

## 1. Introduction

- Approximately 20% of patients with Stage II colon cancer develop recurrent disease following resection of the primary tumor.<sup>1</sup>
- Identification of patients with high risk stage II colon cancer may allow appropriate targeting of adjuvant chemotherapy.<sup>1</sup>
- Colon cancer is likely to be a collection of distinct diseases when studied at a molecular level.
- These molecular subtypes may have prognostic and predictive significance
- Unsupervised approaches can be used to identify molecular subtypes but rely on large numbers of samples to generate valid data.

In this study we have assessed the suitability of the Colorectal Cancer DSA™ microarray platform<sup>2</sup> and formalin fixed paraffin embedded samples for the identification of molecular subtypes in colon cancer.

## 2. Materials & Methods

### Samples

600 Stage II colon cancer 10µm FFPE samples were collected from 12 centres.

### Gene expression profiling

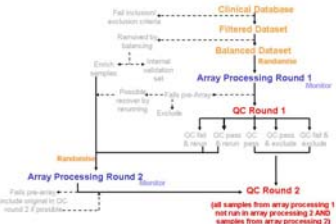
Total RNA from FFPE tumor tissue was extracted using the Roche Extraction kit and amplified using the NuGEN WT-Ovation™ FFPE System. The amplified product was hybridised to the Almac Colorectal Cancer DSA™ research tool.

### Analysis

Expression data was RMA normalized. 27,431 probe sets were included after intensity filtering. Further probe sets were removed, if ANOVA results suggested that a probe set was dominated by technical effects (e.g. Present Call or Physiological Lab vs. clinical parameter). Principal Component Analysis (PCA) and Hierarchical Clustering were used for unsupervised analysis. Expression of probe set for ligand X > 2000 was used to define a sample as Ligand X positive. Differential Expressed probe sets were selected based on a fold change > 2 and a t-test with correction for multiple testing (FDR = 5%). Chi square test was used to test the significance of good/poor distribution for Ligand X positive and negative group.

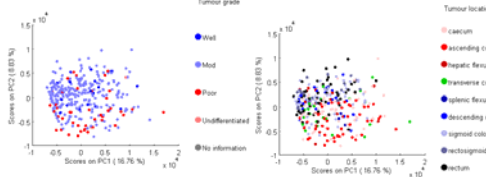
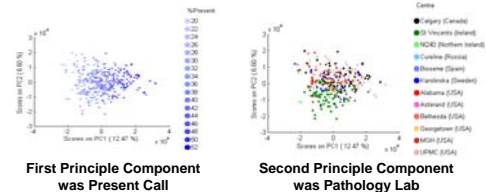
Functional analysis was performed using GeneGo Metacore knowledgebase

## 3. Quality Assessment



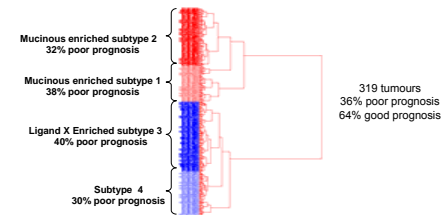
319 samples were selected following the quality assessment

## 4. Identification of Biases in Samples



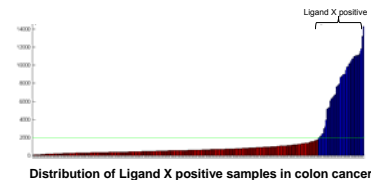
Removal of transcripts from the analysis removed these biases

## 5. Hierarchical Clustering of Data



Hierarchical clustering identified at least four distinct molecular subtypes

## 6. Subtype 3: Enriched for Ligand X



	Ligand X positive	Ligand X negative	P value
Good prognosis	22 (50%)	183 (66%)	0.033
Poor prognosis	22 (50%)	92 (34%)	

The ligand X expression indicates poor prognosis

## 7. Other Subtypes: Characterization

### Subtype 1: Mucinous Enriched

- Expression of Muc genes
- CXCL14 overexpressing
- 38% poor prognosis

### Subtype 2: Mucinous Enriched

- Expression of Muc genes
- Activation of Wnt signalling (p=0.0038)
- Down regulation of REG1A (a known poor prognostic marker in colon cancer)
- 32% poor prognosis

### Subtype 4

- Activation of TGF-Beta signalling (p=0.02)
- Predominantly left sided colon
- 30% poor prognosis

## Conclusions

- It is possible to identify molecular subtypes within colon cancer using archived FFPE samples.
- Some subtypes may have prognostic significance.
- Some subtypes may be defined by pathways that are therapeutically targetable.
- Knowledge of molecular subtypes from archived samples may enhance the development of prognostic or predictive biomarker tests and identify novel targets.

## References

1. Benson, AB. New Approaches to the Adjuvant Therapy of Colon Cancer. *The Oncol* (2006); 11: 973.
2. Farragher S et al. RNA expression analysis from formalin fixed paraffin embedded tissues. *Histochem Cell Biol*. 2008 Sep;130(3):435-45