Bioinformatic Analysis for the Validation of Novel Biomarkers for Cancer Diagnosis and Drug Sensitivity

Laura Anne Rebecca Lockwood

Email: u1150569@unimail.hud.ac.uk

Abstract

Background: The genetic control of tumour progression presents the opportunity for bioinformatics and gene expression data to be used as a basis for tumour grading. The development of a genetic signature based on microarray data allows for the development of personalised chemotherapeutic regimes.

Method: ONCOMINE was utilised to create a genetic signature for ovarian serous adenocarcinoma and to compare the expression of genes between normal ovarian and cancerous cells. Ingenuity Pathways Analysis was also utilised to develop molecular pathways and observe interactions with exogenous molecules.

Results: The gene signature demonstrated 98.6% predictive capability for the differentiation between borderline ovarian serous neoplasm and ovarian serous adenocarcinoma. The data demonstrated that many genes were related to angiogenesis. Thymidylate synthase, GLUT-3 and HSP90AA1 were related to tanespimycin sensitivity (p=0.005).

Conclusions: Genetic profiling with the gene signature demonstrated potential for clinical use. The use of tanespimycin alongside overexpression of thymidylate synthase, GLUT-3 and HSP90AA1 is a novel consideration for ovarian cancer treatment.

Keywords: Bioinformatics, genetics, microarrays, ovarian, cancer, tanespimycin, signature

Acknowledgements

Acknowledgements should be given to my supervisor, Dr Rachel Airley, for the inspiration behind this research project and Dr John Stephenson, for his support with statistical analysis.

http://dx.doi.org/10.5920/fields.2015.115

Except otherwise noted, this article copyright: © 2014 Laura Anne Rebecca Lockwood. This work is licensed under a Creative Commons Attribution 4.0 International License
Introduction
Key Features of Cancer
Traditional theory states that cancer is a simple disease involving down-regulated control over cell proliferation. However, research now suggests that the molecular pathology of cancer is highly complex and cells acquire a plethora of traits that permit tumourgenesis and malignant transformation (Hanahan, & Weinberg, 2000). Whilst all cancers display a number of differences, there are some common themes that have been identified (Figure 1), with key features involving the evasion of apoptosis and angiogenesis (Hanahan & Weinberg, 2000).


Genetic basis of cancer
The cell cycle is a key aspect to consider in oncogenesis; the stages of the cell cycle are tightly regulated by the expression of a series of proteins, with mechanisms existing to prevent tumourgenesis and promote apoptosis (apoptosis is the term used to describe programmed cell death) in cells which are damaged (Shah, M & Schwartz, G. 2001). Cell proliferation is regulated by genes. Tumour suppressor genes, such as p53 (Levine et al 1991), inhibit the growth and division of cells, whilst proto-oncogenes, such as RAS (Furth et al 1987), stimulate cell proliferation and accelerate growth (Chial, 2008). This equilibrium between tumour suppressor genes and proto-oncogenes, in healthy tissues, maintains a balance of cell growth and apoptosis through cell cycling processes (Chial, 2008). Proto-oncogene mutation can
lead to the formation of oncogenes and ultimately increases the probability of tumourgenesis, especially since those involving the amplification of gene frequency, such as that caused by mutations or polymorphisms in the promoter regions of genes (Biondi et al 2000), can increase the expression of proteins involved in malignant progression (Chow, 2010). This is exemplified by a polymorphism in the promoter region of the PR gene, which increases its transcription and thus production of the hPR-B protein in endometrial cancer (De Vivo et al 2002).

Figure 2. Mechanisms Involved in Apoptosis Cascade and Cancer Initiation (Wong, R (2011) Apoptosis in Cancer: From Pathogenesis to Treatment Journal of Experimental Cancer and Clinical Cancer Research Vol. 30 pp.87. doi:10.1186/1756-9966-30-87. This image is © 2011Rebecca SY Wong, used under a Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited).

Mutations in genes can be either inherited or acquired, with inherited mutations substantially increasing the risk of developing a cancer if a first degree relative has been diagnosed with it; this is commonly seen in patients who have inherited mutations of the MET gene which can cause hereditary papillary renal cancer (Linehan et al 2010). As previously mentioned, acquisition of mutations in the promoter regions of the tumour suppressor genes can alter their expression (Vooght, K. et al 2009) and ultimately down-regulate the control of proliferation and cell death (Chial, 2008). The different types of mutation ultimately will impact on the outcome,
with point mutations, insertions and deletions being common types of mutation that can have very dissimilar effects on the cell (Loewe, 2008).

Apoptosis is initiated by numerous factors (Figure 2). However, the caspase enzymes have demonstrated huge significance in the progression of cell death, by cleaving vital cellular proteins and activating DNAase for the degradation of DNA (Lavrik et al 2005). Genetic expression of caspases has been reported as reduced in many cancers, thus reducing the apoptotic ability of cancerous cells (Philchenkov et al 2004). Cells which have acquired the ability to evade apoptosis can be immortalised and replicate uncontrollably, leading to tumour formation (Lowe & Linn, 1999).

Tumour Growth & Hypoxia
Tumour growth often leads to hypoxia, whereby the oxygen demand outweighs the diffusion of oxygen from the local vasculature (Dachs & Tozer, 2000). Hypoxia has the ability to augment or diminish gene expression to facilitate continued tumour growth in an environment which is oxygen deprived. This is achieved by stimulating angiogenesis, mediated by the increased expression of angiogenic factors, such as hypoxia-inducible factor (HIF). This allows a greater permeation of oxygen to tumour tissue by increasing HIF-1α gene expression (which encodes for a regulatory subunit of HIF) through its subsequent impact on signalling cascades and increase in the density of glucose transporters to enhance glucose uptake for glycolysis (Dachs & Tozer, 2000). Angiogenesis involves the novel formation of a new vasculatory network and is an essential process in the development of cancer as tumours utilise this mechanism to obtain oxygenation and nutrients to meet the needs of the growing neoplasm (Nishida, N. et al 2006). The angiogenesis cascade is essential and involves key molecules in the progression of cancer (Figure 3). Heat-shock protein 90 (HSP90) is described as a molecular chaperone heavily involved in the regulation of HIF and ultimately impacts significantly on angiogenesis (Bohonowycz et al 2010). Apoptosis and angiogenesis in healthy cells are under genetic control, whilst in cancerous cells, mutations in these regulatory genes can lead to out of control cell proliferation and tumour growth through changes in expression levels (Hanahan & Weinberg, 2000).
Figure 3. Angiogenic Signalling in Cancer. (Bohonowycz, J.E. Gopal, U. & Isaacs, J.S (2010) Hsp90 as a Gatekeeper of Tumor Angiogenesis: Clinical Promise and Potential Pitfalls Journal of Oncology Vol. 2010, Article ID 412985, doi:10.1155/2010/412985. This image is © 2010 J. E. Bohonowycz et al., used under a Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited).

