A Novel Machine Learning Based Method To Detect Homozygous Deletion Of Homologous Recombination Repair (HRR) Genes In Prostate Cancer

Jianging Wang1*, Xin Ye2*, Peng Cheng3, Usha Singh4, Katherine Bell5, Karen Unshak6, Lihua Liu5, Xuesong Lyu7, Wenfang Ku8, Zhai Huang5, Xuqiu Chen9, Sheng Liu9, Yonghua Guo10, Weijie Shu10, Liangen Zhou11, Shuang Yang11*

1. Amoy Diagnostics, Xiamen, China. 2. Janssen Research & Development, Shanghai, China. 3. Janssen Research & Development, Succas, N.J. 4. Janssen Research & Development, Spring House, PA, USA. 5. GSK, Beijing, China. 6. GSK, Shanghai, China. 7. GSK, Shanghai, China. 8. GSK, Shanghai, China. 9. GSK, Shanghai, China. 10. GSK, Shanghai, China. 11. GSK, Beijing, China.* Correspondence: yangs@amoydx.com; yex21@its.jnj.com

BACKGROUND

- HRR gene alterations are enriched in prostate cancer (PC) and are associated with sensitivity to PARP inhibitors.
- Homozygous deletion (HD), which accounts for up to 15% of HRR gene alterations in PC, is an important type of pathogenic alteration. However, the reliable detection of HD is technically challenging.
- A novel machine learning based method for HD detection in formalin-fixed paraffin-embedded (FFPE) PC tissues was introduced.

METHODS

- AmoyDx HRD Complete was developed to detect HRR gene alterations, with elaborated design for HD detection.
- HD/min HD simulation events (~30000) of varying tumor cellularity and fragments size were constructed using the sequencing data of matched tumor and wild type cell lines.
- Seven artificial features were calculated based on genomic segment depth and minor allele frequency of single nucleotide polymorphisms, tumor cellularity and tumor ploidy.
- The feature data were fed into the algorithm to generate HD recognition model (training validation dataset). The feature data were fed into the algorithm to generate HD recognition model (training validation dataset).
- Clinical validation: 132 FFPE PC tissues showed high consistency with a second validated NGS assay (svNGS) (PPA 100%, NPA 98.4%, OPA 98.5%).

RESULTS

- In silico validation: the trained HD model achieved 99.8% accuracy in the in silico validation dataset.
- Analytical validation: 100% specificity; 100% sensitivity at gene level with 30% tumor cellularity and at exon level with 40% tumor cellularity with 100ng DNA input; 100% reproducibility (10 repeats per sample in 3 lots of critical reagents, 3 days, 3 operators); strong anti-interference capability with all HD events detected and no false positives.
- Clinical validation: 132 FFPE PC tissues showed high consistency with a second validated NGS assay (svNGS) (PPA 100%, NPA 98.4%, OPA 98.5%).

CONCLUSIONS

A machine learning based method was successfully developed and adopted in multiple NGS assays to detect homozygous deletion of HRR genes in prostate cancer tissues.

It provides a reliable approach to effectively identify prostate cancer patients with HRD homozygous deletion.

ACKNOWLEDGMENTS

This study was funded by both Janssen and AmoyDx.

DISCLOSURES

W.K., Singh S., Bell K., Unshak K., Liu L., Lyu K., and Zhou L. are the full time employees of AmoyDx. GSK. Weijie Shu, Xuesong Lyu, Yonghua Guo, and Shuang Yang C. are the full time employees of AmoyDx.