be used as biological replicates. We extracted RNA from 34 cultured EOC cell lines (ovary1847, JHOS-2, OAW28, OAW42, OV7, OV17R, OV56, Caov-2, OV90, OVCA420, OVCA429, OVCA432, OVCA433, OVCAR-2, OVCAR-3, OVCAR-5, OVCAR-8, OVCAR-10, Caov-3, SKOV-3, UWB1.289, A2008, EFO-21, C13, OV2008, FU-OV-1, IGROV-1, TOV-112D, A2780, CH1, DOV 13, TYK-nu, PEO1 and COLO720E) (Methods: Cell line phenotypes *in vitro*) and performed expression assays with Affymetrix Human U133 Plus 2.0 arrays. The data were deposited in Gene Expression Omnibus (GEO) with the accession of GSE28724. Details of the EOC cell lines are given in Supporting Information Table 11.

Cell line phenotypes *in vitro*

Cell lines were cultured in RPMI 1640 media (Invitrogen, Carlsbad, CA) with 10% foetal bovine serum (#S1810-500; Biowest, Nuaillé, France). Measurements of population doubling time and colony formation assays in methylcellulose were described previously (Huang et al, 2008; Liu et al, 2008a; Matsumura et al, 2011; Mori et al, 2009). Mann–Whitney *U*-test of GraphPad Prism was used to statistically evaluate the numerical values for the cell line phenotypes across the subtypes.

Lentivirus library infection and shRNA retrieval by PCR of the genomic DNA

Fourteen cell lines representing Epi-A, Mes or Stem-A were chosen based on the SW for the subtype signature so as to have "more representative" cell lines for a given subtype, and these cell lines were used for shRNA screening. We used a pooled library of shRNA-expressing lentiviruses (80,000 clones targeting 16,000 genes per library, TRC1.0, #CSTVRS; Sigma–Aldrich, St Louis, MA). Optimal lentiviral infection conditions achieved a multiplicity of interest (MOI) of 0.3 to ensure the highest probability of having single shRNA integration into the host genome in each cell (Luo et al, 2008). Each lentiviral vector encodes each shRNA expression cassette with the puromycin resistance gene, allowing the use of puromycin to isolate stable integrants. Under selection pressure from puromycin (5 μg/ml), infected cells were allowed to propagate for ~14 days (~4 or 5 passages), whereby cells expressing shRNA that silence genes that were required for and known to suppress cell growth were depleted from and enriched in the culture, respectively. Hence, the abundance of each shRNA (=shRNA copy number) is reflective of the effect of an shRNA on cell growth. At the endpoint of the incubation, genomic DNA was harvested from the resulting cells by PureLink Genomic DNA kits (#K1820-01, Invitrogen). The integrated shRNA sequences were retrieved from the genomic DNA (100 ng) by PCR amplification using vector primers (shRNA Forward Primer: 5′-atcttgtggaaaggacgaaac-3′ and shRNA Reverse Primer: 5′-tactgccatttgtctcgaggt-3′) with KOD Plus ver. 2 (#KOD-211, Toyobo) and 28–32 cycles of 98°C for 10 s, 56°C for 30 s, and 68°C for 1 min. Products were purified with QIAQuick PCR Purification Kit (#28106, Qiagen, Hilden, Germany).

Next-generation sequencing analysis by Solexa to count copy numbers of individual shRNAs

Amplified DNA (20 ng) from PCR was used to construct a sequencing library using a ChIP-Seq sample preparation kit (#IP-102-1001, Illumina, San Diego, CA). The two sample-multiplexing sequencing method was used individually, with multiplexing index 6 and index 12 primers for each sample (Illumina, #PE-400-1001). Constructed libraries were subjected to a final size-selection step on a 10% Novex TBE gel (#EC6275BOX, Invitrogen, Carlsbad, CA). DNA fragments of 205 bp were excised, recovered and quantified following Illumina's qPCR quantification protocol and guides. Quantified libraries were then sequenced on the Genome Analyzer IIx (Illumina) using the multiplexing single-end sequencing protocol at a length of 58 + 7 bp (#PE-400-2002, Illumina). Image analysis and base calls were performed using the default settings. After stripping off the PCR primer sequences, reads were then aligned to the shRNA library using Bowtie with the specified settings: –solexa1.3-quals -n 0-l 5 -v 0 -k 1 -m 1–best –strata -y –nomaqround. The data were deposited in GEO with the accession of GSE45420.

Statistical identification of the functionally relevant genes in a subtype-specific manner

Using reads with a perfect match to the reference sequences (Sigma–Aldrich), copy number was counted and normalized by total number of reads in a sample. RNAi gene enrichment ranking (RIGER) was used to find phenotype-specific, functionally relevant genes from the scale-normalized copy number count data (Luo et al, 2008). Among 80,000 hairpins included in the library, next-generation sequencing analyses

The paper explained

PROBLEM:

Epithelial ovarian cancer exhibits considerable heterogeneity, which may lead to poor survival rates for patients treated with standard chemotherapeutic regimens. This has prompted the need for a robust classification scheme to unravel this heterogeneity and allow for the development of personalized treatment strategies.

RESULTS:

A large collection of gene expression data enabled the identification of five distinct subgroups of ovarian carcinoma. The existence of these five subgroups was validated in an independent collection. Genome-wide shRNA screening against a panel of ovarian carcinoma cell lines revealed two subtype-specific targets and the pathways that control cancer cell growth.

IMPACT:

We identified five distinct subgroups, allowing rational patient stratification. Subsequent assays uncovered genes and deregulated pathways, which will be instrumental in guiding future therapeutic strategies for ovarian cancer.

detected 60,002 and 65,533 shRNA hairpins in two independent screenings and 57,168 hairpins were intersected in both results. We compiled and subsequently standardized these two datasets by ComBat (Johnson et al, 2007). Binary comparisons were performed on the three subtypes (*e.g.* Epi-A subtype *versus* the others). We adopted the signal-to-ratio as the metric for ranking hairpins, 1000 as the number of permutations, and Kolmogorov–Smirnov in the RIGER settings. The false discovery rate was computed using the Benjamini and Hochberg procedure. Genes were considered significant at $q < 0.005$ in Fig 3B or $q < 0.03$ for the validation study. For heatmap presentation, we retained the hairpins with a hairpin score ≥ 0.2.