Metastasis in cancer has been researched extensively and encompasses a multi-step process (Nakayama et al 2012) involving: the disconnection of adhesions and single-cell separation from tumour masses (Cook et al 2011), tumour intravasation into blood and lymphatic systems (Thiolloy & Rinker-Schaeffer, 2011) and subsequent extravasation (Hata et al 2007), evasion of the host immune system in circulation (Youn et al 2008) and, as mentioned, angiogenesis (Houle et al 2002). Two models exist for cancer metastasis with the first considering 'early dissemination' as a method of metastasis which occurs in precancerous lesions, prior to malignant development, and the second describing 'late dissemination' which involves metastasis from primary tumours (Klein 2008). Once disseminated, cells are carried in the blood and lymphatic systems to sites distant from the primary tumour and begin to grow as a secondary mass (Liotta 2001). These models suggest that early disseminated metastases are genetically diverse from the primary tumour, whilst those that are a product of late dissemination demonstrate a genetically similar profile; early disseminated metastases may therefore behave differently to the primary tumour and could consequently demonstrate a distinct difference in chemoresponse (Klein, 2008).
Chemoresponse

Chemotherapy

Chemotherapy drugs have been in the clinic for in excess of 50 years, with efforts continuing to develop medicines which target neoplastic cells and subsequently display cytostatic or cytotoxic effects (Johnstone, R. et al 2002). Anti-cancer drugs have structural differences and varying specificity. However, all drugs aim to induce cellular changes to precipitate initiation of apoptosis (Johnstone, R. et al 2002). To further understand the effects of anticancer drugs, tumour response and chemosensitivity, their mechanisms must be understood in more detail.

Classical Anticancer Agents

Anticancer drugs can be grouped into three main categories: classical chemotherapy, hormonal and immunotherapy. Chemotherapy can be further subdivided into: alkylating agents, antimetabolites, cytotoxic antibiotics, topoisomerase inhibitors and anti-microtubule agents (Espinosa et al 2003). Classification of anti-cancer drugs according to their basic mechanism is particularly important as it allows for a comprehensive overview of drug availability, considers the potential for toxicity in multi-drug regimens (Espinosa et al 2003) and permits a brief look at the potential for chemoresponse on an individual patient basis.

The capacity of tumour cells to undergo cell division and rapid growth is derived from their ability to undergo DNA replication at a rapid rate; alkylating agents use this ability as their primary target to demonstrate cytotoxic and cytostatic effects (Valeriote & van Putten, 1975). Alkylating agents is a broad category which includes platinum-based drugs and describes a general mechanism whereby all alkylating agents form covalent links with nucleophilic-centered macromolecules, such as DNA, to form ‘cross-links’ and thus inhibit DNA replication and subsequent RNA production and protein synthesis (Siddik, 2002).

Antimetabolites are drugs which intervene in the metabolism within a cell and include dihydrofolate reductase inhibitors and thymidylate synthase inhibitors (Kaye, 1998). Folate is a key component in purine and pyrimidine synthesis, which form the basic units of DNA (Choi Mason, 2000), with dihydrofolate reductase catalysing dihydrofolate reduction to tetrahydrofolate; methotrexate is an example of a dihydrofolate reductase inhibitor (Schweitzer et al 1990). Thymidylate synthase catalyses the production of thymidine phosphate from dUMP; inhibitors of this enzyme, such as 5-fluorouracil, can therefore inhibit the production of thymidine phosphate which is essential for DNA synthesis and repair (Touroutoglou & Pazdur, 1995).

Cytotoxic antibiotics, or ‘antitumour antibiotics’, are derived naturally from microorganisms and generally exert their action directly on DNA, although the precise mechanism is drug dependent (Missailidis, 2008). Anthracyclines, such as daunorubicin and doxorubicin, are thought to exert their action through a combination of DNA intercalation (and subsequent inhibition of DNA synthesis) and free radical formation leading to damage of existing DNA (Minotti et al 2004). Dactinomycin (actinomycin D) is a chromomycin and an additional example of a cytotoxic antibiotic which binds directly to DNA, inhibiting RNA synthesis through interference with RNA elongation (Sobell, 1985). Bleomycin, a further example of an
antitumour antibiotic, binds to and degrades DNA, although it requires oxygen for activation (Stubbe & Kozarich, 1987). Mitomycin C is another example of a naturally occurring chemotherapeutic agent; it forms covalent cross-linkages between DNA strands that are complementary and subsequently inhibits replication (Tomasz et al 1987).

Topoisomerase inhibitors are another class of chemotherapeutic agents and are divided into two groups, those that inhibit topoisomerase I enzyme (which is involved in the initiation of the cleavage of one strand of DNA) and those that inhibit topoisomerase II enzyme (which initiates cleavage of both strands) (Hande, 2006). Through inhibition of the topoisomerase enzyme, these chemotherapeutic drugs can inhibit DNA replication and subsequently promote cell apoptosis (Eweseudo & Ratain, 1997). Topoisomerase I inhibitors, such as irinotecan, are derivatives of the phytochemical camptothecin (Eweseudo & Ratain, 1997), whilst topoisomerase II inhibitors include etoposide, which can prevent re-ligation of cleaved DNA (Hande, 2006).

Microtubules are an essential part of the eukaryotic cell; they are required for intracellular transport, cellular movement and division of the cell and are composed of tubulin heterodimers (Checci et al 2003). Neoplastic cells undergo mitosis more rapidly than normal cells, thus providing a target for drugs which interact with microtubules and block progression of the cell cycle, allowing for the initiation of apoptotic cascades (Checci et al 2003). Anti-microtubule chemotherapeutic agents include the vinca alkaloids and the taxanes (Checci et al 2003), with the taxanes, such as paclitaxel, inhibiting cell proliferation through stabilisation of microtubules and subsequent mitotic inhibition (Rowinsky 1997). The vinca alkaloids, such as vincristine and vinblastine, initiate depolymerisation of microtubules through the binding of β-tubulin and similarly, can inhibit continuation of mitosis (Checci et al 2003).

