13P
#533

Analytical validation of HER2DX test for early-HER2+ breast cancer

E. Sanfeliu Torres[1], M. Marín-Aguilera[2], P. Jares[3], G. Villacampa[2], E, Hernández-lllán[3], B. Gonzalez-Farre[1], F. Brasó-Maristany[4], P. Galván[4], O. Castillo[4], P. Blasco[4], V. Sirenko[4], A. Aguirre[4], J. A. Puig-Butille[3] , A. Vivancos[5], J. Matito[5], C. M. Perou[6], P. Villagrasa Gonzalez[2], A. Prat[2,4,7,8], J. Parker[2], L. Paré Brunet[2]

[1]Pathology, Hospital Clínic Barcelona, Spain; [2]Scientific Department, Reveal Genomics, S.L.; Barcelona, Spain; [3]Molecular Biology Core, Hospital Clinic Barcelona, Bar<elona Spain; [4]IDIBAPS - Translational genomics and targeted therapies in solid tumours lab,- Fundació de Recerca Clínic Barcelona, Spain; [5]Genomics department, Vall d'Hebron University Hospital, Barcelona, Spain; [6]Genetics, UNC - Lineberger Comprehensive Cancer Center, Chapel Hill, NC,US; [7]Medical Oncology Department, Hospital Clinic Barcelona, Barcelona, Spain; [8]University of Barcelona, Spain

Clínic Barcelona | Universitat de Barcelona

Translational genomics and targeted therapies in solid tumors

IDIBAPS Institut D'Investigacions Biomèdiques August Pi i Sunyer

VHIO Vall d'Hebron Institute of Oncology

Background and Objectives

HER2DX[1] is a unique 27-gene based algorithm that predicts the risk of relapse (RR-score), the probability of achieving a pathological complete response (pCR-score), and the individual levels of *ERBB2* expression (*ERBB2*-score) in early-stage HER2-positive (HER2+) breast cancer. It integrates clinical data with gene expression (measured by nCounter (Nanostring Technologies).

HER2DX is clinically available since January 2022 as a Laboratory Developed Test (LDT) in a central laboratory (CLab), *Centre de Diagnòstic Biomèdic*, which belongs to Hospital Clínic Barcelona (Spain),

Here, we aimed to analytically validate the test performed at CLab having as a reference the development laboratory (DLab). Performance of HER2DX by using RNAseq platform was also explored and compared to nCounter.

Methods

- HER2DX was performed as described previously[1].
- Two RNA controls from BT474 cell line were included in every run for a quality check of the procedure.
- All scores ranged from 0 to 100 and pre-defined cut-offs were used to get HER2DX-groups.
- Repeatability and reproducibility were evaluated from different tissue sections, RNA, or formalin-fixed paraffin embedded (FFPE) blocks.
- Simulations (N=1·10[6]) were used to calculate diagnostic values: sensitivity, specificity, positive and negative predictive values (PPV and NPV), and accuracy.
- Robustness was measured by evaluating the interference of non-tumour tissue and by using different RNA quantities. Spiking experiments of RNA from tumour in RNA from stroma was performed in 4 FFPE samples. Four samples were analysed with 3 different quantities.
- Differences due to the nCounter instrument (N=2), Tagset lot (N=2), and Tagset defrost cycles (N=2) were also evaluated.
- RNAseq (Illumina Exome Panel) was performed in 30 RNA samples. Illumina RNA Prep with enrichment and UDI adapters were used to the library preparation. Coverage was 2x101 paired-end reads.

References and acknowledgements

References: 1. Prat A et al. Ebiomedicine, 2022;

Funding: This study was funded by Reveal Genomics (Barcelona, Spain). Authors acknowledges support to technicians at Centre de Diagnòstic Biomèdic. This presentation is the intellectual property of the author/presenter.

COI: The first author does not have conflicts of interest.