Validation of functional determinants in cell growth of Stem-A cell lines by siRNAs

We selected 135 genes as Stem-A-specific growth-promoting genes for further validation via siRNA transfection from the top hit gene list from RIGER analysis of shRNA lentivirus screens ($q < 0.03$). The validation experiments were performed via a process consisting of four steps (Fig 3C). Dharmacon SMART pool siGENOME siRNA (1st and 2nd steps) and Dharmacon SMART pool ON-TARGET*plus* siRNA (OTP; 3rd and 4th steps) formats (Thermo Fisher Scientific, Lafayette, CO) were used to validate the effect of gene knockdown on cell growth of ovarian cell lines (Fig 3C). PA-1 (1st, 2nd, and 3rd steps) and CH1, A2780 and OVCAR-3 (4th step) were used as representative cell line(s) for the Stem-A subtype. As reference(s) for the subtype, HeyA8 (1st step), HeyA8 and OVCA433 (2nd and 3rd steps), OVCA429, PEO1, ovary1847, SKOV-3 and HEY (4th step) were used (Fig 3C). Cells were reverse-transfected with each individual siRNA per well in a 96-well format in the following conditions: OVCA433, 2500 cells with 0.3 μl of DF1 (T-2001); HeyA8, 800 cells with 0.08 μl of DF4 (T-2004); PA-1, 1200 cells with 0.22 μl of DF2 (T-2002); OVCA429, 1500 cells with 0.22 μl of DF4 (T-2004); PEO1, 4000 cells with 0.24 μl of DF4 (T-2004); ovary1847, 2500 cells with 0.12 μl of DF2 (T-2002); SKOV-3, 2500 cells with 0.12 μl of DF2 (T-2002); HEY, 1000 cells with 0.08 μl of DF4 (T-2004); CH1, 1800 cells with 0.17 μl of DF4 (T-2004); A2780, 2000 cells with 0.16 μl of DF1 (T-2001); OVCAR-3, 4000 cells with 0.2 μl of DF3 (T-2003, Thermo Fisher Scientific). We used two negative controls for Dharmacon SMART pool siGENOME siRNA transfection (#D-001206-13-20 and #D-001206-14-20), and one negative control for Dharmacon SMART pool ON-TARGET*plus* siRNA transfection (#D-001810-10-20). Assays were performed in quadruplicate. After 96-h incubation, an MTS assay was used to measure cell growth using a CellTiter 96 AQueous Non-Radioactive Cell Proliferation Assay following the manufacturer's recommendations (#G5430, Promega, Madison, WI). Genes were considered as Stem-A-specific growth-promoting genes when their down-regulation caused $\geq 20\%$ growth suppression on the Stem-A cell line ($p < 0.001$), and showed $\geq 20\%$ more growth suppression on the Stem-A line than on the reference cell lines.

Cell line drug sensitivity *in vitro*

Eighteen ovarian cancer cell lines (12 non-Stem-A: OVCA433, OVCA429, OVCAR-8, PEO1, OVCA432, OVCA420, HeyA8, HEY, HeyC2, SKOV-3, ovary1847 and DOV 13; 6 Stem-A: PA-1, CH1, A2780, OVCAR-3, SKOV-4 and SKOV-6) were tested for their sensitivity to paclitaxel, vincristine and vinorelbine, as described previously (Bild et al, 2006). Paclitaxel (#T7402), vincristine (#V8879) and vinorelbine (#V2264) were purchased from Sigma–Aldrich. Cells were seeded in 96-well plates at an optimal density, which was determined for each cell line to ensure that it reached 80% confluency by the end of the assay. Following an overnight incubation, cells were treated with nine concentrations of each drug (twofold dilution series over a 128-fold concentration range) for 48 h. The percentage of the cell population responding to the drug relative to the negative controls was measured using a CellTiter 96 AQueous Non-Radioactive Cell Proliferation Assay, following the manufacturer's recommendations (#G5430, Promega). Dose-response curves were plotted using GraphPad Prism, to derive a growth inhibitory concentration of 50% (GI50; drug concentration for 50% growth inhibitory effects on cells) for each cell line in at least three independent experiments. Mann–Whitney *U*-test of GraphPad Prism was used to statistically evaluate the averaged GI50s between non-Stem-A and Stem-A cell lines.

Western blotting analysis

Total cell lysates were prepared by direct lysis with RIPA buffer (#R0278, Sigma–Aldrich), supplemented with protease inhibitor cocktail (#539134, Calbiochem, Boston, MA). Protein concentrations were determined using BCA protein assay (#23225, Thermo Scientific,

Rockford, IL). Electrophoresis of the cell lysates were carried out with a BioRad Mini Protean II apparatus and transferred onto PVDF membranes (#IPFL00010, Millipore, Billerica, MA) with a BioRad Mini Trans-Blot apparatus, following the manufacturer's recommendations. Membranes were immunoblotted with primary antibodies directed against PARP (#9542, Cell Signaling, Danvers, MA), Caspase-3 (#9662, Cell Signaling) or β-actin (#A1978, Sigma–Aldrich), followed by immunoblotting with secondary IRDye 800CW conjugated goat anti-rabbit (#926-32211) or IRDye 680 conjugated goat anti-mouse antibodies (#926-32220, LI-COR Biosciences, Lincoln, NE). The western blots were scanned using an Odyssey Infrared Imaging System from LI-COR Biosciences.

Author contributions

SM conceived the idea; SM, JPT, BCG and RYH devised the project and obtained funding; SM, TZT, QHM, JTC, JPT and RYH wrote the paper; SM, QHM, JY, JAL, LZW, NM and MKW performed the experiments; MCW, LHBAH and RS performed next-generation sequencing analysis; SM, TZT and JTC performed bioinformatics analyses; MC performed clinical parameter analyses; BD and JMN provided OSLO ovarian cancer samples; NM, MM, and IK provided ovarian cancer cell lines.

Acknowledgements

We thank Drs. K. Nakayama, T. Hattori, A. Numata, T. Yokomizo, T. Inoue, T. Baba, H. Saya, M. Sudo, P. Koeffler, E. Chen, X. Fu, Y. Ito, M. Araki, J. Chi, T. Noda and J. Nevins for critical reading of the manuscript and/or helpful discussions. We thank Drs. K. Yamaguchi and S. Murphy for kindly providing us with a panel of ovarian cancer cell lines. We thank Dr. R. Jackson for her careful English editing. We thank the financial support from Cancer Science Institute of Singapore, Institute of Molecular Cellular Biology at A*STAR, the Vehicle Racing Commemorative Foundation in Japan and the Princess Takamatsu Cancer Research Fund. The Oslo cohort dataset (Ben Davidson and Jahn Nesland) was obtained and analysed under the ethics approval from the Health Region of South-Eastern Norway (# 04300).

Supporting Information is available at EMBO Molecular Medicine Online.

The authors declare that they have no conflict of interest.