Targeted Anticancer Agents
Targeted anticancer agents and ‘targeted therapy’ are terms used to describe anti-neoplastic drugs which are designed by intention to interfere with a specific target molecule which has been identified as possessing a key role in tumour development and progression (Sawyers, 2004). The limitations of the classical anticancer agents, including poor specificity and subsequent toxicity, high dose requirement and multi-drug resistance, can be overcome through the use of targeted anticancer agents and has led to the development of relatively novel compounds based on known targets (Chari, 2007).

Anti-oestrogens are an additional class of targeted-anticancer therapy and include the pro-drug tamoxifen (Colleta et al 1994). They antagonise the effect of oestrogens through competitive inhibition of the oestrogen receptor (Neven & Vergote, 2001), thus demonstrating usefulness in the pathogenesis of several cancers, such as breast, which have been linked to oestrogen exposure (Yager & Liehr, 1996).

Monoclonal antibodies, such as rituximab and trastuzumab, have also been developed as targeted anticancer agents as they are developed to possess a high specificity in the targeting of tumour antigens (Chari, 2007). To enhance their anti-tumour activity, these antibodies can also be conjugated with cytotoxic drugs or with
radioactive isotopes (Chari, 2007), as seen with doxorubicin and monoclonal antibody IgG2a (Yang & Reisfeld, 1988). The specificity of targeted anticancer agents highlights the importance of identifying targets for individual cancers and promotes development of novel agents based on these targets.

**Factors Affecting Chemoresponse**

Response to chemotherapy is measured in terms of patient survival, tumour size reduction and changes in metabolic activity (Weber, 2005). Chemosensitivity and chemoresistance are terms used to determine the level of response of a particular cancer to specific chemotherapy regimens however, cancer requires a more individualised marker for defining the measurement of response (Toole et al. 2007). Chemoresponse can be analysed by tumour microenvironment and tumour cell characteristics which includes gene expression levels, presence of membrane transporters and genetic polymorphisms (Longley & Johnston, 2005). Gene expression signatures are a relatively novel method in the field of oncology and serve the purpose of identifying novel drug targets, provide prognostic information and predict chemoresponse in individual patients (Rathnagiriswaran et al. 2010). This allows for a personalised approach to the design of chemotherapeutic regimens to improve patient response and clinical outcome.

**Chemosensitivity and Chemoresistance**

Drug resistance of cancer cells has been identified as the primary factor in failure of chemotherapy regimens to treat tumours (Gatti & Zunino, 2005); and thus forms the basis of an argument for the design of personalised chemotherapy regimens. The failure of chemotherapeutic treatment can arise from either intrinsic or acquired drug resistance (Ozben, 2006), with intrinsic resistance existing prior to treatment and acquired resistance occurring during or in response to treatment (Wilson et al. 2006). Chemoresistance in tumours can be exhibited for one class of drug or can extend to many different classes, termed ‘multi-drug resistance’ (Baguley, 2010). Drug resistance is particularly complex, since there are numerous factors which can contribute to the ability of a tumour to evade treatment (Wilson et al. 2006); alterations to prevent drug accumulation inside the cell (Gottesman, 2002), changes to drug targets (Hayes & Wolf, 1990) and factors which impact on cellular response, such as DNA mismatch-repair (Baguley, 2010) are all mechanisms of chemoresistance. Tumour hypoxia can also impact on drug delivery to cells (Baguley, 2010), with the tumour vasculatory system requiring the hypoxic ‘switch’ for angiogenesis (Laderoute et al. 2000), and the resulting supply of blood to the tumour impacting on drug access, since blood flow can often be intermittent (Baguley, 2010).

Tumour cells utilise drug transporters as a method of drug accumulation evasion, with overexpression of ABC transporters having been demonstrated to contribute to multidrug resistance, and their efficacy is owed to their ability to increase drug efflux, thereby reducing intracellular concentration of cytotoxic agents (Ozben, 2006). Drug families most associated with this type of multidrug resistance include the taxanes, vinca alkaloids and antimetabolites (Ozben, 2006). P-glycoprotein is an example of an ABC multi-drug transporter which is well established as a mediator of resistance to a number of drug classes (Stavrovskaya, 2000). The effect of gene expression on drug resistance is exemplified by the multidrug resistance gene, which encodes for the P-glycoprotein transporter, with overexpression of the mdr1 isoform of the gene
correlating to multidrug resistance in specific solid tumour masses (Nooter & Herweijer, 1991).

Alteration of a drug target can occur through loss of a cell surface receptor or simple mutation of the receptor or target gene (Gottesman, 2002). Many drugs have a specific target protein, such as an enzyme, upon which their mechanism of cell destruction depends; the efficacy of these drugs can be vastly reduced when the levels of expression of these targets are elevated or decreased (Stavrovskaya, 2000). Neoplastic cells which display resistance to topoisomerase inhibitors, for example, tend to have a lower quantity of topoisomerase enzyme which permits evasion of cytotoxicity (Stavrovskaya, 2000). Furthermore, once cells that are susceptible to chemotherapy have been damaged by cytotoxic agents, mutations in the genes or proteins (e.g. tumour suppressor gene p53) involved in the apoptotic cascade can also aid in a cell’s ability to resist programmed cell death (Stavrovskaya, 2000).

Neoplastic cells have the additional ability to develop resistance to drugs which target specific stages in the cell cycle in the attempt to induce apoptosis; cells which are not undergoing the targeted cell cycle stage, such as those at the hypoxic centre of a dense solid tumour, may therefore be insensitive to the action of the chemotherapeutic agent (Shah & Schwartz, 2001). The use of combination chemotherapy in particular can propagate this type of resistance, with one agent impacting on the cell cycle to such an extent as to render another agent ineffective, such as that seen with flavopiridol and paclitaxel (Shah & Schwartz, 2001).

