厚生労働科学研究費補助金 （難病・がん等の疾患分野の医療の実用化研究事業）
分担研究報告書

がん免疫逃避機構を標的にした次世代型免疫治療の
臨床応用と新規バイオマーカーの探索に関する研究

研究分担者 岡崎 拓
徳島大学 免疫・分子生物学 教授

研究要旨
がんの免疫療法は難治性がんに対する新規治療法として注目されてきたが、現在まで期待されたほどの効果は得られてない。我々は、免疫抑制受容体 PD-1（Programmed cell death-1）が、がんによる免疫逃避機構に大きく関与していることを見出したことから、PD-1 阻害によるがん免疫の再活性化を目指した新規治療法の開発を行ってきた。これまでの我々の基礎検討において、抗 PD-1 抗体による治療効果が、使用するがん細胞株の種類およびマウスの系統によって大きく異なったことから、本研究では、抗 PD-1 抗体による治療効果を予測し得るバイオマーカーを探索するとともに、抗 PD-1 阻害抗体と組み合わせることにより相乗効果を示す治療法の開発を目的とする。これにより、抗 PD-1 阻害抗体の効果が期待される患者を選別するとともに、抗 PD-1 阻害抗体単独では効果が期待されない患者にはより効果的な治療法を提供することが可能になると期待されるため、その臨床的および学術的意義は大きいと考えられる。

A．研究目的

がんによる死亡は年々増加しており、難治性がんに対する早急な対応が求められている。免疫療法は有望な新規治療法と考えられているが、現在まで期待通りの臨床効果は得られていない。その最大の理由として、癌細胞がみずから腫瘍免疫応答を抑制し、宿主免疫から逃れる『免疫逃避機構』の存在があると考えられる。マウスを用いた実験において、免疫抑制受容体 PD-1（Programmed cell death-1）の機能を阻害することによりがん免疫が増強され、効率的にがんが拒絶されたことから、がんによる免疫逃避機構において PD-1 が中心的な役割を果たしていると考えられる。そこで、PD-1 阻害によるがん免疫の再活性化を目指した新規治療法の開発を行ってきた。

これまでの我々の基礎検討において、抗 PD-1 抗体による治療効果が、使用するがん細胞株の種類およびマウスの系統によって大きく異なったことから、本研究では、抗 PD-1 抗体による治療効果を予測し得るバイオマーカーを探索するとともに、抗 PD-1 阻害抗体と組み合わせることにより相乗効果を示す治療法の開発を目的とする。

B．研究方法

1）治療効果を予測するバイオマーカーの

探索

これまでに、がん細胞の表面上に PD-1 リガンドが発現している場合には、抗PD-1 阻害抗体が治療効果を示す可能性が高いことを示唆する結果が得られているが、リンパ球側のバイオマーカーについては、ほとんど解析されていない。そこで、リンパ球側のバイオマーカーを探索する目的で、昨年度よりT細胞の活性化状態を可視化するシステムの構築を試みている。プロモーター配列、使用する細胞株、刺激条件等について、様々な組み合わせで比較検討を行った。

また、DO11.10 細胞株を用いて、PD-1 による抗原受容体刺激に対する抑制活性を観察できる実験系を、これまでに作製している。そこで、抗原刺激により発現が誘導される遺伝子のうち、PD-1 シグナルにより発現誘導が抑制される遺伝子を、マイクロアレイ解析により探索した。

2）抗 PD-1 阻害抗体と組み合わせることにより相乗効果を示す治療法の開発

PD-1阻害は、がん免疫を増強する反面、自己免疫を惹起する可能性がある。PD-1 欠損マウスが、マウスの系統により異なる種類の自己免疫疾患を自然発症することから、PD-1 欠損による免疫応答の活性化には、PD-1 欠損以外の遺伝要因や環境要因が大きく影響を与えると考えられる。最近我々は、LAG-3（Lymphocyte activation gene-3）という別の免疫抑制受容体が PD-1 と協調的に自己免疫を制御していることを見出した。他のグループによって、がん免疫応答の制御においても PD-1 と LAG-3 が協調的に働くと報告されているが、そのメカニズムについてはほとんど分

かっていない。DO11.10 細胞株を用いて、LAG-3 による抗原受容体刺激に対する抑制活性を観察できる実験系をこれまでに作製するとともに、PD-1 と LAG-3 を同時に介在させることにより、相加相乗的な抑制効果を観察できることを見出している。そこで、抗原刺激により発現が誘導される遺伝子のうち、PD-1 シグナル特異的あるいは LAG-3 シグナル特異的に、発現誘導が抑制される遺伝子を、マイクロアレイ解析により探索した。

（倫理面への配慮）

動物実験にあたっては「動物の愛護及び管理に関する法律」を遵守するとともに、徳島大学動物実験委員会において許可を得た後、徳島大学動物実験管理規則に則って行った。

遺伝子改変動物の使用等、組換え遺伝子実験にあたってはカルタヘナ議定書、遺伝子組換え生物等の使用等の規制による生物の多様性の確保に関する法律を遵守するとともに、徳島大学遺伝子組換え実験安全管理委員会において許可を得た後、徳島大学遺伝子組換え実験安全管理規則に則って行った。

C． 研究結果

1）治療効果を予測するバイオマーカーの探索

プロモーター配列の変更およびリクローニング等により、抗原刺激時の標識率をほぼ 100%に向上させることに成功した。しかし、PD-1 や LAG-3 を介在させた際にもレポーター（EGFP）の発現が中等度に認められたため、抑制を受けた細胞を鋭敏に判別するには、さらなる改良が必要であ

17

ると思われた。

　DO11.10 細胞株において、抗原刺激により発現量が2倍以上上昇した遺伝子の数は2,358個であった。そのうちの1,465遺伝子は、PD-1により50%以上、発現誘導が抑制されていた。一方、406遺伝子は、PD-1による抑制効果が10%以下であった。例えば、IL-2遺伝子は抗原刺激により発現量が15.3倍に増加していたが、PD-1介在時には3.2倍の増加にとどまり、約85%の発現誘導抑制が確認された。また、PD-1介在時に発現量が4倍以上亢進する遺伝子が613遺伝子、認められた。

2）抗 PD-1 阻害抗体と組み合わせることにより相乗効果を示す治療法の開発

　上述の通り、DO11.10 細胞株において、抗原刺激により発現量が2倍以上上昇した遺伝子の数は2,358個であったが、そのうちの1,431遺伝子が、LAG-3により50%以上、発現誘導が抑制されていた。一方、428遺伝子は、LAG-3による抑制効果が10%以下であった。例えば、IL-2遺伝子は抗原刺激により発現量が15.3倍に増加していたが、LAG-3介在時には2.0倍の増加にとどまり、約93%の発現誘導抑制が確認された。また、LAG-3介在時に発現量がむしろ亢進する遺伝子が605遺伝子、認められた。

　PD-1とLAG-3両分子の協調作用については、1,666遺伝子が50%以上の発現誘導抑制を受けていた。一方、275遺伝子は、抑制効果が10%以下であった。また、両分子介在時に発現量がむしろ亢進する遺伝子が483遺伝子、認められた。

D．考察

　活性化 T 細胞を可視化するレポーターコン

ストラクトの作製については、標識率の向上に成功したが、PD-1存在下でもレポーター遺伝子の弱い発現が確認され、PD-1による抑制を受けた細胞の判別が困難であったため、さらなる改良が必要であると思われる。

　抗原受容体刺激によって多くの遺伝子が発現誘導を受けるが、その全てがPD-1あるいはLAG-3によって抑制される訳ではないことが明らかとなった。また、一部の遺伝子については、発現がむしろ誘導されていた。PD-1による発現誘導抑制効果が特に強い遺伝子、ほとんど抑制されない遺伝子、むしろ発現が誘導される遺伝子等に着目して解析することにより、PD-1阻害抗体によるT細胞の活性化をより鋭敏かつ正確にモニターできる可能性がある。

E．結論

　抗 PD-1 阻害抗体は、がんの治療に対して革新的な変化を与えると期待される。しかし、抗 PD-1 抗体による治療効果は、使用するがん細胞株の種類およびマウスの系統によって大きく異なることから、抗 PD-1 阻害抗体による T 細胞の活性化レベルを鋭敏かつ正確にモニターし、治療効果を予測する方法を開発することが重要である。

F．健康危険情報

　なし

G．研究発表

1．論文発表

Satoru Iwamoto, Masahiro Kido, Nobuhiro Aoki, Hisayo Nishiura, Ryutaro Maruoka, Aki Ikeda, Taku Okazaki, Tsutomu Chiba, Norihi-ko Watanabe, "TNF-α is essential in the induction of fatal autoimmune hepatitis in mice through upregulation of hepatic CCL20 ex-

pression", Clin Immunol, vol. 146, No. 1, pp.15-25, 2013

2. 学会発表

岡崎　拓、*aida* マウスを用いた自己免疫疾患発症制御機構の解析、第7回自己免疫疾患研究会、東京、2012 年 7 月 7 日

Taku Okazaki, Immuno-inhibitory receptors in the regulation of autoimmunity, 2013 SKKU International Symposium on Molecular Medicine, Suwon, Korea, February 28th, 2013

岡崎一美、杉浦大祐、高橋涼香、梶原武雄、岡崎　拓、PD-1 欠損マウスを用いた自己免疫疾患のゲノム解析、第11回四国免疫フォーラム、高知県南国市、2012 年 6 月 9 日
岡崎　拓、免疫抑制受容体による免疫応答の制御、第4回　東京編・徳島大学研究者との集い、東京、2012 年 7 月 27 日
Il-mi Okazaki and Taku Okazaki, Genetic re-constitution of autoimmunity in mice, The Second Immunology Symposium at the University of Tokushima, Tokushima, January 25th, 2013
Taku Okazaki, Immuno-inhibitory receptors in the regulation of autoimmunity, The Second Immunology Symposium at the University of Tokushima, Tokushima, January 25th, 2013
Il-mi Okazaki, Daisuke Sugiura, Suzuka Takahashi, Takeo Kajihara, Taku Okazaki, Identification of new therapeutic targets by genetic dissection and reconstitution of auto-immune diseases in mice, JST-CREST International Symposium, Tokyo, February 12-13th, 2013