For more information

National Library of Medicine (Ovarian Cancer):
http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0001891/

RNAi Consortium:
http://www.broadinstitute.org/rnai/trc/lib

BinReg
http://dig.genome.duke.edu/software.html
http://www.stat.duke.edu/~mw/ABS04/Lecture_Slides/5.Stats_BinReg.pdf

ClaNC
http://www.stat.tamu.edu/~adabney/clanc/

RIGER
http://www.broadinstitute.org/cancer/software/GENE-E/

References

Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, *et al* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403: 503-511

Anglesio MS, Arnold JM, George J, Tinker AV, Tothill R, Waddell N, Simms L, Locandro B, Fereday S, Traficante N, *et al* (2008) Mutation of ERBB2 provides a novel alternative mechanism for the ubiquitous activation of RAS-MAPK in ovarian serous low malignant potential tumors. Mol Cancer Res 6: 1678-1690

Armstrong DK, Bundy B, Wenzel L, Huang HQ, Baergen R, Lele S, Copeland LJ, Walker JL, Burger RA (2006) Intraperitoneal cisplatin and paclitaxel in ovarian cancer. N Engl J Med 354: 34-43

Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C, *et al* (2009) Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 462: 108-112

Bast RC Jr, Hennessy B, Mills GB (2009) The biology of ovarian cancer: new opportunities for translation. Nat Rev Cancer 9: 415-428

Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A, *et al* (2006) Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 439: 353-357

Blum A, Kalai A, Langford J (1999) Beating the hold-out: bounds for K-fold and progressive cross-validation. In *Proceedings of the twelfth annual conference on Computational learning theory* pp 203-208. Santa Cruz, California, USA: ACM

Bowen NJ, Walker LD, Matyunina LV, Logani S, Totten KA, Benigno BB, McDonald JF (2009) Gene expression profiling supports the hypothesis that human ovarian surface epithelia are multipotent and capable of serving as ovarian cancer initiating cells. BMC Med Genomics 2: 71

Bray F, Ren JS, Masuyer E, Ferlay J (2013) Global estimates of cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer 132: 1133-1145

Calza S, Hall P, Auer G, Bjohle J, Klaar S, Kronenwett U, Liu ET, Miller L, Ploner A, Smeds J, *et al* (2006) Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients. Breast Cancer Res 8: R34

Chen C, Grennan K, Badner J, Zhang D, Gershon E, Jin L, Liu C (2011) Removing batch effects in analysis of expression microarray data: an evaluation of six batch adjustment methods. PloS ONE 6: e17238

Cheung HW, Cowley GS, Weir BA, Boehm JS, Rusin S, Scott JA, East A, Ali LD, Lizotte PH, Wong TC, *et al* (2011) Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. Proc Natl Acad Sci USA 108: 12372-12377

Chin L, Hahn WC, Getz G, Meyerson M (2011) Making sense of cancer genomic data. Genes Dev 25: 534-555

Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd edition. Lawrence Erlbaum Associates: New Jersey, USA.

Dabney AR (2006) ClaNC: point-and-click software for classifying microarrays to nearest centroids. Bioinformatics 22: 122-123

Denkert C, Budczies J, Darb-Esfahani S, Gyorffy B, Sehouli J, Konsgen D, Zeillinger R, Weichert W, Noske A, Buckendahl AC, *et al* (2009) A prognostic gene expression index in ovarian cancer – validation across different independent data sets. J Pathol 218: 273-280

Ewens WJ, Grant GR (2001) *Statistical Methods in Bioinformatics. An Introduction*. Springer-Verlag Inc., New York

Fava F, Raynaud-Messina B, Leung-Tack J, Mazzolini L, Li M, Guillemot JC, Cachot D, Tollon Y, Ferrara P, Wright M (1999) Human 76p: a new member of the gamma-tubulin-associated protein family. J Cell Biol 147: 857-868

Fox N, Mathers N (1997) Empowering research: statistical power in general practice research. Family Pract 14: 324-329

Gatza ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, Datto MB, Kelley M, Mathey-Prevot B, Potti A, *et al* (2010) A pathway-based classification of human breast cancer. Proc Natl Acad Sci USA 107: 6994-6999

Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307-315

George SL (1996) Reducing patient eligibility criteria in cancer clinical trials. J Clin Oncol 14: 1364-1370

Gilks CB, Prat J (2009) Ovarian carcinoma pathology and genetics: recent advances. Hum Pathol 40: 1213-1223

Haibe-Kains B, Desmedt C, Loi S, Culhane AC, Bontempi G, Quackenbush J, Sotiriou C (2012) A three-gene model to robustly identify breast cancer molecular subtypes. J Natl Cancer Inst 104: 311-325

Helland A, Anglesio MS, George J, Cowin PA, Johnstone CN, House CM, Sheppard KE, Etemadmoghadam D, Melnyk N, Rustgi AK, *et al* (2011) Deregulation of MYCN, LIN28B and LET7 in a molecular subtype of aggressive high-grade serous ovarian cancers. PLoS ONE 6: e18064

Hendrix ND, Wu R, Kuick R, Schwartz DR, Fearon ER, Cho KR (2006) Fibroblast growth factor 9 has oncogenic activity and is a downstream target of Wnt signaling in ovarian endometrioid adenocarcinomas. Cancer Res 66: 1354-11362

Hogdall EV, Christensen L, Kjaer SK, Blaakaer J, Bock JE, Glud E, Norgaard-Pedersen B, Hogdall CK (2003) Distribution of HER-2 overexpression in ovarian carcinoma tissue and its prognostic value in patients with ovarian carcinoma: from the Danish MALOVA Ovarian Cancer Study. Cancer 98: 66-73

Howell SJ (2013) Advances in the treatment of luminal breast cancer. Curr Opin Obstet Gynecol 25: 49-54

Hsu DS, Balakumaran BS, Acharya CR, Vlahovic V, Walters KS, Garman K, Anders C, Riedel RF, Lancaster J, Harpole D, *et al* (2007) Pharmacogenomic strategies provide a rational approach to the treatment of cisplatin-resistant patients with advanced cancer. J Clin Oncol 25: 4350-4357

Huang RY, Wang SM, Hsieh CY, Wu JC (2008) Lysophosphatidic acid induces ovarian cancer cell dispersal by activating Fyn kinase associated with p120-catenin. Int J Cancer 123: 801-809

Hubbert C, Guardiola A, Shao R, Kawaguchi Y, Ito A, Nixon A, Yoshida M, Wang XF, Yao TP (2002) HDAC6 is a microtubule-associated deacetylase. Nature 417: 455-458

Iorio E, Ricci A, Bagnoli M, Pisanu ME, Castellano G, Di Vito M, Venturini E, Glunde K, Bhujwalla ZM, Mezzanzanica D, *et al* (2010) Activation of phosphatidylcholine cycle enzymes in human epithelial ovarian cancer cells. Cancer Res 70: 2126-2135