Current Chemotherapy Guidance
Clinical treatment of cancer is predetermined by guidance given to the National Health Service and outlines where surgery and drug treatment is appropriate, often without considering many individual characteristics of patients and their disease. Characteristics vary within and between cancer types, however, the resource provision for its treatment originates from the National Institute for Health and Care Excellence (NICE), with little consideration for the influence of a patient’s genetics on prognosis and chemoresponse. However, NICE have recently developed guidelines considering genetic predisposition to breast cancer, utilising the presence of BRCA1, BRCA2 and/or TP53 mutations as markers of risk for breast cancers, allowing for guidance on chemoprevention (NICE, 2013). Additionally, genetic testing has been used to identify HER2 positive and negative patients in metastatic gastric (NICE, 2010) and breast cancer. This forms the basis for the choice of specific chemotherapy regimes, particularly where the monoclonal antibody trastuzumab is concerned, since this targets HER2 when overexpressed and should not otherwise be used in patients who prove negative, due to risk of cardiotoxicity (NICE, 2006). The involvement of genetic testing in the clinic has enormous potential for development; personalised medicines would allow for the tailoring of a drug regimen to suit an individual patient, regardless of guidelines, and should combine tumour characteristics, genetic influence and drug sensitivity to optimise chemotherapy outcomes (Jain, 2005).

Biomarkers
Since the discovery that the basic origins of cancer involve genetic alterations in the cell (Bunz, 2008) and that genomic biomarkers can be utilised to indicate disease
state and progression (Strimbu & Tavel, 2010), oncogenomic data has been used to accelerate the scientific understanding behind cancer to rationally develop better treatment choices and regimes (Strausberg et al 2004). Biomarkers have been identified as predictors of tumour chemosensitivity (Lee et al 2007) and can therefore be a key aspect in the development of personalised medicines for cancer patients (La Thangue & Kerr, 2011). Biomarkers fall into two primary categories: predictive and prognostic, with predictive providing information on the cancer which is independent of drug therapy and prognostic providing information on the likelihood of response to targeted therapies (Alymani et al 2010).

There are some notable examples of when biomarkers have successfully predicted therapeutic outcome and have been extrapolated into the clinical setting: the human epidermal growth factor receptor 2 (HER2) now provides a target for trastuzumab; it was discovered to be over-expressed in a quarter of patients with breast cancer who now benefit from more personalised treatment (Alymani et al 2010). Additionally, K-ras is a small G-protein and is significantly involved in signal transduction of the epidermal growth factor receptor (EGFR) and it has been demonstrated that anti-EGFR drugs (Saltz et al 2006), such as panitumumab and cetuximab are ineffective, when K-ras mutations isolate the pathway from providing a target (Alymani et al 2010). Identification of the K-ras mutation can therefore be utilised to identify patients who will not benefit from EGFR inhibitors and may require alternative treatment, as seen in NICE guidance for the treatment of metastatic colorectal cancer (NICE, 2013).

Patients who have cancer which demonstrates a high level of resistance to chemotherapy when given in accordance with guidelines may benefit from gene profiling to identify which therapies the cancer demonstrates chemosensitivity. Research exists which encompasses the analysis of RNA expression data in breast cancer generated from array technology and demonstrates a high level of accuracy for predicting chemosensitivity to docetaxel (Chang et al 2003). This highlights the potential for extrapolation of this technique into the clinical setting for the gene profiling of all patients to identify optimum therapy. As a consequence of this understanding, oncogenomics can use bioinformatics and microarray technology as a platform for processing of data and for its analysis (Rhodes, 2004).

Bioinformatics & Microarray Technology

Bioinformatics is a recently established, multifaceted field of science encompassing molecular biology, information technology and statistical analysis (Yi, 2013). It provides scientists with the tools to mine vast databanks for information and integrate it into programmes of research (Barts Cancer Institute, 2013). Since the introduction of microarray technology in 1995 (Schena, et al 1995), it has become a major factor in genomic research (Manning et al 2006) and science has seen a notable increase in the number of published papers considering predictive biomarkers (Fig. 4) (Alymani et al 2010). Microarrays have the capability to measure the expression of over a thousand genes simultaneously (Manning, A.T et al 2006) which allows for the application of this technology, within the field of bioinformatics, to research into personalised medicines for cancer patients. The data generated from microarrays has been successfully implemented in the identification of biomarkers for the targeting of drugs (Kozian & Kirschbaum, 1999) and has recently been analysed for the generation of gene expression signatures in the classification
of breast cancer subtypes (West et al 2001). This has allowed for a more accurate prediction of prognosis (van't Veer et al 2002). Online databases now exist for microarray data and are available for users to conduct their own research through data manipulation and analysis.


**ONCOMINE**

ONCOMINE is a web-based database of cancer microarray technology which allows for gene expression profiling (Rhodes2004) and has permitted the development and success of numerous research activities within the field of oncology. Recent cancer research involving ONCOMINE has found success when observing gene expression comparatively between localised prostate tumour cells and metastatic tumour cells to further understand tumour progression (Gorlov et al 2010). The study was able to determine gene candidates *in silico* which serve as signatures for tumour progression and development (Gorlov et al 2010). ONCOMINE was also central to a meta-analysis of breast cancer data, with subsequent data mining identifying the NRF-1 gene as up-regulated in oestrogen receptor positive patients when compared to oestrogen receptor negative patients (Kunkle et al 2009). These differences were found to have implications on prognosis, affecting survival rates and chances of disease relapse (Kunkle et al 2009).
Ingenuity Pathways Analysis
Ingenuity Pathways Analysis (IPA) is an entirely web-based software tool which facilitates the identification of signalling cascades and pathways involved in metabolism, whilst allowing for the prediction of the effects of these pathways on disease manifestation and progression (USC, 2011). Effective use of ONCOMINE can be combined with use of IPA to generate further understanding of how gene expression correlates to cell processes in their entirety, allowing for detailed analyses of the pathways that genetic expression has the potential to impact on; this combination technique was employed in the isolation of ‘potential blood-based markers’ for a number of common human cancers and provided data which could support further research into biomarker validity (Yang et al 2008).