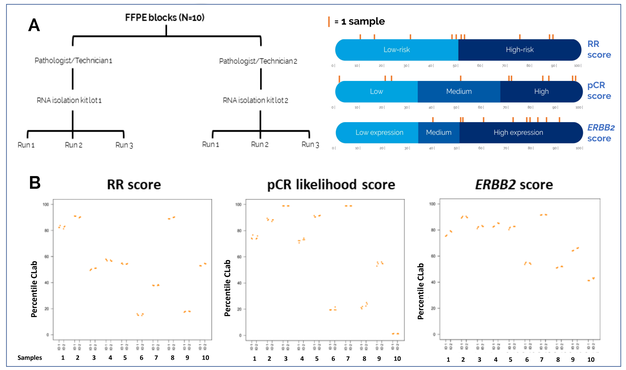
Contact them for permission to reprint and/or distribute at:
esanfeliu@clinic.cat; mmarin@reveal-genomics.com

Results

1- Repeatability within CLab

Experimental design is showed at Fig. 1A. The maximal standard error among the three scores between different technician was 0.94 (scale 0-100) (Fig. 1B).

Fig 1. Experimental design for Repeatability and variability for the three scores

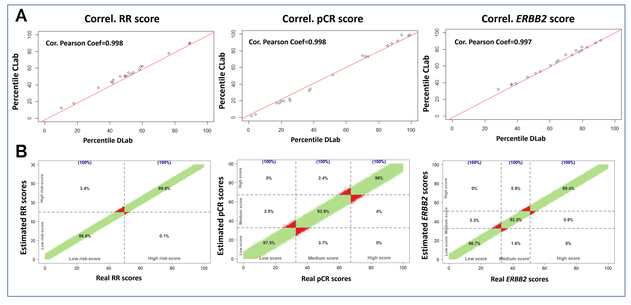


2- Reproducibility between CLab and DLab

Inter-lab HER2DX scores were compared by testing a set of FFPE samples previously analysed at DLab.

Reproducibility starting from RNA was analysed in 20 samples extracted and evaluated at DLab (Fig. 2). Correlations between RR, pCR, and *ERBB2* scores between both labs were >0.997 (Fig. 2A). In the simulation analysis, the probabilities of +/- 5 units difference in the risk-score, pCR-score, and *ERBB2-scores* were 0.5%, 5.2%, and 0.2%, respectively (Fig. 2B).

Fig 2. Reproducibility starting from the same RNA



Reproducibility starting from FFPE blocks was assessed in 29 tumour samples originally evaluated at DLab. Correlation coefficients are >0.97 (Fig. 3A). Bland–Altman plots did not show any relationship of the differences in score values (Fig. 3B). In the simulation (Fig. 3C), diagnostic values are >90% for the three scores (Table 1).

Fig 3. Reproducibility from FFPE block.

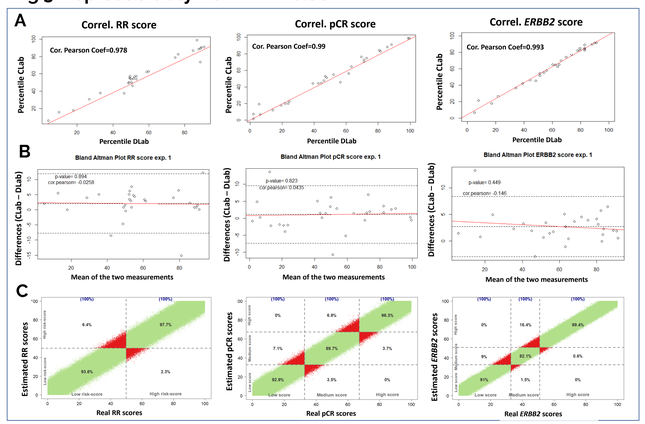


Table 1: Diagnostic values at reproducibility-simulation analysis

Diagnostic values	RR score	pCR score	*ERBB2* score
Sensitivity (%)	97.7	92.9	91
Specificity (%)	93.6	98.2	99.6
PPV (%)	93.9	96.2	99.1
NPV (%)	97.6	96.6	95.8
Accuracy (%)	95.7	96.5	96.8

3- Robustness of HER2DX

Robustness at low tumour content samples (10%) (Fig. 4A) or low RNA quantity (100 ng) (Fig. 4B) was acceptable for the assay. No significant differences were observed across different nCounter instruments (Fig. 4C), Tagset lots (Fig. 4D) and defrost cycles (Fig. 4E) at CLab.