H.　知的財産権の出願・登録状況（予定を含む）
　1.　特許取得　　　なし
　2.　実用新案登録　なし
　3.　その他　なし

19

厚生労働科学研究費補助金 （難病・がん等の疾患分野の医療の実用化研究事業）
分担研究報告書

PD-1 欠損マウスを用いた、バイオマーカー探索の基礎的検討

研究分担者　　竹馬　俊介　　　京都大学　免疫ゲノム医学　助教
研究協力者　　RUI YUXIANG　　　京都大学　博士課程大学院生

研究要旨
　　がんにおける、PD-1 阻害療法の副作用として起こり得る、アレルギーや自己免疫様症状
の、環境危険因子や早期検出を可能にするバイオマーカーを検討するため、PD-1 欠損マ
ウスに既知の自己抗原を投与し、同時投与でT細胞反応に影響するような因子を検討して
いる。その結果、結核死菌を投与した PD-1 欠損マウス由来の脾臓細胞は、炎症性サイトカ
インである IL-6 を大量に産生し、これが T 細胞に作用して炎症性 T 細胞への分化を起こし
やすくなる事がわかった。結核の不顕性感染や、それに類する微生物学的因子が疑われ
るがん患者においては、副作用としての自己免疫症状が起こり易い可能性があり、このよう
な患者への PD-1 阻害薬投与は、より慎重に行う必要があること、自己免疫の兆候をいち早
く察知するマーカーとして、免疫細胞から産生される IL-6 の定量が有用である可能性が示
唆された。

A．研究目的
　ヒト型 PD-1 抗体療法は、がんに対する有効
な免疫増強治療法として国内外で精力的に
開発が進められている。PD-1 阻害による免疫
増強の副作用として、一部の患者には、アレ
ルギーや自己免疫様症状が現れることが予
測され、投与初期に自己免疫疾患の兆候をと
らえるバイオマーカーを得ることが出来れば
有用であると考えられる。また、自己免疫疾患
の危険因子として、患者の先天性因子だけで
なく、基礎疾患や不顕性感染といった環境因
子が考えられる。当研究では、不顕性感染の
代表である、結核菌の菌体抗原が危険性因
子となる可能性を、PD-1 欠損マウスを用いて
検討した。

B．研究方法
　PD-1 欠損マウスに、結核死菌単独、あるい
は結核死菌と、自己抗原（マウス脳ミエリンタ
ンパク由来）ペプチドを投与し、自己免疫様
疾患の発症を 30 日間観察した。また、結核死
菌のみを投与したマウスより、脾細胞を調整し、
これを抗原提示細胞として用いて、野生型マ
ウスの T 細胞を刺激し、結核菌を経験した細
胞の、T細胞分化に与える影響を検討した。こ
の混合培養系における培養上清を採集し、
BD 社が提供する Cytometric Beads Array
(CBA)アッセイを用いて、抗原特異的なサイト
カイン産生を測定した。

（倫理面への配慮）
本研究は、京都大学医学研究科の定める動
物実験実施要綱、および京都大学動物実験

指針を遵守して行った。また、PD-1 欠損マウスは、組み換え DNA 技術を用いて作出されており、事前に、京都大学組み換え DNA 実験委員会の承認を得たうえで行った。

C. 研究結果

昨年度までに、PD-1 欠損マウスに脳由来ペプチドを免疫し、結核死菌（250 □g）を同時投与すると、野生型コントロールマウスに比べ、はるかに強い自己免疫性脳脊髄炎が惹起されること、この症状は、炎症性サイトカインである IL-17 を産生する T 細胞（Th-17）分化の亢進と相関することを見出している。今回はあらたに、野生型コントロールマウスではまったく自己免疫症状が見られない、低用量（50 □g）の結核死菌でも PD-1 欠損マウスでは自己免疫性脳脊髄炎を起こすことを見出した。このことから PD-1 欠損下では、結核菌に対してまず強い反応が起こり、これが脳特異的な炎症性 Th-17 の分化に関与することが考えられた。この点を明らかにするため、結核死菌単独で免疫したマウスより 8 日後に脾臓細胞を分離し、これを抗原提示細胞として、抗原特異的ナイーブT細胞を分化させる実験を行った。結果、結核死菌を経験した PD-1 欠損マウスの抗原提示細胞、特に、CD11bを表現するマクロファージは、野生型に比べ、有意に多くの Th-17 を分化させる事が明らかとなった。試験管内で、PD-1 欠損マウス由来のマクロファージを結核死菌により刺激すると、野生型マクロファージに比べ、高濃度の IL-6 を放出することがわかった。

D. 考察

結核菌体は、抗原提示細胞を刺激して IL-6 産生をうながし、これがナイーブ T 細胞を Th-17 に分化誘導することが知られている。

PD-1 欠損下では、マクロファージから通常より多くの IL-6 産生が起こり、自己反応性 T 細胞に作用して Th-17 細胞を分化させる可能性が示唆された。IL-6 は、PD-1 欠損マウスと野生型マウスで結核菌刺激後初期に差が見られた唯一のサイトカインであり、PD-1 抗体治療開始時の血清 IL-6 を測定すれば、免疫増強療法の副作用としての自己免疫疾患発症を予測できる可能性がある。

E. 結論

結核菌など、抗原提示細胞からの IL-6 を誘導する活性をもつ病原体、およびカビなどの環境因子は、潜在的に自己反応性 Th-17 の分化も起こす可能性を持っている。ここに PD-1 阻害を行うと、自己反応性 Th-17 の活性化を誘導し、思わぬ自己免疫病を発症する可能性がある。よって、患者における、不顕性感染などの微生物学的因子は、PD-1 抗体療法の際に慎重にモニタリングされる必要がある。

F. 健康危険情報
なし

G. 研究発表
1. 論文発表
Chikuma S et al. *Nat. Immunol.*
13：596-603, 2012.

竹馬　俊介　医学のあゆみ　in press

Rui YX, Honjo T and Chikuma S. submited

2. 学会発表
竹馬　俊介「抗 PD-1 抗体によるがん治療法」日台癌のトランスレーショナル研究シンポジウム．神戸　2012 年 11 月 21 日．

竹馬　俊介「核内因子 TRIM28 による自己炎

症性 T 細胞の制御」　北海道大学獣医学部
学術交流資金群講演会「自己免疫疾患研究
の最先端」札幌　2012 年 11 月 1 日.

H．知的財産権の出願・登録状況（予定を含む）
　1．特許取得　　なし
　2．実用新案登録　なし
　3．その他　なし

Ⅲ．研究成果の刊行に関する一覧表

研究成果の刊行に関する一覧表

書籍;なし 雑誌:

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Tan TZ, Miow QH, Huang RY, Wong MK, Ye J, Lau JA, Wu MC, Bin Abdul Hadi LH, Soong R, Choolani M, Davidson B, Nesland JM, Wang LZ, Matsumura N, Mandai M, Konishi I, Goh BC, Chang JT, Thiery JP, Mori S.	Functional genomics identifies five distinct molecular subtypes with clinical relevance and pathways for growth control in epithelial ovarian cancer	EMBO Mol Med.			2013
Kharma B, Baba T, Mandai M, Matsumura N, Murphy SK, Kang HS, Yamanoi K, Hamanishi J, Yamaguchi K, Yoshioka Y, Konishi I.	Utilization of genomic signatures to identify high efficacy candidate drugs for chemo-refractory endometrial cancers.	Int J Cancer.			2013
Abiko K, Mandai M, Hamanishi J, Yoshioka Y, Matsumura N, Baba T, Yamaguchi K, Murakami R, Yamamoto A, Kharma B, Kosaka K, Konishi I.	PD-L1 on tumor cells is induced in ascites and promotes peritoneal dissemination of ovarian cancer through CTL dysfunction.	Clin Cancer Res.	19(6)	1363-1374	2013
Huang RY, Chen GB, Matsumura N, Lai HC, Mori S, Li J, Wong MK, Konishi I, Thiery JP, Goh L.	Histotype-specific copy-number alterations in ovarian cancer.	BMC Med Genomics.			2012
Matsumoto K, Katsumata N, Saito I, Shibata T, Konishi I, Fukuda H, Kamura T.	Phase II study of oral etoposide and intravenous irinotecan for patients with platinum-resistant and taxane-pretreated ovarian cancer: Japan Clinical Oncology Group Study 0503.	JPN J Clin Oncol	42(3)	222-225	2012
Horiuchi A, Hayashi T, Kikuchi N, Hayashi A, Fuseya C, Shiozawa T, Konishi I.	Hypoxia upregulates ovarian cancer invasiveness via the binding of HIF-1α to a hypoxia-induced, methylation-free hypoxia response element of S100A4 gene.	Int J Cancer.	131(8)	1755-67	2012
Yoshioka Y, Ono M, Osaki M, Konishi I, Sakaguchi S.	Differential effects of inhibition of bone morphogenic protein (BMP) signalling on T-cell activation and differentiation.	Eur J Immunol	42(3)	749-59	2012

Ⅳ. 研究成果の刊行物・別刷

Functional genomics identifies five distinct molecular subtypes with clinical relevance and pathways for growth control in epithelial ovarian cancer

Tuan Zea Tan[1†], Qing Hao Miow[1,2†], Ruby Yun-Ju Huang[1,3†], Meng Kang Wong[1], Jieru Ye[1], Jieying Amelia Lau[1], Meng Chu Wu[1], Luqman Hakim Bin Abdul Hadi[1], Richie Soong[1], Mahesh Choolani[3], Ben Davidson[4,5], Jahn M. Nesland[4,5], Ling-Zhi Wang[1,6], Noriomi Matsumura[7], Masaki Mandai[7], Ikuo Konishi[7], Boon-Cher Goh[1,6,8], Jeffrey T. Chang[9], Jean Paul Thiery[1,10,11*], Seiichi Mori[1,11,12**,‡]

Epithelial ovarian cancer (EOC) is hallmarked by a high degree of heterogeneity. To address this heterogeneity, a classification scheme was developed based on gene expression patterns of 1538 tumours. Five, biologically distinct subgroups — Epi-A, Epi-B, Mes, Stem-A and Stem-B — exhibited significantly distinct clinicopathological characteristics, deregulated pathways and patient prognoses, and were validated using independent datasets. To identify subtype-specific molecular targets, ovarian cancer cell lines representing these molecular subtypes were screened against a genome-wide shRNA library. Focusing on the poor-prognosis Stem-A subtype, we found that two genes involved in tubulin processing, *TUBGCP4* and *NAT10*, were essential for cell growth, an observation supported by a pathway analysis that also predicted involvement of microtubule-related processes. Furthermore, we observed that Stem-A cell lines were indeed more sensitive to inhibitors of tubulin polymerization, vincristine and vinorelbine, than the other subtypes. This subtyping offers new insights into the development of novel diagnostic and personalized treatment for EOC patients.

