Cross-study meta-analysis identifies unique bacterial strains separating responder and non-responder patient populations across multiple checkpoint-inhibitor therapy datasets

SECOND GENOME

Jayamary Divya Ravichandar, Erica Rutherford, Nicole Narayan, Yonggan Wu, Thomas Weinmaier, Cheryl-Emiliane Chow, Shoko Iwai, Helena Kiefel, Kareem Graham, Karim Dabbagh, Todd DeSantis

Abstract

The gut microbiota has emerged as an important modulator in cancer progression and a growing body of evidence supports the influence of gut microbiota on response to cancer therapy, especially in the context of checkpoint inhibitor therapy. While several studies present insight into the landscape of microbial shifts modulating response to checkpoint inhibitors, they may be unduly influenced by cohort, sequencing-technology, and data analysis methods. Further, individual studies are often under-powered to detect microbes differentially abundant in responder and non-responder populations, which can limit therapeutic development. Key to microbiome-based drug discovery is the identification of proteins with therapeutic potential that are efficacious across cohorts. Herein, existing published datasets in the checkpoint-inhibitor space were mined and integrated via a cross-study meta-analysis to identify bacterial strains separating responder and non-responder melanoma patient populations.

We compared the baseline gut microbiota associated with stool samples collected from five discrete cancer patient cohorts undergoing checkpoint-inhibitor therapy. Samples were sequenced on one or more technologies (Illumina 16S NGS, 454 16S NGS, and Illumina shotgun metagenomics) and a total of seven publicly-available datasets were analyzed herein. Leveraging our multi-faceted bioinformatics platform, which enables appropriate method-specific quality filtering and statistical testing to identify differentially abundant bacteria at the strain-level, we were able to successfully integrate analysis results across multiple microbiome-profiling technologies. We performed a random effects model based meta-analysis and identified strains that were concordantly enriched in responder populations across datasets. The strains identified herein present opportunities for mining proteins with potential to improve response to checkpoint inhibitors.

This cross study meta-analysis demonstrates the power of Second Genome's bioinformatics pipeline to leverage publicly available datasets and systematically integrate microbial shifts not only across samples from multiple cohorts but also across samples sequenced on different technologies. Our in-house strain database that enables taxonomic annotation down to the strain-level allowed for comparison of fine-grained bacterial identities across datasets, resolving a key challenge with microbiome meta-analysis. This systematic and statistically-driven integration of datasets enabled identification of strains associated with response across multiple responder populations that were not previously reported in the independent analysis of these datasets.

Challenges

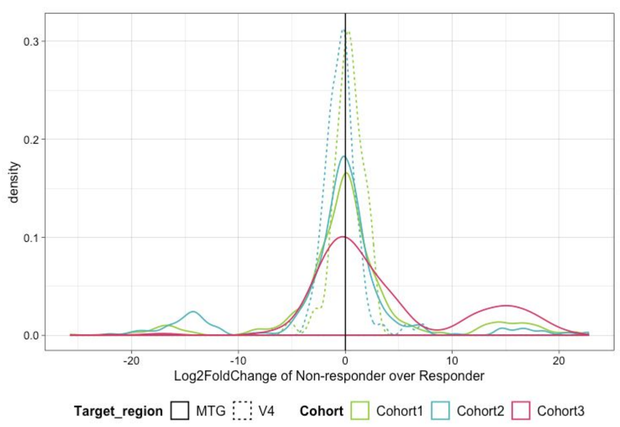
- Gut microbiota has emerged as key in modulating response to immune checkpoint inhibitor therapies (ICT) in melanoma.
- Batch effects are predominant in independent studies and may bias conclusions .
- **Can a robust bioinformatics pipeline identify bacterial modulators of ICT-response in melanoma patients supported across multiple cohorts?**