Fig 4. Robustness of HER2DX. A) HER2DX was performed and compared between 3 dilutions of tumour RNA in stroma RNA, respect to undiluted sample (dilution % varied between samples, being Dil_1 the lowest and Dil_3 the highest); **B)** RNA starting concentrations to perform HER2DX were 100, 250 and 500 ng; **C)** Two different nCounter instruments within CLab, M1 and M2, using data from 20 and 24 samples, respectively, were compared; **D)** Two different lots of the Tagset reagent (Nanostring) were tested in 10 RNA samples; **E)** Two defrost cycles of the tagset were compared in 4 selected RNA samples.

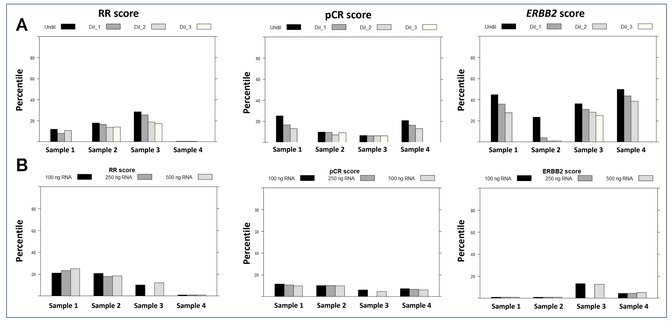
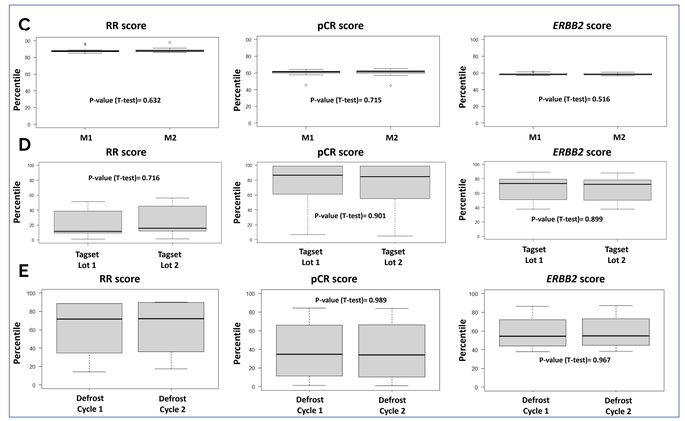


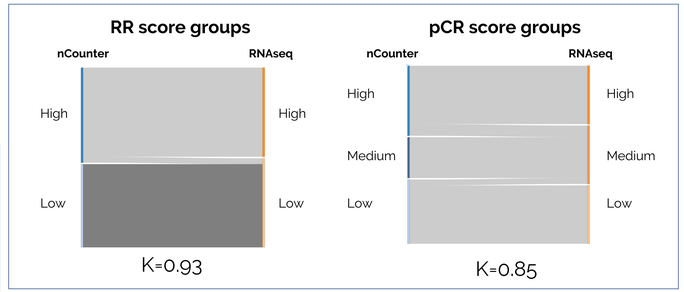
Fig 4. (cont.)



4- RNAseq and nCounter platforms HER2DX performance

Data from 25 of the 27 genes included in the HER2DX algorithm were analysed (two genes were missing at the exome panel). Correlation of gene expression between platforms was high (mean r=0.89, sd= 0.16). Only two genes (*ASPM* and *NTN3*) had a correl. Coef. <0.75. Coefficients of variation (CV) for each gene among the 30 samples within each platform was calculated and compared between platforms. The mean difference in CV was 0.01 (sd=0.28). Two genes (*NTN3* and *TCAP*) had a CV difference higher than 0.5. Concordance between the two platforms in RR, and pCR groups was 96.7%, and 90.0%, respectively. The Cohen kappa (K) was greater than 0.8 in both cases, indicating a high level of agreement between platforms (Fig. 5).

Fig 5. RNAseq and nCounter platforms comparison.



Conclusions

- There is a high association in HER2DX-groups between C- and D-Labs.
- HER2DX from RNAseq and nCounter platforms were highly correlated.
- Analytical validation of HER2DX has proven to be suitable for its intended purpose.