(1) Cancer Science Institute of Singapore, National University of Singapore, Singapore
(2) NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore
(3) Department of Obstetrics and Gynecology, National University Health System, Singapore
(4) Division of Pathology, Norwegian Radium Hospital Oslo University Hospital, Oslo, Norway
(5) Faculty of Medicine, University of Oslo, Institute of Clinical Medicine, Oslo, Norway
(6) Department of Pharmacology, National University of Singapore, Singapore
(7) Department of Obstetrics and Gynecology, Kyoto University, Kyoto, Japan
(8) Department of Hematology and Oncology, National University Health System, Singapore
(9) Department of Integrative Biology and Pharmacology, University of Texas Health Science Center at Houston, TX, USA

(10) Institute of Molecular and Cell Biology, A*STAR (Agency for Science, Technology and Research), Singapore
(11) Department of Biochemistry, National University of Singapore, Singapore
(12) Division of Cancer Genomics, Cancer Institute of Japanese Foundation for Cancer Research, 3-8-31 Ariake, Koto-ku, Tokyo, Japan

*Corresponding author: Tel: +65 6516 3241/42; +65 91472036;
Fax: +65 6779 1453;
E-mail: jpthiery@imcb.a-star.edu.sg
**Corresponding author: Tel: +81 3 3570 0450; Fax: +81 3 3570 0454;
E-mail: seiichi.mori@jfcr.or.jp

†These authors contributing equally to this work.
‡Present Address: Division of Cancer Genomics, Cancer Institute of Japanese Foundation for Cancer Research, Koto-ku, Tokyo, Japan

INTRODUCTION

Epithelial ovarian cancer (EOC) is the most lethal gynaecologic malignancy. The global disease burden is approximately 225,000 new cases per year with a survival rate of 30% (Bray et al, 2013). EOC, like most other cancers, represents a heterogeneous collection of distinct diseases that arise as a consequence of varied somatic mutations and epigenetic changes acquired during the process of tumourigenesis and tumour progression. This heterogeneity is apparent in tumour histopathology such as serous, mucinous, endometrioid and clear cell histotypes. It is now established that the discrete histological types differ with respect to variable clinical features, including epidemiological risk, spread patterns, somatic mutations, chemotherapeutic response and patient prognosis (Gilks & Prat, 2009). The histologically distinct subtype, high-grade serous adenocarcinoma, is the most common subtype and accounts for approximately 70% of all ovarian carcinoma. Although this histotype has distinguishing clinical characteristics from the other subtypes, patients with this histological subtype still show diverse outcomes and usually low survival rates, even after the same or very similar treatment regimens (Gilks & Prat, 2009). One possible reason for this low survival rate is that the high degree of heterogeneity of EOC is not considered in the current standard of care (Vaughan et al, 2011). Thus, it is critically important to develop a systematic scheme to dissect the heterogeneity of EOC (Bast et al, 2009; Vaughan et al, 2011).

Genome-scale expression data has been instrumental in characterizing the complex biological diversity of human cancer (Alizadeh et al, 2000; Perou et al, 2000; Verhaak et al, 2010). Subtypes identified through expression microarray analyses are coupled with multiple clinical parameters, such as patient prognosis, age of onset and molecular marker expression (Alizadeh et al, 2000; Perou et al, 2000; Verhaak et al, 2010). Efforts to dissect EOC heterogeneity have correlated expression patterns with clinical features, such as histological types, aggressiveness and patient outcomes (Denkert et al, 2009; Helland et al, 2011; Mok et al, 2009; The Cancer Genome Atlas Research Network, 2011; Tothill et al, 2008). However, due to varied sample sizes and analytical criteria, the reported subtypes of EOC are similar but not completely the same (Helland et al, 2011; The Cancer Genome Atlas Research Network, 2011; Tothill et al, 2008; Verhaak et al, 2013), with reports of six molecular subtypes in 285 serous and endometrioid EOC (Tothill et al, 2008), yet only four molecular subtypes in 489 high-grade serous EOC (The Cancer Genome Atlas Research Network, 2011). Thus, a refined classification scheme with intense phenotypic characterization remains to be established. Also, the molecular targets relevant to cancer cell growth in these transcriptional subtypes have not been identified. The development of diagnostic and therapeutic strategies based on such a scheme is paramount for improving therapeutic efficacy in patients with EOC.

Despite recent successes with molecular targeted therapies for chronic myelogenous leukaemia, ER- or Her2-positive breast cancer, and EGFR-mutated lung cancer, targeted therapies for EOC have not been as encouraging (Quintas-Cardama et al, 2009; Rosell et al, 2010; Yaziji et al, 2004). One approach for the identification of specific targets for EOC subtypes is the use of a genome-wide, systematic, functional assessment of cancer cell growth (proliferation and/or viability). The recent success in suppressing the growth of cultured lung cancer cells with activating *EGFR* mutations by siRNA (Sordella et al, 2004) unveiled the sensitivity of siRNA-based approaches in distinguishing drivers of tumour growth. RNAi libraries, such as The RNAi Consortium (TRC) lentiviral library (Moffat et al, 2006; Root et al, 2006), have enabled systematic genetic studies in mammalian cells, and have identified the genes responsible for proliferation and viability in human cancer cell lines, particularly in the context of synthetic lethality (Barbie et al, 2009; Luo et al, 2008; Scholl et al, 2009).

The TRC library contains 80,000 lentivirally expressing short hairpin RNAs (shRNAs), corresponding to 16,000 human genes. In a systematic screen, a library such as this could be employed to help isolate key regulators of cancer cell growth on a genome-wide scale in a pooled format. Cultured cells would be infected with a pool of the shRNA-expressing lentivirus library such that a typical cell is subjected to only one integration event of an shRNA-expressing lentiviral genome into the host. Infected cells would then be allowed to proliferate for a period of time to permit the amplification or depletion of hairpins accordingly. Although the vast majority of shRNAs have minimal effects on cell proliferation and/or viability, an shRNA that silences the expression of a critical gene will be relatively depleted. Conversely, the relative amplification of an shRNA suggests that it targets a gene with an inhibitory role in cell growth. These integrated hairpins are then subsequently retrieved from the genomic DNA by PCR amplification, and the abundance of each shRNA sequence can be measured with microarray hybridization (Luo et al, 2008) or with next-generation sequencing technology (Sims et al, 2011).

Notably, the successful application of this platform led to the discovery of *PAX8* as having a more essential role in proliferation and survival in ovarian cancer cell lines than in cell lines from other tissues (Cheung et al, 2011). Furthermore, *TBK1* was identified as a synthetic lethal partner of oncogenic *KRAS* in an earlier report using this method (Barbie et al, 2009). Despite these successes, this technology has not been used to identify subtype-specific growth-promoting genes, particularly in the context of ovarian cancer.

Here, we describe a functional genomic approach to dissect the heterogeneity of EOC. We established a large-scale meta-analysis of EOC microarray datasets to determine EOC molecular subtypes. Next, we integrated EOC cell line data into the molecular subtyping scheme to derive an *in vitro* working model representative of each molecular subtype. Finally, we utilized genome-wide shRNA screening to identify molecular targets crucial for cell growth in a selected subtype, which linked the subtype with tubulin polymerization inhibitory drugs.

RESULTS

Molecular heterogeneity of epithelial ovarian cancer
We used a large collection of ovarian tumour gene expression data ($n = 1538$; serous: 1335, mucinous: 27, clear cell: 25,

endometrioid: 96, and others: 55 samples; note that the histological distribution is largely biased toward serous adenocarcinoma as opposed to typical clinical setting) derived from 16 independent studies (Supporting Information Table 1) (Anglesio et al, 2008; Bild et al, 2006; Bowen et al, 2009; Denkert et al, 2009; Hendrix et al, 2006; Hogdall et al, 2003; Hsu et al, 2007; Iorio et al, 2010; Jochumsen et al, 2007, 2009; Mok et al, 2009; Pejovic et al, 2009; The Cancer Genome Atlas Research Network, 2011; Tone et al, 2008; Tothill et al, 2008; Tung et al, 2009). Among the 16 datasets, the dataset from TCGA was the largest in sample number ($n = 406$; 26.4% of all samples). All publicly available datasets were included at the time of the study (April 2010), and compiled with an Oslo cohort dataset (BD and JMN). A strong batch-effect was removed by ComBat, eliminating technical differences across data collection sites, while conserving meaningful variations (Supporting Information Fig 1A and B) (Chen et al, 2011; Johnson et al, 2007). A preliminary statistical power analysis showed that 1500 or more samples were required to achieve sufficient statistical power (≥ 0.8) in capturing the complexity and dynamicity of EOC (Supporting Information Fig 2; Supporting Information Materials and Methods) (Fox & Mathers, 1997). In this collection, known prognostic factors were correlated with patient overall survival by univariate and multivariate Cox proportional hazards analyses (Table 1).