Jochumsen KM, Tan Q, Holund B, Kruse TA, Mogensen O (2007) Gene expression in epithelial ovarian cancer: a study of intratumor heterogeneity. Int J Gynecol Cancer 17: 979-985

Jochumsen KM, Tan Q, Hogdall EV, Hogdall C, Kjaer SK, Blaakaer J, Kruse TA, Mogensen O (2009) Gene expression profiles as prognostic markers in women with ovarian cancer. Int J Gynecol Cancer 19: 1205-1213

Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8: 118-127

Kim J-H (2009) Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. Comput Stat Data Anal 53: 3735-3745

King ER, Tung CS, Tsang YT, Zu Z, Lok GT, Deavers MT, Malpica A, Wolf JK, Lu KH, Birrer MJ, *et al* (2011) The anterior gradient homolog 3 (AGR3) gene is associated with differentiation and survival in ovarian cancer. Am J Surg Pathol 35: 904-912

Konavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In 14th International Joint Conference on Artificial intelligence pp 1137-1143

Konstantinopoulos PA, Spentzos D, Karlan BY, Taniguchi T, Fountzilas E, Francoeur N, Levine DA, Cannistra SA (2010) Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. J Clin Oncol 28: 3555-3561

Liu H, Yang R, Tinner B, Choudhry A, Schutze N, Chaqour B (2008a) Cysteine-rich protein 61 and connective tissue growth factor induce deadhesion and anoikis of retinal pericytes. Endocrinology 149: 1666-1677

Liu Y, Hayes DN, Nobel A, Marron JS (2008b) Statistical significance of clustering for high-dimension, low-sample size data. J Am Stat Assoc 103: 1281-1293

Lobert S, Vulevic B, Correia JJ (1996) Interaction of vinca alkaloids with tubulin: a comparison of vinblastine, vincristine, and vinorelbine. Biochemistry 35: 6806-6814

Luo B, Cheung HW, Subramanian A, Sharifnia T, Okamoto M, Yang X, Hinkle G, Boehm JS, Beroukhim R, Weir BA, *et al* (2008) Highly parallel identification of essential genes in cancer cells. Proc Natl Acad Sci USA 105: 20380-20385

Manfredi JJ, Horwitz SB (1984) Taxol: an antimitotic agent with a new mechanism of action. Pharmacol Therapeut 25: 83-125

Maruyama K, Ochiai A, Akimoto S, Nakamura S, Baba S, Moriya Y, Hirohashi S (2000) Cytoplasmic beta-catenin accumulation as a predictor of hematogenous metastasis in human colorectal cancer. Oncology 59: 302-309

Matsumura N, Huang Z, Mori S, Baba T, Fujii S, Konishi I, Iversen ES, Berchuck A, Murphy SK (2011) Epigenetic suppression of the TGF-beta pathway revealed by transcriptome profiling in ovarian cancer. Genome Res 21: 74-82

McGuire WP, Hoskins WJ, Brady MF, Kucera PR, Partridge EE, Look KY, Clarke-Pearson DL, Davidson M (1996) Cyclophosphamide and cisplatin compared with paclitaxel and cisplatin in patients with stage III and stage IV ovarian cancer. N Engl J Med 334: 1-6

Meyniel JP, Cottu PH, Decraene C, Stern MH, Couturier J, Lebigot I, Nicolas A, Weber N, Fourchotte V, Alran S, *et al* (2010) A genomic and transcriptomic approach for a differential diagnosis between primary and secondary ovarian carcinomas in patients with a previous history of breast cancer. BMC Cancer 10: 222

Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK, *et al* (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell 124: 1283-1298

Mok SC, Bonome T, Vathipadiekal V, Bell A, Johnson ME, Wong KK, Park DC, Hao K, Yip DK, Donninger H, *et al* (2009) A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor: microfibril-associated glycoprotein 2. Cancer Cell 16: 521-532

Monk BJ, Huang HQ, Burger RA, Mannel RS, Homesley HD, Fowler J, Greer BE, Boente M, Liang SX, Wenzel L (2012) Patient reported outcomes of a randomized, placebo-controlled trial of bevacizumab in the front-line treatment of ovarian cancer: a Gynecologic Oncology Group Study. Gynecol Oncol 128: 573-578

Monti S, Savage KJ, Kutok JL, Feuerhake F, Kurtin P, Mihm M, Wu B, Pasqualucci L, Neuberg D, Aguiar RC, *et al* (2005) Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. Blood 105: 1851-1861

Mori S, Chang JT, Andrechek ER, Matsumura N, Baba T, Yao G, Kim JW, Gatza M, Murphy S, Nevins JR (2009) Anchorage-independent cell growth signature identifies tumors with metastatic potential. Oncogene 28: 2796-2805

Moritz M, Braunfeld MB, Sedat JW, Alberts B, Agard DA (1995) Microtubule nucleation by gamma-tubulin-containing rings in the centrosome. Nature 378: 638-640

Moritz M, Zheng Y, Alberts BM, Oegema K (1998) Recruitment of the gamma-tubulin ring complex to Drosophila salt-stripped centrosome scaffolds. J Cell Biol 142: 775-786

Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, *et al* (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10: 515-527

Pejovic T, Pande NT, Mori M, Mhawech-Fauceglia P, Harrington C, Mongoue-Tchokote S, Dim D, Andrews C, Beck A, Tarumi Y, *et al* (2009) Expression

profiling of the ovarian surface kinome reveals candidate genes for early neoplastic changes. Transl Oncol 2: 341-349

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, *et al* (2000) Molecular portraits of human breast tumours. Nature 406: 747-752

Quintas-Cardama A, Kantarjian H, Cortes J (2009) Imatinib and beyond – exploring the full potential of targeted therapy for CML. Nat Rev Clin Oncol 6: 535-543

Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP (2006) GenePattern 2.0. Nat Genet 38: 500-501

Root DE, Hacohen N, Hahn WC, Lander ES, Sabatini DM (2006) Genome-scale loss-of-function screening with a lentiviral RNAi library. Nat Methods 3: 715-719

Rosell R, Viteri S, Molina MA, Benlloch S, Taron M (2010) Epidermal growth factor receptor tyrosine kinase inhibitors as first-line treatment in advanced nonsmall-cell lung cancer. Curr Opin Oncol 22: 112-120

Scholl C, Frohling S, Dunn IF, Schinzel AC, Barbie DA, Kim SY, Silver SJ, Tamayo P, Wadlow RC, Ramaswamy S, *et al* (2009) Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. Cell 137: 821-834

Shen Q, Zheng X, McNutt MA, Guang L, Sun Y, Wang J, Gong Y, Hou L, Zhang B (2009) NAT10, a nucleolar protein, localizes to the midbody and regulates cytokinesis and acetylation of microtubules. Exp Cell Res 315: 1653-1667