Ovarian Cancer
Ovarian cancer is the predominant cause of mortality in gynaecological malignancies, with the five-year survival rate remaining at just 53% (Ries, LAG. et al 2003), which can be attributed to the majority of cases being diagnosed at an advanced stage (Landen et al 2008). Current NICE guidance for treatment of ovarian cancer recommends either a platinum-based compound (carboplatin or cisplatin) or a combination of cisplatin with paclitaxel as first-line therapy (NICE, 2003). However, current research shows that some ovarian cancers are displaying resistance to cisplatin-induced cytotoxicity and reduced tumour cell sensitivity (Hu et al 2003), which demonstrates a possibility for the NICE recommended therapy to be largely ineffective.

Pathogenesis
Tumourgenesis occurs in ovarian cancer through malignant transformation of cells which can originate from the outer epithelial surface of the ovary, or from the superficial epithelial layer of the fallopian tubes (Dubeau, 2008). Ovarian tumours have been partitioned into two categories as part of a model to facilitate identification; type I tumours include serous, endometrioid, mucinous and clear-cell carcinomas and are often low grade, whilst type II tumours tend to be more aggressive and include high-grade serous carcinoma and malignant mesodermal tumours. The majority of ovarian tumours are type II high-grade serous carcinomas which spread into the abdomen and undergo distant metastasis (Kurman & Shih, 2010).
Ovarian cancer is a heterogeneous disease and each of the cancers within the classification type outlined in the model have demonstrated genetic similarity, with type I cancers rarely displaying mutations in the p53 gene and commonly displaying mutations in KRAS, BRAF and ERBB2 genes. In contrast, type II tumours have a high frequency of p53 gene mutations and are significantly more homogenous (Kurman & Shih, 2010). The introduction of protein expression analysis into ovarian cancer research has presented opportunities for the identification of signatures for the different histotypes which fall under ovarian cancer, and whilst the pathogenesis of the various histotypes can be deduced, knowledge is sought to determine how these genes impact on tumour progression and chemosensitivity (Toss et al 2013) with recently proposed signalling pathways being a key starting point (Fig. 5).

Research suggests that long term prognosis for this cancer is significantly related to the degree of microvascular development within a tumour. Growth and capacity of ovarian carcinoma to metastasise has been demonstrated to be highly angiogenesis dependent (Alvarez et al 1999), with tumour hypoxia resulting in the secretion of proangiogenic growth factors to stimulate the development of tumour vasculature (Ramakrishnan et al 2005). Angiogenesis is integral to tumour survival, particularly in ovarian cancer, since tumour size is often large in relation to the relatively small ovaries and revascularisation is key to disease progression (Hazelton...
et al 1999). In addition to growth enhancement, angiogenesis can also facilitate the circulation of tumour cells to promote metastasis (Álvarez et al 1999).


Biomarkers and Signalling
As previously identified, the genetic mechanisms behind ovarian cancer will determine its histotype and behaviour. There are numerous genes implicated in tumour pathogenesis (Fig. 6), and a variety of genes have been found to be implicated in the different grades of ovarian cancer (Fig. 7). In addition to the genes identified, there are a number of genes which have been researched more thoroughly and are widely accepted as biomarkers that are strongly affiliated with ovarian cancer pathogenesis and progression.
As mentioned, the key to tumour growth in ovarian cancer is the hypoxic stimulus of growth factors. Hypoxia inducible factor (HIF) is a complex which is heavily involved in transcription and which is highly sensitive to changes to intracellular oxygen levels (Poon, E. Harris, A. & Ashcroft, M. 2009). The overexpression of HIF in ovarian cancer is correlated with a poorer prognosis and survival rate (Semenza, 2002). HIF-1α is a ‘regulatory subunit’ of HIF and its overexpression is highly associated with advanced neoplasms (Poon, E. et al. 2009). Hypoxia often stimulates changes in cellular metabolism, with the glucose transporter GLUT-1 having been identified for its expression levels providing a relationship with the stage of cancer; significant overexpression has been noted in malignant epithelial tumours when compared to borderline disease (Canturia et al 2000) which is expected, since GLUT-1 expression is known to be controlled by HIF-1 (Behrooz & Ismail-Beigi, 1997).

The proangiogenic growth factor, vascular endothelial growth factor (VEGF), was shown to be significantly implicated in angiogenesis in both normal ovaries (to maintain the function of the menstrual cycle) and in neoplastic ovaries (Ramakrishnan et al 2005). VEGF binds to tyrosine kinase receptors, activating the PI3 and AKT/MAP Kinase pathways and was identified as over-expressed in ovarian tumour cells comparative to normal, providing adequate vascularisation for neoplastic survival and promoting cell immortality. VEGF expression appears also to provide prognostic information on disease staging (Ramakrishnan et al 2005). Research in non-human cell lines has also found that overexpression of VEGF in ovarian epithelial surface cells can undergo malignant transformation to cells which form ascites (Ramakrishnan et al 2005), thus highlighting the potency of VEGF as a prognostic biomarker for ovarian cancer. Gene polymorphisms of VEGF were also highlighted as potential factors in prognosis, with research leading to the consideration that differing combinations of genotype might affect circulating levels of VEGF (Hefler et al 2007).

Additionally, there are numerous growth factors which can have an impact on tumour cell survival, namely placental growth factor (PGF) (Hu et al. 2003). PGF has a role in the reorganisation and normalisation of tumour vessels (Hedlund et al 2012). Cellular enzymes have also been the subject of research in ovarian cancer: data revealed that the expression of the gene encoding the enzyme thymidylate synthase is significantly higher in epithelial ovarian cancer compared to normal ovaries (Fujiwaki et al 2000), which suggests a heavy involvement of thymidylate synthase in the pathogenesis and progression of ovarian cancer. Another enzyme, topoisomerase II, was subject to investigation, with the genes topoisomerase II alpha and
topoisomerase II beta being upregulated in ovarian cancer; topoisomerase II beta was considered to have potential as a novel chemotherapeutic target (Withoff et al 1999). Furthermore, the enzyme caspase 1-α has been discovered as pro-apoptotic in ovarian cancer cells and its down-regulation in these cells was suggested as a factor in the resistance of ovarian cancer cells to apoptosis (Feng et al 2005).