To identify EOC molecular subtypes, we applied consensus clustering (CC) to the collection and detected five clusters (Fig 1A) that were characterized by markers of differentiation or cell-type status and stromal components, including the presence of infiltrated inflammatory cells (Supporting Information Table 2). Subtypes were annotated by applying single sample gene set enrichment analysis (ss-GSEA) (Verhaak et al, 2010) with literature-curated gene signatures for epithelial, mesenchymal and stem cells (Supporting Information Text), and confirmed this characterization with the use of appropriate markers. The silhouette plot and SigClust (Liu et al, 2008b) analysis confirmed tumour similarity within each subtype, indicating the robustness of the classification (Supporting Information Fig 3A). The subtype distribution by cohorts and histology is presented in the Supporting Information Text and Supporting Information Figs 4A and B. Subtype distribution within the samples, taken by laser capture microscopy (GSE10971, GSE14407 and GSE18520), implied that the subtypes were intrinsic to cancer cells, and not dependent on stromal cells (Supporting Information Text).

We compared our subgrouping with a previous classification (285 samples; GSE9891) included in our combined dataset (Tothill et al, 2008). An overall concordance of 82.9% for all of the subtypes was found (Supporting Information Table 3; Supporting Information Fig 3B); thus, our large-scale analysis confirmed the previous study, and provided finer distinctions not detectable with fewer samples. Also, we noted that the proposed molecular subtypes were akin to that of serous ovarian carcinoma as proposed by The Cancer Genome Atlas Research Network (2011) (Supporting Information Fig 3B). However, the subtyping schemes from the previous studies did not show a one-to-one match with our proposed classification (Supporting Information Table 3; Supporting Information Fig 3B; Supporting

Information Text; see the mutual relationships among Epi-A or Epi-B/C2, C3 or C4/Immunoreactive or Differentiated). This discrepancy may suggest a shared biological feature across these subgroups and hence may cause an imperfect distinction among the subtypes with predictive models as described later (Fig 1D; Supporting Information Fig 8C; Supporting Information Table 8; Supporting Information Text). We also noted that TCGA molecular subtyping did not include a Stem-B/C6 population (Supporting Information Fig 3B; Supporting Information Text). The proposed subtypes in the current study are similar to the previously identified molecular subtypes yet reveal novel biological features.

Correlation of subtype with clinicopathological parameters

We correlated the subtypes with various clinicopathological parameters to ascertain their clinical relevance (Supporting Information Fig 6A; Supporting Information Tables 4A and B; note that the clinicopathological information obtained with each dataset was neither standardized nor centrally reviewed across the datasets; therefore, there might be misdiagnosed or mis-evaluated samples included). We found a significant correlation between subtype and patient outcome: Epi-A, Epi-B and Stem-B subtypes had a better prognosis in a Kaplan–Meier analysis (Fig 1B), while Mes and Stem-A tumours were linked with poorer outcomes. The Mes subtype included more advanced staged and metastasized tumours (Supporting Information Fig 6A; Supporting Information Tables 4A and B), whereas some Stem-A tumours were already found to be at stages 1 and 2 (Supporting Information Fig 6B), with poorer outcomes than those of other subtypes, even at stages 1 and 2 (Supporting Information Fig 6B), Furthermore, Stem-A tumours were enriched in older patients (Supporting Information Fig 6A; Supporting Information Tables 4A and B). The Stem-B subtype, on the other hand, was characterized by multiple histological types, including the majority of mucinous, endometrioid and clear cell carcinoma and some serous carcinoma (Supporting Information Figs 4B, 5 and 6A; Supporting Information Tables 4A and B). Focusing solely on serous tumours (Supporting Information Fig 6D), the frequency of Epi-A-classified tumours decreased significantly as tumour classification moved from serous tumours with low malignant potential (LMP) through to high-grade tumours, whereas the opposite shift in pattern was true for Mes and Stem-A serous tumours. All subtypes displayed high-grade serous carcinoma, with distinctions in survival in Kaplan–Meier curves (Supporting Information Fig 6C). The effect of molecular subtyping on prognosis was significant in both the univariate and multivariate Cox regression analyses with multiple combinations of clinically relevant parameters and status (Table 1; Supporting Information Tables 5A–E; Supporting Information Text).

Clear distinctions were also observed in the enrichment of the gene expression signatures for various pathways. The ss-GSEA analysis of 1538 samples using 6898 gene sets (GSEA databases Supporting Information Table 6) revealed a subtype-specific enrichment of 207 gene sets (Fig 1C; Supporting Information Table 7) (Subramanian & Simon, 2011). Mes tumours correlated with *Metastases* and *TGF-β-related* pathways, consistent with

Table 1. Univariate and multivariate Cox proportional hazards regression analysis for multiple clinical variables and tumour subtypes.

Clinical variables	Sample size (total $n = 539$)	Univariate (HR, 95% CI)	p-value	Multivariate (HR, 95% CI)	p-value
Age (year)					
<55	175 (32.47%)	1		1	
≥55	364 (67.53%)	1.403 (1.071–1.839)	0.0141	1.285 (0.9781–1.687)[a]	0.07173[a]
Stage					
I or II	47 (8.72%)	1		1	
III or IV	492 (91.28%)	3.907 (1.843–8.285)	0.00038	3.429 (1.591–7.389)[a]	0.00165[a]
Grade					
1	17 (3.15%)	1		1	
≥2	522 (96.85%)	2.58 (0.9578–6.949)	0.0608	1.365 (0.494–3.763)[a]	0.54799[a]
Metastasis					
Primary	500 (92.76%)	1		1	
Metastasis	39 (7.24%)	1.349 (0.8323–2.185)	0.224	1.391 (0.854–2.27)[a]	0.1853[a]
Subtype					
Non-Epi-A	483 (89.61%)	1		1	
Epi-A	56 (10.39%)	0.7103 (0.4498–1.122)	0.142	0.9449 (0.5834–1.53)[b]	0.8176[b]
Non-Epi-B	384 (71.24%)	1		1	
Epi-B	155 (28.76%)	0.69 (0.5206–0.9144)	0.0098	0.7347 (0.5532–0.976)[b]	0.033[b]
Non-Mes	361 (66.98%)	1		1	
Mes	178 (33.02%)	1.171 (0.907–1.513)	0.225	1.01 (0.7771–1.324)[b]	0.9164[b]
Non-Stem-A	411 (76.25%)	1		1	
Stem-A	128 (23.75%)	1.417 (1.075–1.868)	0.0135	1.382 (1.045–1.83)[a]	0.0234[a]
Non-Stem-B	517 (95.92%)	1		1	
Stem-B	22 (4.08%)	1.204 (0.6383–2.271)	0.567	1.14 (0.6033–2.149)[b]	0.6886[b]

Epi-A, epithelial-A; Epi-B, epithelial-B; Mes, mesenchymal; Stem-A, stem-like-A; Stem-B, stem-like-B.
p-values below 0.05 are shown in red.
[a]Multivariate Cox regression analysis of clinical variables with Stem-A subtype.
[b]For multivariate Cox regression, each subtype was independently analysed with the other clinical variables (age, stage, grade and metastasis) from the remaining subtypes.

their link with epithelial–mesenchymal transition (EMT) and metastasis (Supporting Information Fig 6A) (Maruyama et al, 2000; Yin et al, 1999). In comparison, *chromatin modification* gene sets were highly enriched in the Stem-A subtype (Fig 1C; Supporting Information Table 7). Overall, this expression-based subtyping scheme dissected ovarian serous carcinoma heterogeneity into subgroups with similar biological properties.

Predictive framework for EOC subtype classification

We next developed a predictive model with BinReg as a potential diagnostic tool for quantitative gene expression-based subgroup assignment (Supporting Information Fig 7A and B) (Gatza et al, 2010). This was performed using microarrays of representative samples for each subtype ($n = 50$ per subtype). Fig 1D shows predicted probabilities for subtype status of the remaining samples ($n = 1413$) not used in building predictive model. A comparison of the subtype predicted by BinReg with that classified by the CC (Fig 1A) revealed an overall 78.8% concordance for all subtypes (78.5% for core samples) (Fig 1D; Supporting Information Table 8), and a highly similar pattern of patient outcomes (Fig 1B; Supporting Information Fig 7C). This demonstrated the powerful predictive capability of the method, with the concordance comparable with those reported in previous studies for multiple breast cancer cohorts (Supporting Information Text) (Calza et al, 2006; Haibe-Kains et al, 2012). We affirmed the accuracy of this method using 10-fold cross-

validation (Supporting Information Figs 8A–C) (Blum et al, 1999; Kim, 2009; Konavi, 1995), 3-way split cross-validation (Ewens & Grant, 2001), and also by comparing BinReg to ClaNC (Supporting Information Fig 9; Supporting Information Materials and Methods).