Cohorts included in the meta-analysis

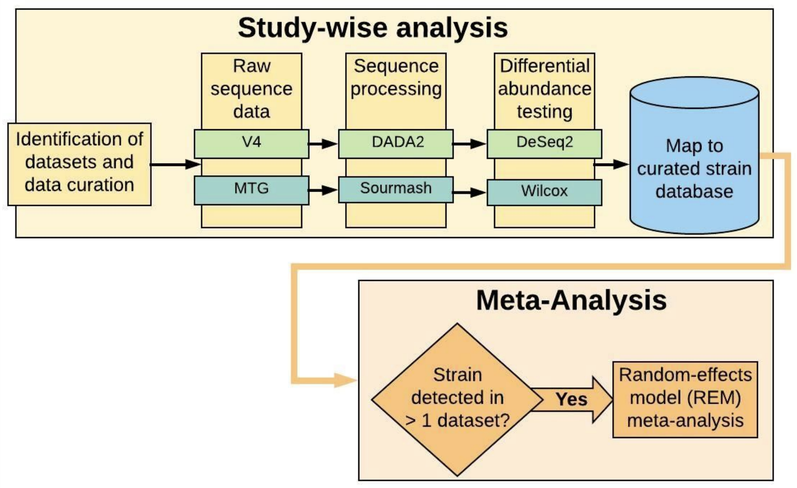
- Stool from 3 cohorts (5 datasets) of melanoma patients on anti-PD1 ICT therapy was included.
- Microbiome of responders (R) & non-responders (NR) was compared.

Dataset	Target region	Design
Cohort 1, *Matson et al.*	V4	R:15, NR:26
	MTG	R:14, NR:24
Cohort 2, *Gopalakrishnan et al.*	V4	R:30, NR:13
	MTG	R:14, NR:11
Cohort 3, *Frankel et al.*	MTG	R:3, NR:7

- **Distribution of effect sizes are cohort and platform dependent.**

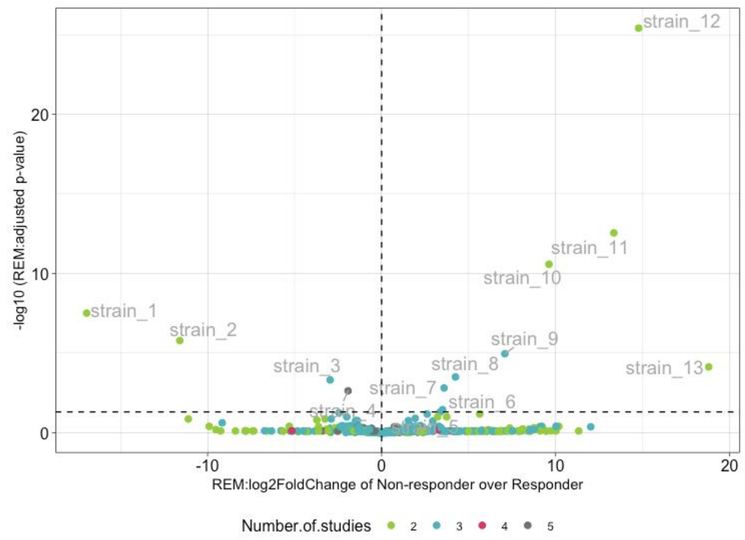


Methods

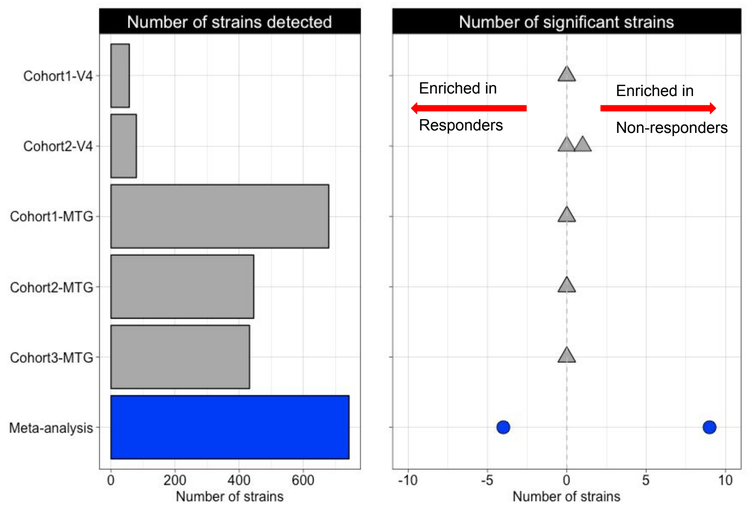


Cross-study meta-analysis

- **Meta-analysis identified 4 strains significantly enriched in responders and 9 significantly enriched in non-responders.**
- Significance: Benjamini-Hochberg-corrected REM p values < 0.05.