Sims D, Mendes-Pereira AM, Frankum J, Burgess D, Cerone MA, Lombardelli C, Mitsopoulos C, Hakas J, Murugaesu N, Isacke CM, *et al* (2011) High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. Genome Biol 12: R104

Sordella R, Bell DW, Haber DA, Settleman J (2004) Gefitinib-sensitizing EGFR mutations in lung cancer activate anti-apoptotic pathways. Science 305: 1163-1167

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, *et al* (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 98: 10869-10874

Subramanian J, Simon R (2011) An evaluation of resampling methods for assessment of survival risk prediction in high-dimensional settings. Stat Med 30: 642-653

Syrjanen K, Syrjanen S (2013) Detection of human papillomavirus in sinonasal papillomas: systematic review and meta-analysis. Laryngoscope 123: 181-192

The Cancer Genome Atlas Research Network. (2011) Integrated genomic analyses of ovarian carcinoma. Nature 474: 609-615

Therneau TM, Grambsch PM (2000) *Modeling Survival Data. Extending the Cox Model*. Springer-Verlag Inc., New York

Tone AA, Begley H, Sharma M, Murphy J, Rosen B, Brown TJ, Shaw PA (2008) Gene expression profiles of luteal phase fallopian tube epithelium from BRCA mutation carriers resemble high-grade serous carcinoma. Clin Cancer Res 14: 4067-4078

Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, Johnson DS, Trivett MK, Etemadmoghadam D, Locandro B, *et al* (2008) Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. Clin Cancer Res 14: 5198-5208

Tung CS, Mok SC, Tsang YT, Zu Z, Song H, Liu J, Deavers MT, Malpica A, Wolf JK, Lu KH, *et al* (2009) PAX2 expression in low malignant potential ovarian tumors and low-grade ovarian serous carcinomas. Mod Pathol 22: 1243-1250

Vaughan S, Coward JI, Bast RC Jr, Berchuck A, Berek JS, Brenton JD, Coukos G, Crum CC, Drapkin R, Etemadmoghadam D, *et al* (2011) Rethinking ovarian cancer: recommendations for improving outcomes. Nat Rev Cancer 11: 719-725

Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding L, Golub T, Mesirov JP, *et al* (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell 17: 98-110

Verhaak RG, Tamayo P, Yang JY, Hubbard D, Zhang H, Creighton CJ, Fereday S, Lawrence M, Carter SL, Mermel CH, *et al* (2013) Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. J Clin Invest 123: 517-525

Yaziji H, Goldstein LC, Barry TS, Werling R, Hwang H, Ellis GK, Gralow JR, Livingston RB, Gown AM (2004) HER-2 testing in breast cancer using parallel tissue-based methods. JAMA 291: 1972-1977

Yin JJ, Selander K, Chirgwin JM, Dallas M, Grubbs BG, Wieser R, Massague J, Mundy GR, Guise TA (1999) TGF-beta signaling blockade inhibits PTHrP secretion by breast cancer cells and bone metastases development. J Clin Invest 103: 197-206

Utilization of genomic signatures to identify high efficacy candidate drugs for chemo-refractory endometrial cancers

[1]Budiman Kharma, [1]Tsukasa Baba, [1]Masaki Mandai, [1]Noriomi Matsumura, [2]Susan K. Murphy, [1]Hyun Sook Kang, [1]Koji Yamanoi, [1]Junzo Hamanishi, [1]Ken Yamaguchi, [1]Yumiko Yoshioka, and [1]Ikuo Konishi

[1]Department of Gynecology and Obstetrics, Kyoto University Graduate School of Medicine, Kyoto, JAPAN

[2]Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, NC

Correspondence;

54 Shogoin Kawahara-cho, Sakyo-ku, Kyoto, Kyoto 606-8507, JAPAN

Phone; 81-75-751-3269, Fax; 81-75-761-3967, email; babatsu@kuhp.kyoto-u.ac.jp

Conflict of interest; the authors declare no conflict of interest.

Novelty and Impacts; we declare that all data are novel, developed by our own experiments and have not been published or submitted for publication. We have used a pharmacogenomics approach with drug-specific signatures as a targeted method to identify new efficacious candidate drugs through array-based analysis, and confirmed using *in vitro* and *in vivo* experiments. These findings may not only provide promising drugs for improving prognosis of chemorefractory patients, but also support the utility of pharmacogenomics approaches.

Running title; Fludarabine is a therapeutic candidate for refractory endometrial cancer

Keywords; chemoresistant, fludarabine, chemodynamics, endometrial cancer

Abbreviations;

Abstract

Objectives: Endometrial cancer, one of the most common gynecologic malignancies, is increasing in Japan, nearly doubling over the last decade. High grade disease patients are often resistant to conventional chemotherapy with platinum agents; therefore discovery of efficacious new drugs in this setting is required to benefit chemo-refractory cases. **Methods:** The 50% Growth-Inhibitory (GI50) concentration of 27 clinically relevant drugs was measured in the NCI60 panel of cell lines. Gene expression data were analyzed using Bayesian binary regression, to first generate a response signature for each drug, then to calculate individual susceptibility scores using *in vivo* endometrial cancer data (GSE2109; http://www.ncbi.nlm.nih.gov/geo) and *in vitro* data (GSE25458), as well as to identify candidate drugs for chemo-refractory cases. Using these candidates, cell proliferation, apoptosis, and caspase assays were performed *in vitro*. The tumor growth-inhibitory effect of the candidate was also assessed *in vivo* using nude mice. **Results:** Through microarray analysis, Fludarabine and Temsirolimus showed higher susceptibility scores in high grade cases compared with cisplatin, doxorubicin, and paclitaxel. Fludarabine significantly inhibited cell proliferation and increased apoptosis in the cisplatin-resistant endometrial cancer cell line, HEC1A, relative to HEC50B ($p<0.001$). Fludarabine treatment also enhanced Caspase 3/7 activity in HEC1A relative to HEC50B cells ($p<0.001$), and inhibited the growth of HEC1A xenograft tumors relative to cisplatin ($p<0.05$). **Conclusions:** These results support that identification and use of genomic signatures can lead to identification of new therapeutic candidates that may prove beneficial to chemo-resistant cases. Fludarabine may be useful in targeting high grade, chemo-refractory endometrial cancer.