To ensure the robustness of the classifier, we performed validation on five independent ovarian cancer datasets (total $n = 418$; Supporting Information Table 1) (King et al, 2011; Konstantinopoulos et al, 2010; Meyniel et al, 2010) that were not included in the prediction modelling. We observed high concordance for the gene expression patterns and clinicopathological characteristics in the predicted molecular subtype (Fig 1E; Supporting Information Tables 4A, C and D). Using 260 samples from the validation set (GSE19829 [$n = 28$], GSE30311 [$n = 47$] and GSE26712 [$n = 185$]), for which patient outcome information was supplied (Konstantinopoulos et al, 2010), the Kaplan–Meier analysis on the BinReg-predicted molecular subtypes revealed a similar pattern of patient prognoses with that of the original CC analysis ($p = 0.0372$ by the log-rank test; Fig 1B; Supporting Information Fig 7D) for subtypes other than Stem-B (Supporting Information Text). ClaNC (Dabney, 2006; Verhaak et al, 2010) further confirmed the highly comparable and predictive capability of this EOC subtyping (Supporting Information Fig 8D). Thus, the molecular subtype prediction model can assign clinical samples with unknown subtype status with high accuracy.
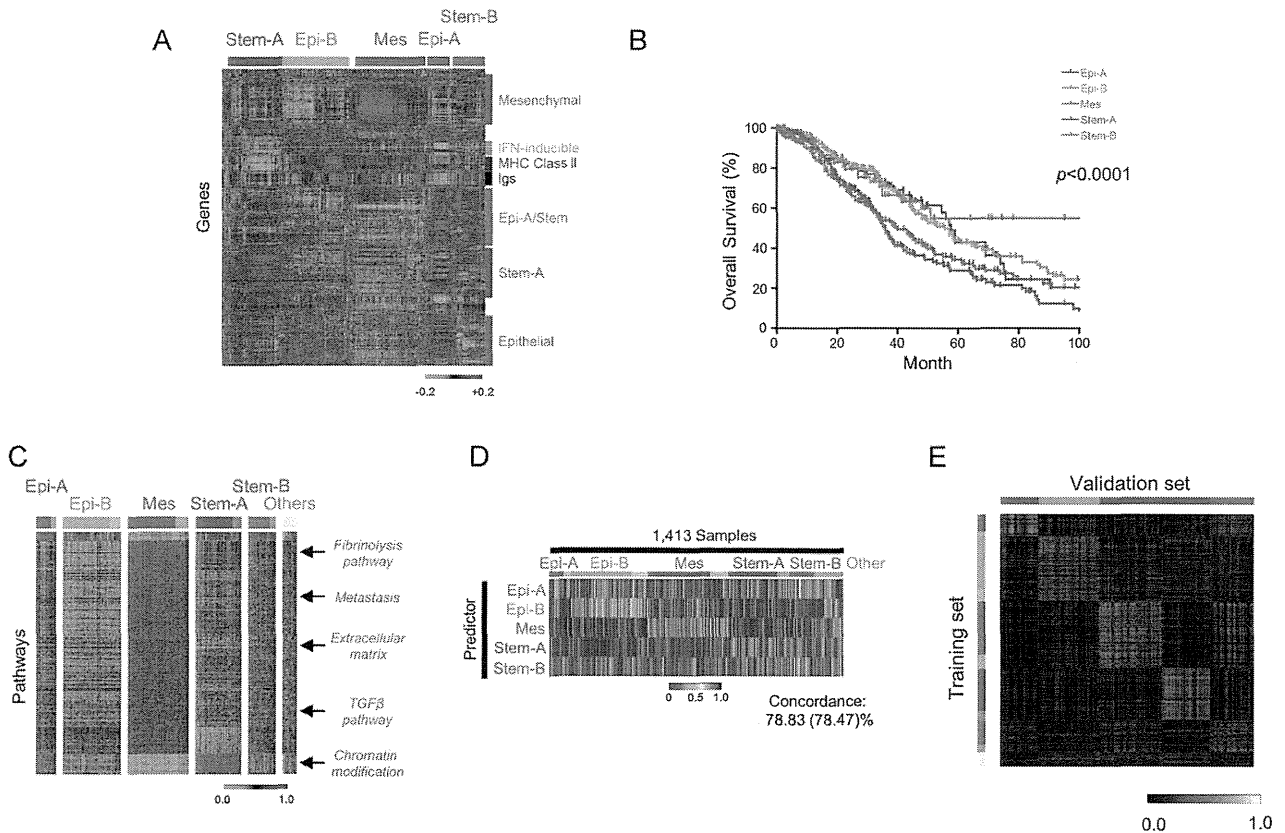
Figure 1. CC analysis revealed five subtypes of epithelial ovarian carcinoma.

A. Gene expression heatmap for the five tumour clusters (red = high; green = low expression). CC of 1538 samples identified five subtypes, designated by the associated gene components. Note the similarities between Epi-A/Stem-B subtype tumours, between Epi-A/Epi-B subtypes for epithelial genes, and the expression pattern of Epi-A/Stem genes. Also note that none of cultured cell-line data was included in this analysis.

B. Kaplan–Meier survival analysis for each subtype. Among data for 1538 patient samples, survival information for 978 samples was available (GSE3149: 143, GSE9891: 277, TCGA: 400, GSE14764: 80, GSE18520: 53 and Oslo cohort: 25 samples) (Epi-A: 80, Epi-B: 264, Mes: 284, Stem-A: 220, Stem-B: 61 and others: 69 samples) and used for the Kaplan–Meier analysis.

C. Subtype-specific pathway enrichment. Heatmap shows subtype-specific single sample gene set enrichment analysis (ss-GSEA) scores (false discovery rate (FDR) in significance analysis of microarrays (SAM) $q = 0\%$, receiver operating characteristic (ROC) >0.85) for 1538 ovarian cancer samples. Red = high; green = low enrichment scores. Gene sets are aligned in descending value of ROC. Samples are aligned by subtype classification and SW. Deep colour = positive SW (core samples); pale colour = samples classified, but negative SW. "Others" indicates the unclassified samples not grouped in any of the five subtypes in the initial CC analysis in Fig 1A. Arrows indicate positions of selected pathways.

D. Ovarian cancer subtype predictors (BinReg). A heatmap is shown for the predicted probabilities of subtype status on 1413 clinical samples not used in the subtype predictor generation. Red = high; blue = low. Samples were aligned according to subtype classification by CC and SW. Colour as for (C). "Others" is represented as for (C).

E. Heatmap of Spearman correlation *Rho* between the subtype of training data ($n = 1538$) and the BinReg predicted subtype of samples in five independent datasets (GSE19829, GSE20565, GSE30311, GSE26712 and GSE27651; total $n = 418$). The validation samples are aligned horizontally according to the predicted subtype, whereas the training samples are aligned vertically according to the subtype. Yellow = high correlation; black = low correlation. Abbreviations: Epi-A, epithelial-A; Epi-B, epithelial-B; Mes, mesenchymal; Stem-A, stem-like-A; Stem-B, stem-like-B.

Identification of representative cell lines for each subtype

Cell lines corresponding to each EOC subtype were identified for *in vitro* modelling. We performed two rounds of CC on a pool of datasets from 142 cultured EOC cell lines, resulting in Epi-A: 29, Epi-B: 10, Mes: 34, Stem-A: 42 and Stem-B: 27 cell lines (Supporting Information Figs 10A and B); the results were unambiguously supported by similarity matrices, the silhouette values with significant *p*-value by SigClust (Fig 2A) (Liu et al, 2008b), as well as consistent subtype assignments amongst

biological replicates of 28 cell lines (Supporting Information Table 9; Supporting Information Text). The cell-line subtype predictors (Fig 2B) were then applied to tumour core samples to estimate the molecular similarity of the subtypes between *in vivo* tumours and *in vitro* cell lines. We observed a high level of accuracy in the *area under the curve* (AUC: 0.744 to 0.918) and a high concordance between the predicted tumour subtype by a cell-line classifier with the initially assigned tumour subtype (75.8–87.9%) (Fig 2B). Furthermore, we found a high
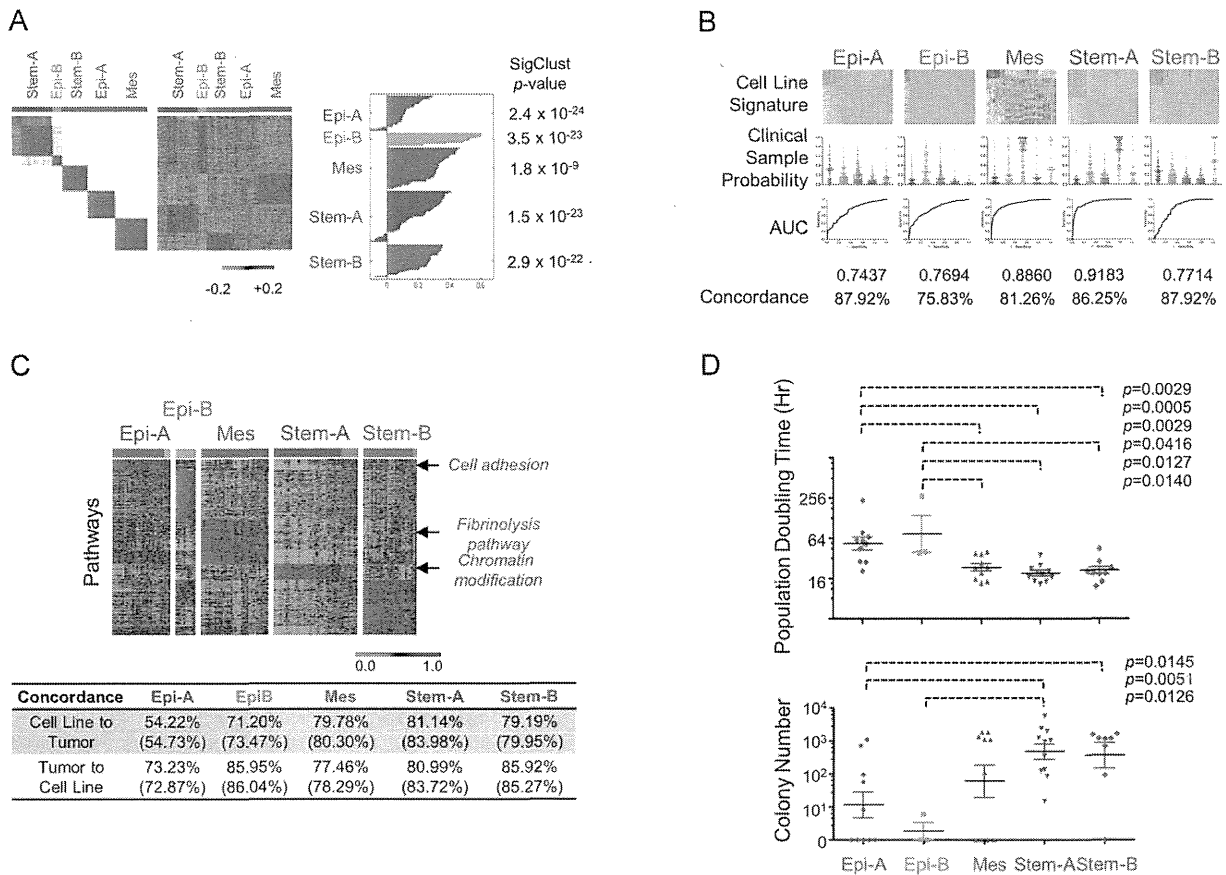
Figure 2. Identification of cell line subtype status.