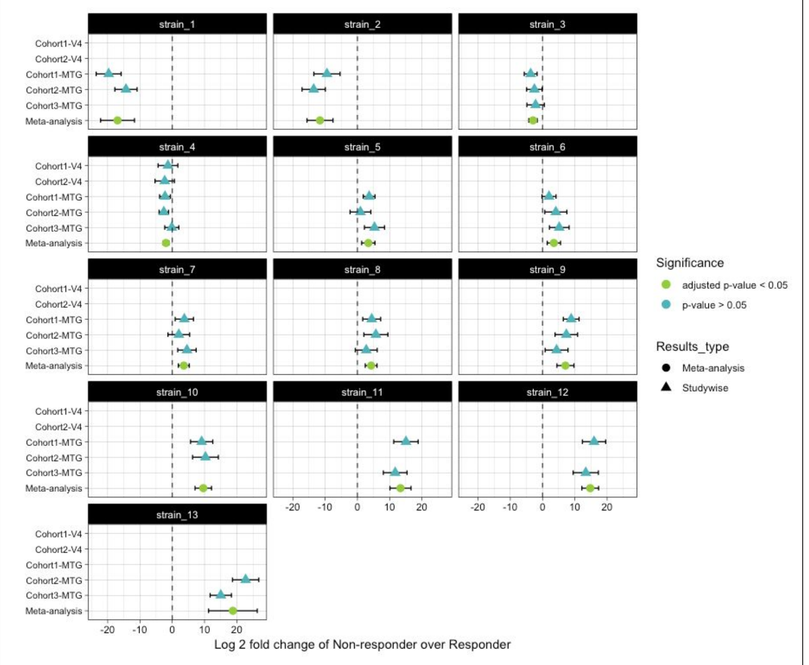
- **Meta-analysis is more sensitive at identifying significant changes compared to individual studies.**



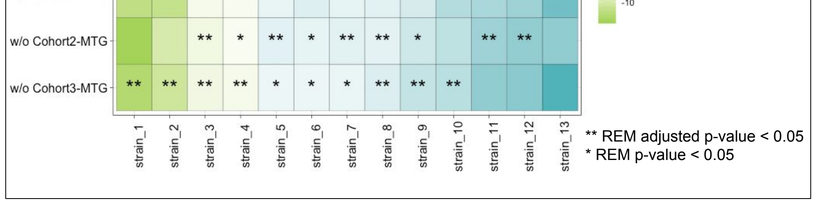
- *Matson et al.* (2018) Science.
- *Gopalakrishnan et al.* (2017) Science.
- *Frankel et al.* (2017) Neoplasia.

Strain signatures are supported across cohorts

- **Meta-analysis identified significantly enriched strains that were not inferred as enriched in independent analysis of each dataset.**



- Datasets were re-analyzed in meta-analysis using the leave one-out method.
- Strain signatures identified here, are not highly cohort-specific.
- The strain signature, however, may be biased by sequencing methodology-specific variation, with MTG datasets driving meta-analysis results.



Conclusions

- **Our bioinformatics pipeline successfully integrates data sequenced on different platforms and identifies trends not observed in independent analysis of studies.**
- **Our in-house strain database that enables taxonomic annotation down to the strain-level allowed for comparison of fine-grained bacterial identities across datasets, resolving a key challenge with microbiome meta-analysis.**
- ICT-response associated bacterial signatures identified from this systematic and statistically-driven integration of datasets present novel therapeutic opportunities for mining proteins modulating response to therapy.
- Alternatively, bacterial signatures supported across multiple cohorts can be explored as biomarker targets for determining response to ICT-therapy.