Introduction

Endometrial cancer is the leading cause of gynecologic malignancy with 43,470 estimated cases diagnosed per year and 7,950 annual death in the United States, respectively consisting of 6% of new cancer cases and 3% of all cancer deaths, and disease incidence has been steadily increasing[1, 2]. The majority of endometrial cancers, more than 80%, are diagnosed at an early stage with the disease located within the uterus. When diagnosed at an early stage, primary surgery is frequently curative enough to be associated with a favorable prognosis. In contrast, extra-uterine spread of cancer cells profoundly impacts patient prognosis as previous studies revealed high hazard ratios for stage III and stage IV compared with stage I disease[2]. Clear cell and papillary serous carcinomas of the uterus are associated with aggressive behaviors, even at an early stage, with five year survival between 60-66%[3]. Besides staging and histology, several pathological factors, such as tumor grade, depth of invasion, and lymph vascular invasion are well-known to determine the prognosis of each patient with an aggressively metastatic phenotype. Recently, adjuvant chemotherapy has been introduced after primary surgery as part of the first-line management for preventing recurrence of such high-risk disease[4]. First line chemotherapy typically consists of a combination regimen followed by treatment with a single agent on disease progression. Throughout the phase II-III studies of the Gynecologic Oncology Group (GOG), platinum combined with doxorubicin and/or taxane has played an important role in the treatment of high-risk disease, but was also associated with infrequent complete response with recurrence in nearly half of these patients. Furthermore, there is no active second-line agent after failure of primary chemotherapies, as the response rate to paclitaxel was at most 25% for recurrent patients previously treated with doxorubicin and cisplatin[5]. With the objective of

improving the prognosis of those with high-risk disease, it is essential to identify candidate cytotoxic agents that are effective against patients with resistance to conventional chemotherapies (chemo-refractory tumors), or supportive agents that increase sensitivity to primary chemotherapies.

Identification of effective second-line agents for chemo-refractory cancers has been a long-sought goal, and over the past few years various new cytotoxic agents have been synthesized for the treatment of malignancies. Clinical trials for endometrial cancer have also historically been conducted using drugs identified as effective for other solid malignancies, especially ovarian cancer. However, these trial-and-error approaches to drug mining are inefficient for this heterogeneous entity "endometrial cancer", and have mostly failed. To elevate the efficacy of second-line chemotherapy, individualized therapy will be necessary based on biological features of the patient and tumor, such as molecular mechanisms and clinical phenotypes.

Recent development of genome-wide analysis with microarray has revealed chemo-refractory cancers possess characteristic gene expression profiles, so called chemo-resistant signatures[6, 7]. A computational analysis using Bayesian binary regression methods enabled to project a phenotype-signature extracted from one microarray onto another microarray to predict the phenotype probability of each sample in projected microarray. In this study, we utilized this bioinformatics approach to demarcate that fludarabine had potential efficacy in chemo-refractory endometrial cancers. Furthermore, we performed several *in vitro* and *in vivo* approaches using endometrial cancer cell lines to demonstrate that fludarabine may be a potential alternative treatment for chemo-refractory endometrial cancers.

Material and Methods

Patients.

Clinicopathological information of 262 patients treated for endometrial cancer during 2004-2011 in Kyoto University Hospital were obtained with written consent from each patient and used under protocols approved by the Kyoto University Institutional Review Board. The prognostic risk of each case was determined as low, intermediate, or high as previously described[8].

Cell lines and culture.

Human endometrial cancer cell lines, AN3CA, HEC1A, HEC1B, KLE, RL95-2, TEN (ATCC, Rockville, MD), ACC230, ACC564 (DSMZ, Brunswick, Germany), HHUA, Ishikawa, JHUEM-1, JHUEM-2, JHUEM-3, JHUEM-7, JHUEM-14, Sawano (RIKKEN BRC, Tsukuba, Japan), HEC50B, HEC108, HEC265, SNG-M (JCRB, Osaka, Japan), were maintained in RPMI1640 (Nikken, Kyoto, Japan) or DMEM/Ham's F12 (Invitrogen, Carlsbad, CA) supplemented with 10% heat-inactivated fetal bovine serum (v/v; Biowest, France) and penicillin–streptomycin (100 IU/ml penicillin, 100 μg/ml streptomycin; Nacalai Tesque, Kyoto, Japan). All cells were seeded into Cellstars® tissue culture plates (Greiner, Frickenhausen, Germany) and used for experiments after 18hr incubation.

Chemicals.

Following the manufacturer's instructions, a 5mM stock solution of fludarabine (Alexis Biochemicals; San Diego, CA) was prepared in cold sterile water and stored at -20° C. Fresh

thawed dilutions were used for each experiment. A 3.33mM stock solution of cisplatin (Sigma Aldrich; St Louis, MO) was also prepared in sterile water following the manufacturer's instructions.

Bioinformatics Analyses.

Total RNA was extracted from cell lines using the RNeasy® Mini Kit (QIAGEN, Valencia, CA), and gene expression microarray data (Affymetrix U133 Plus 2.0) were generated in triplicate and RMA-normalized as described previously[9]. Expression microarray data of endometrial cancers (GSE2109) was also obtained from the Gene Expression Omnibus web site (http://www.ncbi.nlm.nih.gov/geo). Gene expression data for NCI60 cell lines was obtained from the National Cancer Institute along with 50% Growth Inhibitory (GI50) values for 27 commercially available chemotherapeutic agents (http://www.dtp.nci.nih.gov/webdata.html). GI50 value data for 27 drugs from the NCI60 cell line data were normalized with median centering using Cluster 3.0 (http://rana.lbl.gov/EisenSoftware.htm), and converted into a visual representation using JavaTreeView (http://jtreeview.sourceforge.net/). Heat maps were generated using Matlab (Mathworks, Natick, MA) or R with Bioconductor (http://www.r-project.org/) as described previously[9, 10, 11]. Genomic signatures of drug susceptibility were generated using Bayesian binary regression[10] from gene expression data of 10 sensitive and 10 resistant cell lines from the NCI60 drugs screening panel. Probabilities of chemosensitivity in response to NCI60 drugs were scored for each sample in dataset GSE 25458 as previously described[9].

Cell Proliferation Assay.

AN3CA, HEC1A, and HEC50B cells were seeded into 96-well tissue culture plates at 2×10^3 cells per well. The cell culture medium was replaced with fresh medium containing 0, 10, 25, 50,

100, 250, or 500μM fludarabine and incubated for 24hr. The number of viable cells in each well was examined using the WST-1 assay kit (Premix WST-1®, Takara, Otsu, Japan) following the manufacturer's instructions.

Apoptosis Detection.

AN3CA, HEC1A, and HEC50B cells were treated with medium containing 0, 25, 50, or 100μM fludarabine for 24hr. Cells were harvested by trypsinization for apoptosis detection by flow cytometry. Following washing with phosphate-buffered saline, cells were resuspended in 200μL of 1× annexin-V binding buffer (BD Pharmingen, Inc.). Next, 5μL of 7-AAD (BD Pharmingen, Inc.) and 5μL of annexin-V were added to the tubes and incubated for 10 minutes at 4°C in the dark. Cells were washed with 500μL of 1× annexin-V binding buffer and then resuspended in 200μL of 1× annexin-V binding buffer followed by filtration through a 70-μm nylon filter (Becton Dickinson Labware, Franklin Lakes, NJ). The samples were placed on ice and analyzed by FACSCalibur (Becton Dickinson). The data were analyzed using FlowJo v.7.6.3 (Tree Star, Ashland, OR).