A. Five subtypes in ovarian cancer cell line classification. Left panel. CC matrix of 142 ovarian cell lines. Red = high; white = low similarity. Middle panel. Gene expression heatmap of ovarian cell lines. Red = high; green = low expression. Right panel. Silhouette analysis for each subtype. Column to the right of silhouette plot is the SigClust (Liu et al, 2008b) p-value indicative of cluster significance for each subtype.

B. Prediction of clinical samples by cell line predictors using BinReg. Upper panel. Gene expression heatmaps for subtype predictors based on cell line expression data. Red = high; blue = low expression. Middle panel. Predicted probability of core clinical samples for cell-line subtype predictor by BinReg. Each subtype signature detected the probability difference between the corresponding subtype from the remaining subtypes with statistical significance ($p < 0.0001$; Mann–Whitney U-test). Lower panel. Receiver operating characteristic (ROC) analyses of subtype predictors. Overall accuracy is shown by the area under the ROC curve (AUC) (Pejovic et al, 2009). Concordance (%) of the subtype status derived from CC with the prediction based on the cell line subtype predictors.

C. Upper panel. Cell line subtype-specific pathway enrichment. Subtype-specific single sample gene set enrichment analysis (ss-GSEA) scores (false discovery rate (FDR) of the significance analysis of microarrays (SAM) $q = 0\%$, ROC > 0.85 as overexpressed gene sets) for 142 ovarian cell lines are shown as a heatmap. Red = high; green = low enrichment scores. Gene sets aligned in descending value of ROC; samples are aligned according to the subtype classification by CC and the SW. Deep colour = positive SW (core samples); pale colour = samples classified to a subtype, but negative SW. Arrows indicate positions of selected pathways. Lower panel: Concordance (%) of the subtype status (from CC by genes) with the prediction result (from BinReg based on the subtype predictors by enrichment scores). The number in parentheses indicates the accuracy of the prediction against core samples.

D. Characterization of *in vitro* phenotypes of cell lines in each subtype. Upper panel. Population doubling time of a cell line was measured with the MTS assay (Matsumura et al, 2011) and is shown as dot plots. Lower panel. Anchorage-independent cell growth ability for each cell line was measured using the methylcellulose assay (Mori et al, 2009). $\log_{10}$-transformed colony number is shown. p-values were computed by Mann–Whitney U-test. Abbreviations: Epi-A, epithelial-A; Epi-B, epithelial-B; Mes, mesenchymal; Stem-A, stem-like-A; Stem-B, stem-like-B.

correlation between clinical tumour subtype and cell line subtype in the Spearman correlation map analysis (Supporting Information Fig 10C). These findings indicated a high level of similarity between ovarian cancer cell lines and tumour transcriptomic expression patterns (Fig 2B; Supporting Information Fig 10C).

We next compared the pathway activation for these 142 cell lines with that of the clinical tumours using ss-GSEA analysis

(Figs 1C and 2C; Supporting Information Table 10). Epi-A cell lines were characterized by cell adhesion-related gene sets, reflecting enrichment of epithelial cell markers. Importantly, 33 of the 402 cell line subtype-specific gene sets were shared with tumours, including enrichment of *fibrinolysis pathway* and *chromatin modification* in the Mes and Stem-A subtypes, respectively (Supporting Information Table 10); this was confirmed with BinReg analyses using a statistical model with

pathway enrichment scores (Fig 2C). We estimated the subtype status of clinical samples by fitting a Bayesian probit regression model with the subtype-specific enrichment scores for cell lines. Reverse estimations were also performed from the tumour samples to the cell lines. By applying the same method as in Fig 2B, we observed high levels of concordance between the predicted subtype of tumours by the cell-line ss-GSEA pathway classifier with the initially assigned tumour subtype (54.2–81.1%) and reciprocally high concordance between the predicted cell line subtype by a tumour ss-GSEA pathway classifier with the original cell line subtype (72.9–86.0%). These results indicated strong similarity between cell lines and tumours in the pattern of pathway enrichment (Fig 2C). We then correlated the *in vitro* phenotypes for the molecular subtypes, and identified a significant correlation between cell line subtypes with population doubling time and anchorage-independent cell growth potential (Fig 2D). Epi-A and Epi-B cell lines had longer population doubling times and decreased colony-forming ability, which may reflect the less-aggressive behaviour of clinical tumours. Overall, these cell lines can serve as good experimental models for each molecular subtype.

Genome-wide shRNA screens identified subtype-specific growth-promoting genes

Genes essential to each subgroup were investigated via genome-wide screens using the pooled TRC shRNA library, with the presumption that tumours within the same subtype would share molecular mechanisms for their growth (proliferation and/or survival). The experimental strategy of the screen is shown in Fig 3A. Briefly, we conducted pooled shRNA screens on 14 ovarian cell lines, representing Epi-A, Mes and Stem-A subtypes, that differ profoundly in gene expression and clinical properties (Figs 1A and B) (4 Epi-A: OVCA429, OVCAR-8, OVCA433, PEO1; 5 Mes: ovary1847, HEY, HeyA8, HeyC2, SKOV-3; and 5 Stem-A: A2780, CH1, PA-1, SKOV-4, SKOV-6). These 14 cell lines were selected based on their high silhouette width (SW) values for the subtype signature in order to screen with "more representative" cell lines for a given subtype, with a notion of PA-1 as a teratocarcinoma cell line (Supporting Information Table 11).

Two independent screens were performed to ensure reproducibility. The initial assay was designed to determine concordance among four experimental replicates of a single cell line per subtype (OVCA433, HeyA8 and PA-1 was used to represent Epi-A, Mes and Stem-A subtypes, respectively). Spearman correlations confirmed tight correlations among the quadruplicates in the screen (Spearman $rho = 0.7528 \pm$ SEM $0.0113, p < 10^{-16}$). The second screen was performed in 14 cell lines with the intention to detect differences across subtypes as well as concordance among different cell lines within a subtype. Since the screenings detected similarity in subtype-specific depletions or amplifications of hairpins, we combined both datasets and further performed RIGER analyses (Luo et al, 2008) on the compiled data. Supporting Information Fig 12A illustrates highly distinctive genome-wide patterns in the copy number of subtype-specific shRNAs that were depleted or amplified. The effect size was reasonably large (Cohen, 1988; Monk et al, 2012; Syrjanen & Syrjanen, 2013): the mean effect sizes of depleted

hairpins were Epi-A $= -0.9098$; Mes $= -0.7681$ and Stem-A $= -0.7818$, and those of amplified hairpins were Epi-A $= 0.8128$, Mes $= 0.8282$ and Stem-A $= 0.7486$ (Supporting Information Fig 12B; Supporting Information Table 12).

The primary aim of the screens was to identify genes that, when inhibited, would render growth suppression on a certain molecular subtype. To this end, we identified depleted shRNAs targeting 77 genes for Epi-A, 85 genes for Mes, and 88 genes for Stem-A subtypes (Fig 3B), with high significance in subtype enrichment ($q < 0.005$) and Hairpin Score (>0.2). These genes are potentially involved in growth promotion of the cells in a given subtype (Supporting Information Table 12). Conversely, we identified amplified hairpins targeting 43 genes for Epi-A, 72 genes for Mes, and 44 genes for Stem-A (Fig 3B) that may have a suppressive effect on cell growth of the given subtype under conventional culture conditions (Supporting Information Table 12). For most of the growth-related functional genes, the abundance of shRNAs did not show significant correlation to gene expression, implying that the functional relevance of the genes was independent of their expression levels. Differences in experimental design and detection platforms hampered the integration of the results from this screen with that of another published screen using the same shRNA library (Supporting Information Materials and Methods) (Cheung et al, 2011).

Validation of subtype-specific growth promoting genes

To validate the effects of the genes identified from the screens, we focused on the Stem-A subtype (given its worse clinical outcome) and targeted individual genes with siRNA (Fig 3C). We chose 135 genes depleted in Stem-A subtypes based on a less stringent q-value cut-off of 0.03 from RIGER analysis (note that a more stringent q-value was used in Fig 3B; Supporting Information Table 13). The validation of these 135 genes was performed in a process that consisted of four steps (Fig 3C; with more details available in "Materials and Methods") in order to identify siRNAs that inhibited growth on Stem-A cells but had a minimal effect on other cells. Stem-A-specific essential genes were identified as positive hits based on the following comparisons using Student t-tests: (1) comparison between the growth inhibitory effect of silencing the gene of interest with that of the siRNA negative controls in the Stem-A cells; and (2) comparison between the effect on Stem-A cells with that on the references for the subtype (non-Stem-A cells) (Fig 3C). Relying on criteria of $\geq 20\%$ growth suppression in PA-1 with $p < 0.001$ in a Student's t-test comparing control with the gene of interest and $\geq 20\%$ growth suppression in PA-1 as compared with the reference cell line, 28 genes were found in the first step of validation to be selective for PA-1 cell growth (Supporting Information Table 13). In the second step, we examined the effect of these 28 genes in PA-1, HeyA8 and OVCA433, and further confirmed the growth suppressive effect of 14 of these 28 genes (Supporting Information Table 13). For the third step, we switched platforms from "siGenome" to "On-Target Plus siRNA" to further validate our observations using different sets of target sequences in the genes as well as to reduce possible off-target effects. After this step, five genes (*TUBGCP4, NAT10, GTF3C1, BLOC1S1* and *LRRC59*) were validated as PA-1-relevant