**Chemoresponse**

VEGF overexpression has additionally been found to mediate cytoprotection against cisplatin-induced cytotoxicity and reduce tumour cell sensitivity (Hu et al 2003). The demonstration of hypersensitivity to anti-VEGF drugs in placental growth factor expressing tumours is supported by research which shows that the inactivation of PGF in human tumours leads to a greater resistance to anti-VEGF drugs (Hedlund et al 2012). Proposed theories suggest that PGF-normalised tumour blood vessels could potentially augment delivery of anti-VEGF drugs to the microenvironment of the tumour (Hedlund et al 2012). PGF has also been shown in vitro to sensitise tumour cells to the anti-angiogenic effects of anti-VEGF drugs (Hedlund et al 2012).

High grade ovarian neoplasms are also often known to overexpress indoleamine 2,3-dioxygenase (Okamoto et al 2005), encoded for by the IDO-1 gene (Soliman et al 2010), which has been linked with paclitaxel resistance and overall impaired survival in serous tumours (Okamoto et al 2005). Furthermore, βIII tubulin overexpression has also been associated with paclitaxel resistance in ovarian cancer (Kamath et al 2005).

**Method**

Owing to the increased clinical need for a chemotherapeutic regimen which is effective in treating ovarian cancer, the generation of a novel genomic signature for at least one histotype (ovarian serous adenocarcinoma) was undertaken in order to determine how these genes impact on disease staging and to facilitate the selection of a drug to which the histotype is chemosensitive.

Highly ranked gene expression values were generated from the online database, ONCOMINE, for ovarian cancers. Ten genes were chosen for further analysis based on preliminary research and expression values were generated for both borderline ovarian serous cancer and ovarian serous adenocarcinoma to facilitate comparison. Data validation utilising another dataset was also undertaken to evaluate the predictive capability of the gene signature model. Additionally, data was then viewed for the expression values of the genes in chemoresistant and chemosensitive cell lines for relevant chemotherapeutic agents to demonstrate how the expression of the genes within the signature might impact on chemoresponse. Further analysis of the genomic signature was undertaken utilising IPA with the aim of identifying networks and pathways between genes and in order to further understand how the genomic signature impacts on the disease and chemosensitivity.
Results & Data Analysis

Initially a database search of ONCOMINE was performed which observed gene expression in ovarian cancer cells comparative to normal ovarian cells. Ten genes were selected based on preliminary research into the existence of previously identified relationships of the genes to ovarian cancer pathogenesis or progression. Fig. 8 displays a heat map for the expression of the chosen genes in the two conditions: normal ovary cells and ovarian serous adenocarcinoma.

Gene expression values were then obtained for the chosen genes in an additional dataset of borderline ovarian serous neoplasm when compared to ovarian serous adenocarcinoma (Fig. 9). With the raw data obtained, a logistic regression was completed which considers the binary outcome of cancer state (comparing the two stages), utilising a sample size of 74 patients. The intent of the model is to use the gene expression data as a predictor of cancer state. ONCOMINE uses log2 median-centred intensity expression values.
The classification table (Table 1) shows that 98% of patients were correctly classified into their respective cancer state, which demonstrates that the model has a good predictive capability and is adaptive for use in patients who do not form part of the dataset.

<table>
<thead>
<tr>
<th>Observed Outcome</th>
<th>Predicted Outcome</th>
<th>Percentage Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borderline</td>
<td>30</td>
<td>100.0</td>
</tr>
<tr>
<td>Severe</td>
<td>1</td>
<td>97.7</td>
</tr>
<tr>
<td>Overall Percentage</td>
<td></td>
<td>98.6</td>
</tr>
</tbody>
</table>

Table 1. Classification of Cancer State Based on Gene Signature

The parameter coefficients in a multiple logistic regression equation (Table 3.2) may be interpreted as the change in the log odds of the cancer being the latter grade for a one-unit change in the expression of each gene. An economical logistic regression was also undertaken which found the genes ARNT, HSP90AA1 and TOP2A to be significantly associated with the outcome of the latter grading and a model which utilised these three genes had 95.9% predictive capability. Preliminary validation of the original model was undertaken utilising a novel dataset which demonstrated a 35% predictive capability. An additional cross-validation was conducted using only the three genes found to be significantly associated with the outcome under the economical multiple regression model and resulted in a
25% predictive capability.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Parameter Coefficient</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOP2A</td>
<td>3.113</td>
<td>22.495</td>
</tr>
<tr>
<td>ARNT</td>
<td>7.644</td>
<td>2088.549</td>
</tr>
<tr>
<td>VEGFA</td>
<td>-.231</td>
<td>.793</td>
</tr>
<tr>
<td>TYMS</td>
<td>-.639</td>
<td>.528</td>
</tr>
<tr>
<td>PGF</td>
<td>2.707</td>
<td>14.986</td>
</tr>
<tr>
<td>CASP10</td>
<td>1.662</td>
<td>5.271</td>
</tr>
<tr>
<td>MTOR</td>
<td>.447</td>
<td>1.563</td>
</tr>
<tr>
<td>HSP90AA1</td>
<td>3.880</td>
<td>48.405</td>
</tr>
<tr>
<td>SLC2A12</td>
<td>.386</td>
<td>1.471</td>
</tr>
<tr>
<td>SLC2A3</td>
<td>.706</td>
<td>2.026</td>
</tr>
<tr>
<td>Constant</td>
<td>-32.549</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 2. Parameter Coefficients & Odds Ratios: Multiple Logistic Regression Model

Observing the results to be sensitive to the values of ARNT in the validation sample, a further cross-validation was conducted excluding the ARNT gene from the ten gene signature. This resulted in a 75% predictive capability which may be compared with the 98.6% predictive capability of the original model; it is to be expected that the predictive capability would be lower in the validation set since the model is tailored to the original dataset. The performance of the model with the validation sample in the cross-validation of all genes appeared to be due to the underexpression of the ARNT gene in the validation sample data (which was overexpressed in the original data set) and led to a fourth sensitivity analysis which manipulated the gene expression values of ARNT to demonstrate overexpression and thus increased the predictive capability of the model from 25% to a more respectable 70%.

