Clustering from structural variation in endometrial cancer

Julian Stanley, Shreepriya Das, Michele Cummings, Nicolas Orsi, & Jeremy Gunawardena
Department of Systems Biology, Harvard Medical School, Boston, MA 02115

Endometrial cancers are the fourth most common cancer in women worldwide, and identification of molecular characteristics that define subtypes of the cancer may complement traditional diagnostic techniques. We obtained 95 hysterectomies of different grades of endometrial cancer and analyzed structural variation events from formalin-fixed samples. Hierarchical clustering of copy number variation (CNV) revealed two clusters of endometrial cancer that also correspond to survival. Clustering can also be accomplished by a simple technique of plotting patients in terms of aggregate copy number gain and loss. Further work needs to be done to identify genetic markers that may aid diagnosis, and how those markers compare across cancer types.

A simplified view of endometrial cancer subtypes

Grade 1 or 2 Endometroid
"Type 1"
Better Prognosis

Grade 3 Endometroid
"Type 1" or "Type 2"
Worse Prognosis

Non-Endometroid
"Type 2"
Better Prognosis

Copy number variation (CNV), allelic imbalance, and loss of heterozygosity (LOH) are types of structural variation

Structural variation analysis

Grade 1 Endometroid (11)
Better Prognosis

Grade 2 Endometroid (21)
Worse Prognosis

Grade 3 Endometroid (32)
Non-Endometroid (31)

Copy number variation (CNV), allelic imbalance, and loss of heterozygosity in 95 endometrial cancer FFPE samples

Does the method also apply to ovarian cancer?

Low grade endometroid are least severe → Low copy number

Mucinous is intermediate, but based on CN we predict less severe

Low grade serous and clear cell are intermediate; they distribute normally

High grade serous are most severe → High copy number

Moving forward: comparative analyses

Figure courtesy of Watkins et al. 2014 in Breast Cancer Research. Green bar presents copy number gain, red bar represents copy number loss.

Discussion

Both endometrial and ovarian cancer are extremely heterogeneous diseases that affect hundreds of thousands of women worldwide. Microarray analysis is a powerful tool that may provide information that can lead to more personalized cancer treatments.

The Oncoscan platform, based on the molecular inversion probe technique, can consistently detect variation heavily-degraded FFPE clinical samples. Clustering techniques and comparative analyses are necessary to understand what differences between cancers are visible from genomic analyses, and how different molecular profiles can complement traditional diagnostics.

Acknowledgements

I am extremely grateful to HMS Systems Biology and the Summer Internship in Systems Biology program for the welcoming working environment. Thanks to Shreepriya Das and Jeremy Gunawardena for their invaluable mentorship during this short internship, and to Nicolas Orsi and Michele Cummings for providing samples and helpful clinical expertise.

References

2. Zhang, J. 2018 CNBrick: Convert segment data into a region by sample matrix to allow for other high level computational analyses. R package version 1.36.0