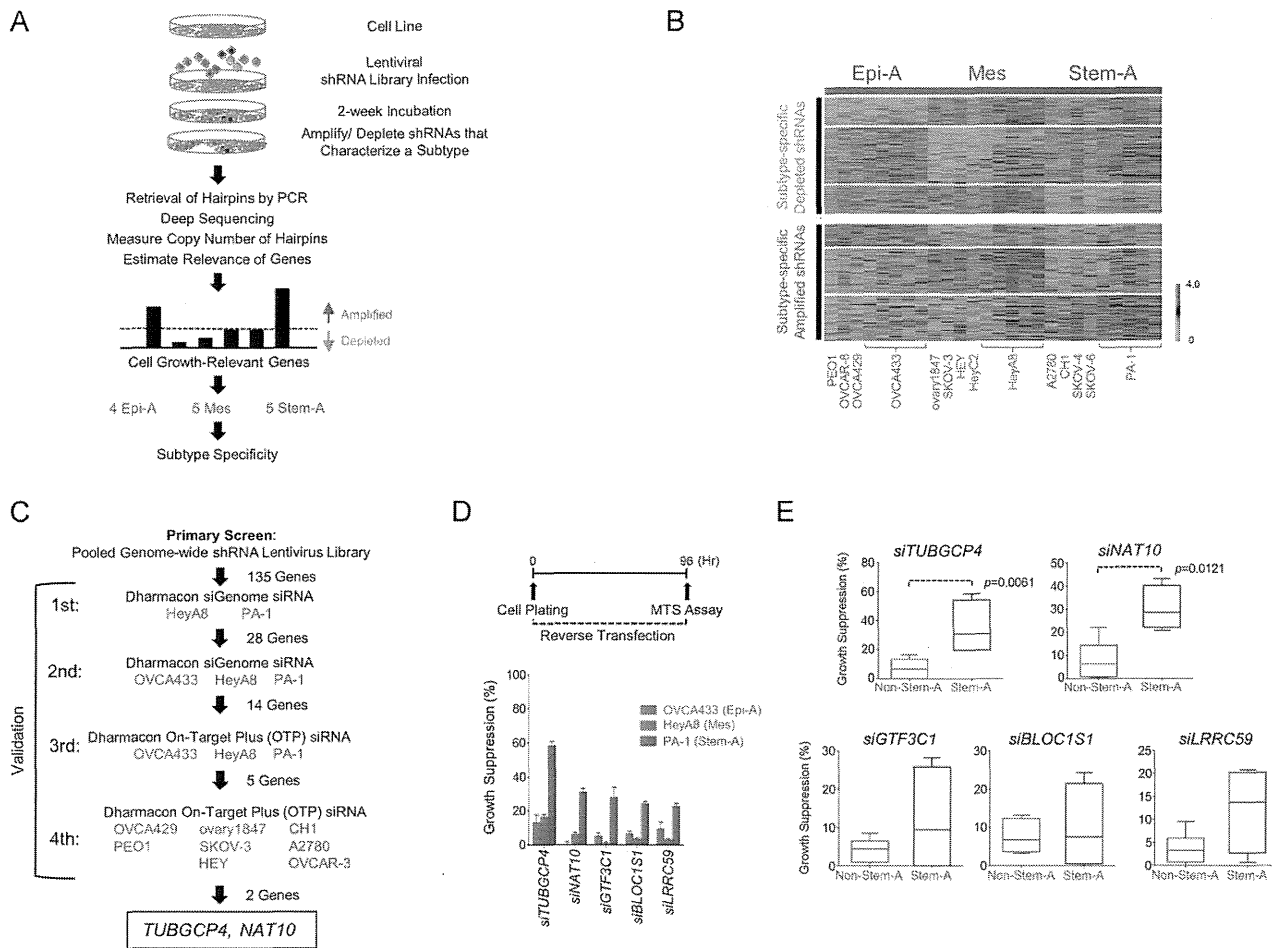**Figure 3. Subtype-specific functional relevance genes.**

A. Schematic showing identification of functionally relevant genes for cell growth in a subtype-specific manner.

B. Gene centred and normalized heatmap, compiled from two independent screens, shows hairpins selectively depleted or amplified in each subtype. The quadruplicates of three cell lines (OVCA433; Epi-A, HeyA8; Mes and PA-1; Stem-A) were assayed in the initial screen, while the second screen used one experimental replicate of 14 different cell lines (4 Epi-A: OVCA429, OVCAR-8, OVCA433, PEO1; 5 Mes: ovary1847, HEY, HeyA8, HeyC2, SKOV-3 and 5 Stem-A: A2780, CH1, PA-1, SKOV-4, SKOV-6). Using reads with a perfect match to the reference sequences (Sigma–Aldrich), the copy number of each hairpin was counted and normalized against the total number of reads in a sample and then rendered to RIGER analysis to find phenotype-specific, functionally relevant genes (Luo et al, 2008). Top panel. Subtype-specific depleted hairpins in Epi-A, followed by Mes and Stem-A subtypes. Each row represents shRNA hairpin copy number and is sorted according to the hairpin score identified in RIGER (Luo et al, 2008). Only hairpin scores ≥ 0.2 and genes significantly enriched in a subtype ($q < 0.005$) are shown. Bottom panel. Subtype-specific amplified hairpins arranged as in the top panel. Red = higher; green = lower copy number counts.

C. Schematic of siRNA experiments validating the identified Stem-A-specific growth-promoting genes. This analysis led to the identification of two functionally relevant genes specific to Stem-A: *TUBGCP4* and *NAT10*.

D. Validation of subtype-selective effect of the genes on cell growth by siRNAs. Upper panel. Timeline of assay performed for the siRNA reverse-transfection experiment. Lower panel. Effect of gene knockdown on cell growth (bar plots) as a percentage ratio of growth suppression, normalized against the negative controls. Error bar indicates the SEM of three independent experiments. Stem-A-selective growth suppression effect is shown for the inhibition of the five validated PA-1 (Stem-A)-specific growth-promoting genes in OVCA433, HeyA8 and PA-1, respectively. Green = OVCA433 (Epi-A); red = HeyA8 (Mes); blue = PA-1 (Stem-A).

E. Effect of silencing PA-1 (Stem-A)-selective genes on cell growth in other ovarian cancer cell lines. The five PA-1-selective genes were silenced individually by siRNA in non-Stem-A (OVCA433, OVCA429, PEO1, HeyA8, ovary1847, SKOV-3 and HEY) and Stem-A (PA-1, CH1, A2780 and OVCAR-3) cell lines in three independent experiments, and examined for their effect on cell growth relative to the negative control. Averaged percentages of growth suppression in each group are shown as a box plot and were statistically evaluated using Mann–Whitney U-test with GraphPad Prism. Bottom, middle and top lines of each box represent the 25th percentile, median and 75th percentile, respectively, and whiskers extend to the most extreme values of the group. Inhibition with si*TUBGCP4* or si*NAT10* significantly suppressed cell growth of Stem-A cell lines as compared to non-Stem-A cell lines. Grey = non-Stem-A cell lines; blue = Stem-A cell lines. Abbreviations: Epi-A, epithelial-A; Mes, mesenchymal; Stem-A, stem-like-A.

genes (Fig 3D). Importantly, PA-1 cells showed increased cleavage of Caspase-3 and PARP after treatment with si*TUBGCP4*, si*NAT10*, si*GTF3C1* or si*LRRC59*, indicating activation of apoptosis in these cells (Supporting Information Fig 13C). Finally, as the fourth step of the validation process, the experiments were conducted with use of additional non-Stem-A (Mes: ovary1847, SKOV-3 and HEY; Epi-A: OVCA429 and PEO1) and Stem-A (CH1, A2780 and OVCAR-3) cell lines to ensure its reproducibility and to exclude any possible impact of PA-1 cells being derived from a different cell-of-origin (teratocarcinoma), even though it had the highest SW of the Stem-A cell lines. *TUBGCP4* or *NAT10* siRNA treatment reproducibly resulted in a statistically significant reduction in cell growth for the Stem-A cell lines, while cell growth for non-Stem-A cell lines was not affected (Fig 3E). These multiple stages of rigorous validation confirmed the dependence of Stem-A cell lines on *TUBGCP4* and *NAT10* in cell growth and ensured that this effect was not limited to PA-1 cells. Silencing of the other three genes (*GTF3C1*, *BLOC1S1* and *LRRC59*), albeit not statistically significant, also exhibited a tendency toward differential toxicity in Stem-A cells (Fig 3E). These observations demonstrate that subtype classification based on gene expression is indeed mirrored by patterns of functional genetic determinants of cell viability. Moreover, the validated genes can provide us with an insight into the molecular mechanisms of Stem-A tumour growth.

Microtubules as potent targets in Stem-A subtype

TUBGCP4 is a component of γ-tubulin ring complex, which is critical for nucleation of tubulin complexes in the cell (Fava et al, 1999; Moritz et al, 1995, 1998). NAT10 is reported as a possible acetyl transferase of α-tubulin that may be involved in the stabilization of microtubules (Hubbert et al, 2002; Shen et al, 2009). The selective effect of si*TUBGCP4* or si*NAT10* on Stem-A cell lines (Fig 3E) may suggest that the Stem-A cell lines are more susceptible to mitotic inhibition than other subtype cell lines. An examination of the expression data of clinical tumours and cell lines revealed higher activity in the enrichment score of microtubule/tubulin-related pathways for Stem-A than that for non-Stem-A subgroups ($p = 6.6 \times 10^{-67}$ and $p = 2.1 \times 10^{-6}$ by Mann–Whitney *U*-test, respectively; Fig 4A; Supporting Information Table 16) (Verhaak et al, 2010). In addition, *TUBGCP4* knockdown resulted in a down-regulation of the *Microtubule* gene set in the transcriptome across Epi-A, Mes and Stem-A cell lines (Supporting Information Fig 13B; Supporting Information Table 14; Supporting Information Text).