Caspase 3/7 activity was measured using the Caspase-Glo® 3/7 Assay System (Promega, Madison, WI) following the manufacturer's instructions. After 2hr incubation with the Caspase-Glo® 3/7 Reagent, luminescence was measured using a Glo-Max®Luminometer (Promega) as directed by the manufacturer.

Western Blot.

AN3CA, HEC1A, and HEC50B cells were treated with medium containing 0, 25, 50, or 100μM fludarabine for 24hr. Cells were lysed in radio-immunoprecipitation assay (RIPA) buffer

(Thermo Fisher Scientific, Waltham, MA) with a protease inhibitor cocktail (EMD, Madison, WI) and a phosphatase inhibitor cocktail (Nacalai Tesque). Protein was quantified using the DC Protein Assay Kit (Bio-Rad, Hercules, CA). Twenty micrograms of sodium dodecyl sulfate (SDS)-treated protein was loaded onto a 10-20% Tris-Tricine Mini Protean® gel (Bio-Rad). Gels were electroblotted onto polyvinylidene fluoride (PVDF) membranes (Bio-Rad). Nonspecific binding of the antibody was blocked by 1hr incubation at room temperature in Blocking One-P (Nacalai Tesque). The membranes were incubated overnight at 4°C with Caspase-3 antibody (1:1,000, Caspase-3 Rabbit mAb, Cell Signaling Technology, Danvers, MA) or Cleaved Caspase-3 (Asp175) antibody (1:1,000, Cleaved Caspase-3 Rabbit polyclonal Ab, Cell Signaling Technology). After washing in tris-buffered saline (TBS)-T, the blots were incubated with the appropriate peroxidase-coupled secondary antibody (1:6,000; Anti-rabbit HRP, GE Healthcare Life Sciences, Uppsala, Sweden). β-actin was used as an endogenous loading control, detected using an anti-human β-actin antibody (1:8,000; Rabbit mAb, Abcam, Cambidge, MA). Specific proteins were detected using ECL Plus Western Blotting Reagent (GE Healthcare Life Sciences). The bands were visualized using Molecular Imager® Gel Doc™ XR+ and ChemiDoc™ XRS+ Systems with Image Lab 2.0 software (Bio-Rad).

Real Time qPCR.

Total RNA was extracted from cell lines using the RNeasy® Mini Kit (QIAGEN). To monitor gene expression, quantitative reverse transcriptase (RT)-PCR amplification of human Caspase 3α, Caspase 3β, and *GAPDH* mRNAs was done by Light Cycler 480-II (Roche, Basel, Switzerland) using a Dual Color Hydrolysis Universal Probe System (Roche). The following primers which were obtained from the Universal Probe Library Assay Design Center (Roche)

were used for analysis: Caspase 3α, 5'-CTG GTT TTC GGT GGG TGT-3' (forward), 5'-CCA CTG AGT TTT CAG TGT TCT CC-3' (reverse); Caspase 3β, 5'-TGG AAT TGA TGC GTG ATG TT-3' (forward), 5'-TGG CTC AGA AGC ACA CAA AC-3' (reverse); GAPDH, 5'-AGC CAC ATC GCT CAG ACA C-3' (forward); and 5'-GCC CAA TAC GAC CAA ATC C-3' (reverse). Cycling parameters were 95°C for 10 seconds followed by 40 cycles of 95°C for 5 seconds and 60°C for 30 seconds, followed by a dissociation cycle of 95°C for 15 seconds, 60°C for 20 seconds, and 95°C for 15 seconds. The expression of human Caspase 3α and Caspase 3β mRNAs was estimated by dividing the Caspase 3α and Caspase 3β threshold cycle (CT) values by the *GAPDH* CT values.

In Vivo Experiment.

Female CD-1 Foxn/Nu mice were purchased from Oriental Bioservice, Inc. (Kyoto, Japan). Animal care and experimental procedures under specific pathogen-free conditions were performed in accordance with the guidelines of the Institute of Laboratory Animals Graduate School of Medicine, Kyoto University. Subcutaneous xenografts were established in the flanks by inoculating 5×10^6 cells of the AN3CA, HEC1A, or HEC50B cell lines. Three days after inoculation, each mouse was treated twice a day for 5 days with intraperitoneal administration of 125mg/kg fludarabine, 1mg/kg cisplatin, or 50μL sterile distilled water as a control treatment (n=4 in each arm). Therapeutic effects were monitored by measuring tumor growth for 30 days after inoculation. The differences in tumor growth were analyzed statistically.

Statistical Analysis.

Group comparisons were done using Mann–Whitney U tests. Prognostic analysis was done using Log-rank test and Fisher's exact test. Statistical analyses were done using GraphPad Prism 5.5 software. Probability values below 0.05 were considered significant.

Results

Clinical significance of current chemotherapy in the treatment of endometrial cancer

Clinical features of the 262 patients studied here are listed in Table 1. Forty-three patients exhibited progression or recurrence of the disease (PD). Clinicopathological analysis revealed that each known prognostic factor including advanced stage, outer-half myometrial invasion, lymphovascular space invasion, and histological grade showed significant differences both in terms of the rate of PD and in the overall survival of patients (OS). Based on the prognostic risk classification, eighty-six patients were categorized as "low-risk" without any PD. Three out of eighty-two "intermediate-risk" patients recurred, but there was no significant difference in OS compared with "low-risk" patients, while "high-risk" patients exhibited higher PD and poor OS ($p<0.0001$, Suppl. Fig.1). PD was significantly lower among patients who received chemotherapies with cisplatin+doxorubicin and paclitaxel+carboplatin in the "intermediate-risk" group ($p<0.05$, Table 2), while not in the "high-risk" group ($p=0.72$).

Drug Sensitivity Prediction from the NCI60 Data.

As a first step toward identifying compounds that exhibit efficacy toward chemoresistant endometrial cancer cells, we queried the NCI-60 database for the 50% Growth Inhibitory doses (GI50) of conventional chemotherapeutic agents including cisplatin, paclitaxel, and doxorubicin.

GI50 values were normalized (see Methods) and visualized using a heatmap. Based on this representation, normalized GI50 values > 0.04 were colored in red or orange and 0.04 was thus determined as a threshold for sensitivity (Fig. 1A). By this criterion, 24 of 62 cell lines (39%) were resistant to cisplatin, doxorubicin, and paclitaxel, and this rate, 39%, was very close to the rate of progressive disease (PD, 44%) among the "high-risk" patients in our clinical data.