Gene expression values were then obtained in chemoresistant cell lines and these were compared to the chemosensitive cell lines for the drug tanespimycin (Fig. 10), as initial observation of the heat maps illustrated a notable difference in gene expression for the two different categories of chemoresponse (resistant and sensitive). The statistical significance of the individual genes when considering the use of multiple genes must be considered, as the usual ‘cut-off’ forms 5%, however with multiple genes, the value must be stricter. Since ten genes have been selected, the new point below which indicates statistical significance is 0.5% (p≤0.005) which subsequently indicates that TYMS and SLC2A3 are the two genes which demonstrate statistical significance for overexpression and are likely to impact the outcome of chemosensitivity to tanespimycin. HSP90AA1 is also borderline for its impact on chemosensitivity. All other genes failed to demonstrate statistical significance (p≤0.005) for expression differences and ultimately will have no significant effect on the outcome of chemoresponse.
Following the analysis utilising data generated from ONCOMINE, Ingenuity Pathways Analysis (Ingenuity® Systems, 2013) was utilised as a tool to search for and demonstrate where existing research has identified relationships between the molecules that make up the signature and their interactions with other molecules relating to ovarian cancer (Fig. 11). Since previous research indicated angiogenesis was a key aspect to tumour progression, the relevant genes to angiogenesis were then distinguished (Fig. 12) and the specific relationships to the signature were then isolated and highlighted (Fig. 13). Downstream effects of activating cisplatin (Fig. 14) and subsequently tanespimycin (Fig. 15) independently were also assessed.

Discussion

Outcome of Disease

Data analysis revealed that the genes selected to form the genetic signature, based on preliminary research, demonstrated varying degrees of overexpression in ovarian serous adenocarcinoma when compared to normal ovarian tissue. Further data analysis also revealed a difference in expression values in the two different stages of cancer: borderline ovarian serous neoplasm and ovarian carcinoma. The classification model, having a 98.6% predictive capability, indicated that the gene signature was highly representative of changing cancer state which could demonstrate the importance of these genes in the progression of ovarian cancer. When the three genes significantly associated with the outcome of disease controlling the expression of hypoxia inducible factor 1 beta (HIF-1β – alias ARNT), heat-shock protein 90 (HSP90AA1) and topoisomerase 2-α (TOP2α)) were utilised in an economical logistic regression, the predictive capability remained high at 95.9%, demonstrating the potential for the gene profiling of just three genes to predict disease staging.

When the data relating to the genes is considered, it can be seen that HIF-1β, HSP90AA1 and placental growth factor (PGF) were over-expressed in ovarian cancer compared to normal ovarian cell lines and all had a large impact on the outcome of disease state (although PGF was not significantly associated). Research suggests that HIF-1α gene expression is upregulated in cancer, as are glucose
transporters (Dachs, G. & Tozer, G. 2000). However HIF1-β was overexpressed i.e. the beta subunit of hypoxia inducible factor 1. The overexpression of HIF1-β is likely to be a consequence of the requirement to participate in mechanistic dimerisation with HIF-1α for receptor function (Semenza, et al 1997). The role of HIF, previously identified as an angiogenic stimulant, is likely to explain the overexpression when the dependency of ovarian cancer on the angiogenesis cascade is considered. It was also noted that there was a large amount of variation between the two datasets in the expression of HIF-1β; which, when just the three significantly associated genes were utilised in a cross-validation of the economical multiple regression model, negatively impacted on its predictive capability, reducing it to 25%. This could be attributed to sample variation, as the origin of the samples is unknown. However, when the expression values in the validation sample were manipulated in a further cross-validation to display overexpression, the predictive capability increased to 70% which is more comparable to the model with the original dataset. It was expected that the predictive capability would be lower, since the model was tailored to match that of the original dataset and there will have been natural inter-sample characteristic variation.

The anticipated upregulation of glucose transporters in cancer was reflected in the data, with both GLUT-3 and GLUT-12 overexpressed in ovarian serous adenocarcinoma when compared to normal ovarian cells. Since glucose transporters are essential for enhancing glucose uptake for glycolysis in hypoxic conditions, it would be expected that GLUT-3 and GLUT-12 gene expression would increase with cancer progression, leading to a greater impact on disease outcome. However, GLUT-3 and GLUT-12 did not have a significant impact on the outcome of disease state when borderline compared to established cancer was considered. Additionally, the data corresponds to results of published research since higher expression of hypoxia-inducible factor would be expected to increase the expression of glucose transporters (Dachs & Tozer, 2000).

Angiogenesis has been identified as a key process in ovarian cancer progression and research supports the important role played by vascular endothelial growth factor alpha (VEGFα) in angiogenesis. Whilst the data correlated with this research and shows its overexpression in cancer when compared to normal cells, the VEGF expression levels had little impact on disease outcome. As was the case with GLUT-3 and GLUT-12, the overexpression of VEGF could be a consequence of HIF overexpression. It would also have been expected that VEGFα would have presented a greater influence on disease outcome, since revascularisation has been identified as essential to disease initiation and progress. HSP90, involved in the regulation of HIF and VEGF in the angiogenesis cascade (Fig. 1.3), also had a significant impact on disease outcome. This was as expected, since HSP90 functions to stabilise growth factors such as VEGF and has also been demonstrated to stabilise p53 mutations in cancer (Asher et al 2001). As previously identified, the pathogenesis of ovarian cancer may progress along one of two pathways: low-grade pathway and high-grade pathway (Fig. 1.6), with the high-grade pathway commonly including tumours that have a characteristic p53 mutation and tend to be much more aggressive. Additionally, the category of type II tumours includes high grade serous carcinomas. It may be of consideration that the cell lines utilised in the original dataset and validation sample are likely to fall into the high-grade pathway and that of type II tumours, thus supporting the suggestion that there may have been an
inherent p53 mutation, with HSP90 serving a protective function over the unstable protein.

Ingenuity pathways analysis was also utilised to generate common relationships to the chosen genes which form the genetic signature. When an ‘angiogenesis overlay’ was applied to the molecules (Fig. 3.5), the IPA database supported research in showing that vascular endothelial growth factor stimulates and increases angiogenesis through activation of the vascular endothelial growth factor receptor-2 (VEGF-2) (Carmeliet et al. 2001). IPA also demonstrated the relationship between VEGF and PGF, highlighting research which determines that PGF is involved in angiogenesis through its heterodimerisation with VEGF (Cao et al. 1996). Whilst the mammalian target of rapamycin (MTOR) was overexpressed in cancer when compared to normal cells, it did not appear to have a significant outcome on disease state. Data on IPA demonstrated a link between MTOR and angiogenesis, identifying that the MTOR protein is necessary for angiogenesis in human breast tissue (Wen et al. 2012). Similarly, the same effect may occur in ovarian tissue and overexpression of this protein may be attributable to its role in angiogenesis.