These findings prompted us to examine the *in vitro* sensitivity of Stem-A cells to microtubule-targeted drugs such as paclitaxel, vincristine and vinorelbine using a panel of 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). A growth inhibitory concentration of 50% (GI50; drug concentration for 50% growth inhibitory effects on cells) was measured for each cell line in at least three independent experiments. The Stem-A cell lines were found to be more sensitive to inhibitors of tubulin polymerization, vincristine and vinorelbine (Lobert et al, 1996), than non-Stem-A cell

lines (Fig 4B). In contrast, paclitaxel, a drug that stabilizes microtubules (Manfredi & Horwitz, 1984), resulted in no significant distinction between the two subgroups (Fig 4B). Moreover, 48-h vincristine treatment caused apoptosis in Stem-A cell lines at 1.2 nM (Fig 4C), whereas minimal or no apoptosis was observed in non-Stem-A cell lines, even at 10 nM concentrations (Fig 4C). Taken together, these findings provide evidence that drugs targeting tubulin polymerization can be useful in treating patients with Stem-A EOC with poor clinical outcomes.

DISCUSSION

Using a large collection of EOC samples, we identified five molecular subtypes (Epi-A, Epi-B, Mes, Stem-A and Stem-B) that exhibited distinct clinicopathological characteristics and rates of overall survival. Of these, Epi-B and Stem-A subtypes were found to be independent prognostic factors. We established a prediction model for these subtypes and validated this model on an independent dataset. For the first time, using a genome-wide shRNA screen, we found that subtype-matched cell lines have distinct vulnerabilities. In particular, the poor-prognosis Stem-A subtype exhibited elevated microtubule activity and was sensitive to several microtubule polymerization inhibitor drugs, such as vincristine and vinorelbine. These results offer possible therapeutic strategies to target specific subtypes of EOC.

Multiple clinicopathological parameters are linked with prognosis in EOC patients, such as age at diagnosis, peritoneal dissemination, metastasis to distant organs/lymph nodes, and response to platinum-based standard chemotherapy (Gilks & Prat, 2009). Here, we add transcriptional subtype as an additional prediction parameter. Although a correlation between the Mes subtype and patient prognosis was detected with a log-rank test, it was masked in the multivariate Cox analysis; this suggests that the Mes subtype may be confounded in the analysis because it is significantly enriched in tumours at a more advanced stage. Nevertheless, since Stem-A and Epi-B subtypes were detected as significant independent prognostic factors in both the univariate and multivariate analyses, this demonstrates the clinical importance of our classification scheme. Of note, a previous study of 489 samples could not correlate their molecular classification with patient overall survival, although a more recent study correlated two TCGA subtypes with relapse-free survival using the same cohort (The Cancer Genome Atlas Research Network, 2011; Verhaak et al, 2013). This is perhaps derived from a bias internal to the cohort, and suggests the need for a substantial number of samples, which is provided by combining multiple datasets, as presented here.

Genomic profiling aimed at dissecting the complexity of cancer could provide further opportunities for the identification of relevant molecular targets. However, a major challenge is to identify cell lines that reflect the relevant underlying tumour biology (Chin et al, 2011). Expression studies of cultured breast cancer cell lines have shown that *in vitro* cells retain subtype characteristics corresponding to those of their *in vivo*
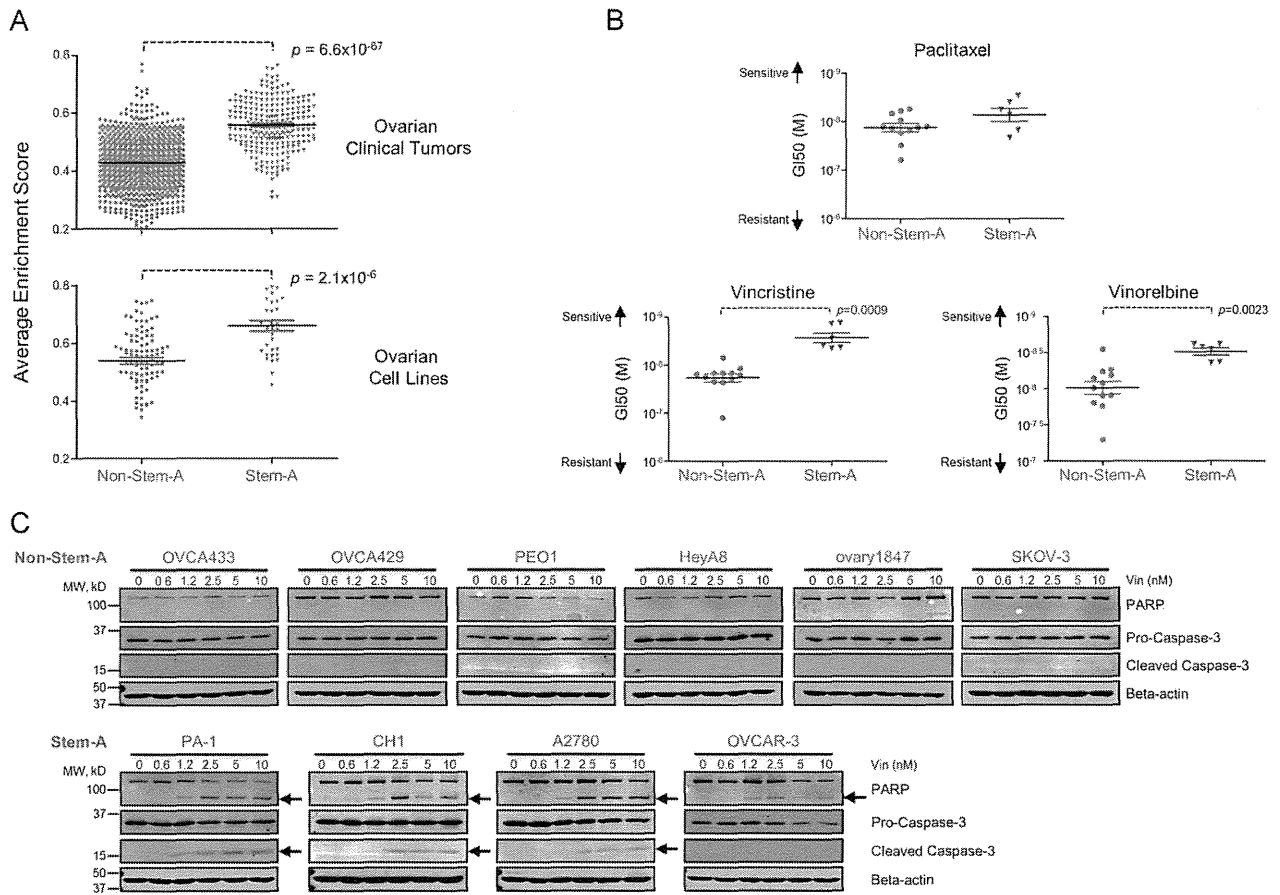
Figure 4. Susceptibility of Stem-A cells to microtubule assembly inhibitors.

A. Estimated microtubule activity in non-Stem-A and Stem-A subgroups of ovarian cancer. Microtubule activity in 1142 core samples of ovarian clinical tumours (Top panel) and in 129 core samples of ovarian cell lines (Bottom panel) was estimated based on the average single sample gene set enrichment analysis (ss-GSEA) enrichment score of 19 microtubule-related gene sets (Supporting Information Table 16) acquired from GSEA databases (Supporting Information Table 6). Differences in microtubule activity between non-Stem-A and Stem-A subgroups were statistically evaluated with Mann–Whitney U-test in Graphpad Prism. Grey = non-Stem-A subgroup; blue = Stem-A subgroup.

B. Specificity of drug sensitivity in ovarian cancer cell lines. A panel of 18 ovarian cancer cell lines was classified into non-Stem-A (OVCA433, OVCA429, OVCAR-8, PEO1, OVCA432, OVCA420, HeyA8, HEY, HeyC2, SKOV-3, ovary1847 and DOV 13) or Stem-A (PA-1, CH1, A2780, OVCAR-3, SKOV-4 and SKOV-6) groups and analysed for their sensitivity to paclitaxel (Top panel), vincristine (Left bottom panel) and vinorelbine (Right bottom panel). GI50 values were calculated with the results from cell proliferation assays for each cell type in three independent experiments, and the mean GI50s are shown as dot plots. A non-parametric Mann–Whitney U-test in Graphpad Prism was used to evaluate the results statistically. A higher value along the y-axis indicates increased sensitivity to the drugs. Colour as for (A).

C. Detection of apoptotic activity upon vincristine treatment. Six non-Stem-A (Upper panel) and four Stem-A (Lower panel) cell lines were subjected to increasing concentrations of vincristine (0 to 10 nM) for 48 h. The presence of apoptotic activity was determined by immunoblotting for cleaved PARP and Caspase-3, as indicated by arrows. Abbreviations: Stem-A, stem-like-A.

counterparts. Hence, matching breast cancer cell lines by expression data could represent *in vivo* tumours (Gatza et al, 2010; Neve et al, 2006; Perou et al, 2000). Whilst we acknowledge that cell lines may be divergent from their ancestral tumour and not wholly representative of the full diversity of ovarian cancer, we believe our classification represents a foundation for further development, particularly since ovarian cell lines can be assigned to unique ovarian tumour subtypes and are not derived from any random scheme. This concept is supported by the similarities in the expression and pathway activation between the cell lines and tumours of a

given subtype, and could be further supported by shared cell functions, such as anchorage-independent cell growth and population doubling time. The availability of representative cell lines would facilitate the quest for functionally relevant targets and bring us a step forward in developing therapeutics that could be matched with the characteristics of individual patients.

Loss-of-function studies using pooled shRNA libraries have identified essential genes in specific human cancer cell lines in the context of synthetic lethality (Barbie et al, 2009; Luo et al, 2008; Scholl et al, 2009) and lineage-specificity (Cheung et al, 2011). Extending this concept, we utilized the pooled shRNA