Secondly, the GI50 values were obtained from the NCI-60 database for another 24 commonly used chemotherapeutic agents for the 24 chemoresistant cell lines. An unsupervised hierarchical clustering analysis using the normalized GI50 values for the chemoresistant cells revealed that these 24 drugs were divided into two clusters: Cluster 1 contains the anti-cancer drugs while cluster 2 contains molecular targeting drugs (Fig.1B). Some of the drugs exhibited efficacy against several cell lines, but fludarabine appeared to effectively target ten out of these 24 chemoresistant cells.

Predicted susceptibility of chemotherapeutic agents in endometrial cancers.

The NCI60 drug screening panel also contains genetic information. By selecting 10 sensitive and 10 resistant cell lines, a genomic signature of drug susceptibility for each chemotherapeutic agent was derived using Bayesian binary regression[10] from the gene expression data. Such genomic signatures can be applied to an independent gene expression dataset to predict drug susceptibility for each sample in the dataset as previously described[12]. Drug-susceptibility signatures developed from the NCI60 dataset were applied to the microarray data of endometrial cancers in GSE2109 for predicting the probabilities of sensitivity to cisplatin, doxorubicin, and paclitaxel. Meanwhile, drug-susceptibility signatures were also developed for fludarabine, Ara-C, irinotecan, and 5-FU, as representative cluster 1 anti-cancer drugs, and for imatinib, gefitinib,

sunitinib, and temsirolimus, as representative molecular targeting drugs from cluster 2; these drugs were chosen as representative drugs for each sub-cluster based on chemo-signatures for the NCI60 cell lines. The sensitivity probabilities of the three conventional chemotherapeutic agents in grade 3 endometrioid adenocarcinoma and serous papillary adenocarcinoma were not superior to those in low grade endometrioid adenocarcinoma (Fig.2A). On the other hand, the probability of sensitivity to fludarabine was significantly higher in grade 3 and serous, and the probability of sensitivity to temsirolimus was significantly higher in serous (Fig.2B; $p<0.001$ and $p<0.05$, respectively).

Gene expression microarray analysis was performed in 20 endometrial cancer cell lines, and the drug-susceptibility signatures of conventional chemo-agents were applied to predict the sensitivity of each cell line. There was a statistically significant correlation between cisplatin GI50 values of 37 ovarian cancer cell lines[10] and cisplatin sensitivity probability scores derived from the gene expression microarray data (GSE25428, r=0.3776, $p=0.02$, data not shown).

Seven cell lines exhibited low probabilities of sensitivity to cisplatin, doxorubicin, and paclitaxel, while another 13 cell lines exhibited higher probabilities of sensitivity. Intriguingly, five out of these seven cells showed relatively high probabilities of sensitivity to fludarabine (Fig.2C), and three had favorable probability scores to temsirolimus (Suppl. Table 2). We further investigated HEC1A as a representative chemo-refractory cell line with AN3CA (sensitive to all conventional chemo-agents) and HEC50B (partly sensitive to conventional chemo-agents) as counterpart controls for further experiments. Fludarabine was selected for further experiments as the most potent alternative agent for chemo-refractory cases.

Cytotoxic activities of fludarabine in vitro.

The cytotoxic effect of fludarabine was examined using proliferation assays. AN3CA, HEC1A, and HEC50B were exposed to fludarabine (from 0 to 500μM) for 24hr. There were dose-dependent growth-inhibitory responses in the AN3CA and HEC1A cells ($p<0.01$), and more than 90% growth abrogation was observed in HEC1A cells at 500μM ($p<0.001$). There was no significant difference in proliferation following fludarabine treatment for the HEC50B cells at any dose (Fig.3A).

To investigate the growth-inhibitory mechanism of fludarabine in endometrial cancer cells, Annexin-V/7-AAD apoptosis assays were performed using flow cytometry. After 24hr exposure to fludarabine, apoptosis in AN3CA and HEC1A cells was increased in a dose-dependent manner, with 20% and 25% of the cells undergoing apoptosis following treatment with 100μM fudarabine (Fig.3B). Conversely, no induction of apoptosis was observed in the HEC50B cells.

Fludarabine-induced caspase-3 activity.

Caspase-3 activation is involved in cell death signaling and induces chromatin condensation and DNA fragmentation, resulting in apoptosis. Caspase-3α and Caspase-3β mRNA expression in HEC1A and AN3CA cells were up-regulated by fludarabine in a dose-dependent manner ($p<0.05$, Suppl. Fig.2A). HEC1A cells were still responsive to fludarabine at the lowest dose, 25μM, while a similar low dose response was not evident in AN3CA cells. There was no change in expression levels of these genes in the HEC50B cells, even at 100μM. Similarly, Caspase-3 protein expression increased in HEC1A cells following fludarabine treatment in a dose-dependent manner accompanied by a concomitant increase of cleaved-

Caspase-3 (Suppl. Fig.2B). Next, Caspase 3/7 activity was assessed using a luminometer to investigate whether or not the up-regulation of Caspase-3 expression in fludarabine-treated cells was related to the activity of Caspase-3 in facilitating and inducing apoptosis. After 24hr exposure to fludarabine, the luminescence was gradually elevated in HEC1A cells in a dose-dependent manner to reflect the activity of Caspase 3/7, and this induction was almost 6-fold higher than that observed in HEC50B cells ($p<0.001$, Fig.3C). The AN3CA cells also showed increasing luminescence but the slope of the increase was lower than that in HEC1A cells.

Therapeutic effects of fludarabine in a mouse xenograft model of endometrial cancer.

We investigated the *in vivo* therapeutic effects of fludarabine, cisplatin, and sterile water on subcutaneously-inoculated HEC1A, AN3CA, and HEC50B xenografts in CD-1 Foxn/Nu mice. Tumor growth was completely inhibited by cisplatin in mice inoculated with AN3CA cells, consistent with the high probability of cisplatin-sensitivity predicted from the microarray analysis ($p<0.05$, Fig.3D). Growth inhibitory effects of fludarabine were not evident in AN3CA-inoculated mice, and there was no therapeutic effect of cisplatin or fludarabine on tumor growth of HEC50B-inoculated mice. On the other hand and also consistent with the predicted sensitivity, tumor growth in HEC1A-inoculated mice was indeed significantly inhibited by fludarabine ($p<0.05$, Fig.3D).

Discussion

Endometrial carcinoma is frequently diagnosed at an early stage, at which point it is usually surgically curable. Surgical treatment includes hysterectomy, bilateral salpingo-oophorectomy, and staging lymphadenectomy which is a controversial but still common procedure in treating endometrial cancers. Adjuvant therapies to prevent relapse are mainly