Topoisomerase II alpha (TOP2A) also had a large impact on disease outcome and its overexpression in cancer cells when compared to healthy cells corresponded with existing research (Trinh et al. 2013). The role of TOP2A as a key agent in DNA metabolism (Depowski et al. 2000) would explain the increase in expression in cancer cells. Thymidylate synthase, however, also had minimal impact on disease outcome but its overexpression is likely to be a marker of poor prognosis, as seen in other forms of cancer (Rahman et al. 2004). It was also noted that caspase 10 was over-expressed in ovarian cancer cells when compared to normal ovarian cells, which contrasts with published research that shows caspases to be under-expressed in cancer (Philchenkov et al. 2004). It would be expected that an increase in caspase expression would prove proapoptotic in these cancer cells.

Furthermore, key relationships were identified between the signature and other common molecules on IPA (Fig. 3.4). VEGF-α, PGF and MTOR were all identified as being involved in ovarian cancer signalling, as was the p53 tumour suppressor gene. IPA also revealed that HSP90AA1 has been discovered as binding to p53 (Yu et al. 2002). Additionally, HIF-1α, HIF-1β and PGF were further identified as involved in VEGF signalling pathways which also supports the research already covered and the expression patterns noted within the data analysis.

**Chemosensitivity**

Although TOP2A was not significantly over-expressed in Tspimycin sensitive cell lines, its overexpression in cancer when compared to normal cells reaffirms the current use of topoisomerase inhibitors as cytotoxic agents. However, research has shown that in tumour cells previously treated with platinum compounds, TOP2A expression levels decreased which could be a key factor in the development of chemoresistance to topoisomerase inhibitors (Chekerov et al. 2006).

Thymidylate synthase was the most statistically significant gene of the signature for sensitivity to tanespimycin, and was ranked third of all genes analysed in the original study. The role of thymidylate synthase as an enzyme involved in DNA synthesis and repair explains its overexpression in tumour cells, however its role and that of GLUT-3 in tanespimycin sensitivity is unknown and further experimental research
may yield more detailed information. HSP90AA1 demonstrated borderline significance for sensitivity to tanespimycin, however it was expected that it would have been ranked higher due to tanespimycin mechanistically acting as a HSP90 inhibitor.

Using the molecule activity predictor on IPA allowed for the prediction of the downstream effects of cisplatin (Fig. 3.7). The results demonstrated that cisplatin inhibits HSP90AA1 (Donnelly & Blagg, 2008) and also inhibits hypoxia-inducible factor-1 which subsequently leads to inhibition of VEGF-α (McMahon et al 2006). The same method predicted that activation by tanespimycin (Fig. 3.8) would increase the expression of VEGF and would increase the inhibition of HSP90AA1. Whilst the predicted increase in VEGF by tanespimycin would be expected to promote tumour vascularisation, it would also be expected to enhance drug delivery to tumour cells through a greater vascular network (and thus blood supply) to the cells.

The discovery that overexpression of vascular endothelial growth factor may reduce the cytotoxicity of cisplatin has highlighted the importance of developing new chemotherapeutic regimes for patients with this characteristic and consequently, combination therapy with paclitaxel and tanespimycin has recently been developed (Katragadda et al 2013). Further research has shown that tanespimycin inhibits the Akt pathway through its inhibition of HSP90 which sensitises tumour cells to the effects of paclitaxel. The synergistic relationship between tanespimycin and paclitaxel could be exploited to maximise the therapeutic outcome of this regime. This demonstrates the effectiveness of tanespimycin in the sensitisation of neoplastic cells to a proapoptotic stimulus (Solit et al 2003). Data from experimental research revealed that HSP90α, HSP90β, HSP70, HSP72 and HSP27 are the proteins in ovarian adenocarcinoma cells which are responsive to treatment with tanespimycin and suggested, contrary to findings, that ovarian cancer cells which have upregulated these proteins are more likely to develop greater resistance to tanespimycin therapy. Furthermore, the expression of HSP27 was identified as increased by tanespimycin and lower expression of this protein was associated with tanespimycin sensitivity (Maloney et al 2007). The inconsistencies in these findings provide further research opportunities.

Conclusion
The high predictive capability of the model (p≤0.05), despite being lower in the validation dataset, demonstrated potential for the implementation of genetic profiling of ovarian cancer patients to predict the grade of their cancer and thus form the basis for selecting a personalised chemotherapeutic regime. The gene expression data in the comparative grades of ovarian cancer combined with the information generated from ingenuity pathways analysis demonstrated that many of the genes in the gene signature were likely to be involved in angiogenesis which would underpin their role in disease progression and significant overexpression in the higher grade of ovarian cancer.

Additionally, the impact of overexpression of thymidylate synthase, glucose transporter 3 and heat-shock protein 90-AA1 on tanespimycin sensitivity could indicate that patients who exhibit the characteristic overexpression of these genes would benefit from treatment with this drug. Further experimental research to determine the correlation between increased tanespimycin sensitivity and the
overexpression of these genes would be necessary to determine the mechanism of the effect their overexpression has on tanespimycin.

References


Hedlund, E. et al (2012) Tumour Cell-derived Placental Growth Factor Sensitises Antiangiogenic and Antitumour Effects of Anti-VEGF drugs *Proceedings for the*


Kunkle, B. et al (2009) ONCOMINE meta-analysis of breast cancer microarray data identified up-regulation of NRF-1 expression in human breast carcinoma 18th World IMACS/MODSIM Congress, Australia


Loewe, L. (2008) Genetic Mutation Nature Education Vol.1 pp.113


Ries, LAG. et al (2003) Cancer Survival and Incidence from the Surveillance, Epidemiology and End-results (SEER) program Oncologist vol. 8 pp.541-542 http://dx.doi.org/10.1634/theoncologist.8-6-541


Trinh, B. Ko. S Barengo, N. Lin, S. & Naora, H. (2013) Dual Functions of the Homeoprotein DLX4 in Modulating Responsiveness of Tumor Cells to Topoisomerase II-Targeting Drugs *Cancer Research* Vol. 73 pp. 1000 [http://dx.doi.org/10.1158/0008-5472.CAN-12-3538](http://dx.doi.org/10.1158/0008-5472.CAN-12-3538)


