Novel machine-learning bioinformatics reveal distinct metabolic alterations for enhanced colorectal cancer diagnosis and monitoring

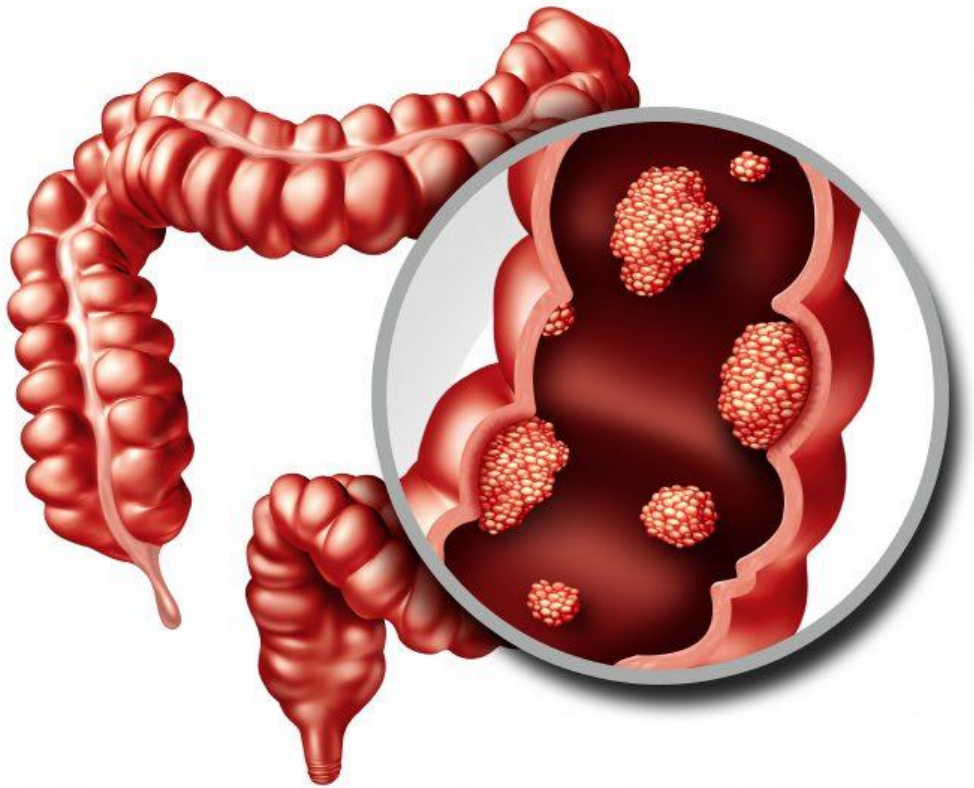
Rui Xu[1,2], Hyein Jung[3], Fouad Choueiry[1, 2], Shiqi Zhang[1,2], Rachel Pearlman[2,4], Heather Hampel[5,6], Ning Jin[2], Jieli Li[7]*, Jiangjiang Zhu [1,2]*

[1]Human Nutrition Program, Department of Human Sciences, The Ohio State University, Columbus 43210, The United States
[2]Comprehensive Cancer Center, The Ohio State University, Columbus, 43210, The United States
[3]Department of Chemistry and Biochemistry, The Ohio State University, Columbus 43210, The United States
[4]Department of Internal Medicine, The Ohio State University, Columbus 43210, The United States
[5]Department of Medical Oncology & Therapeutics Research, City of Hope National Cancer Center, Duarte 91010, The United States
[6]Division of Clinical Cancer Genomics, Beckman Research Institute, City of Hope 91010, The United States
[7]Department of Pathology, The Ohio State University, Columbus 43210, The United States

Rui Xu, Hyein Jung, Fouad Choueiry, Shiqi Zhang, Rachel Pearlman, Heather Hampel, Ning Jin, Jieli Li, Jiangjiang Zhu. 2025. Novel machine-learning bioinformatics reveal distinct metabolic alterations for enhanced colorectal cancer diagnosis and monitoring. ***iMetaOmics*** 2: e70003.
https://doi.org/10.1002/imo2.70003

Introduction

Early diagnosis of colorectal cancer (CRC) reduces mortality and improves treatment outcome

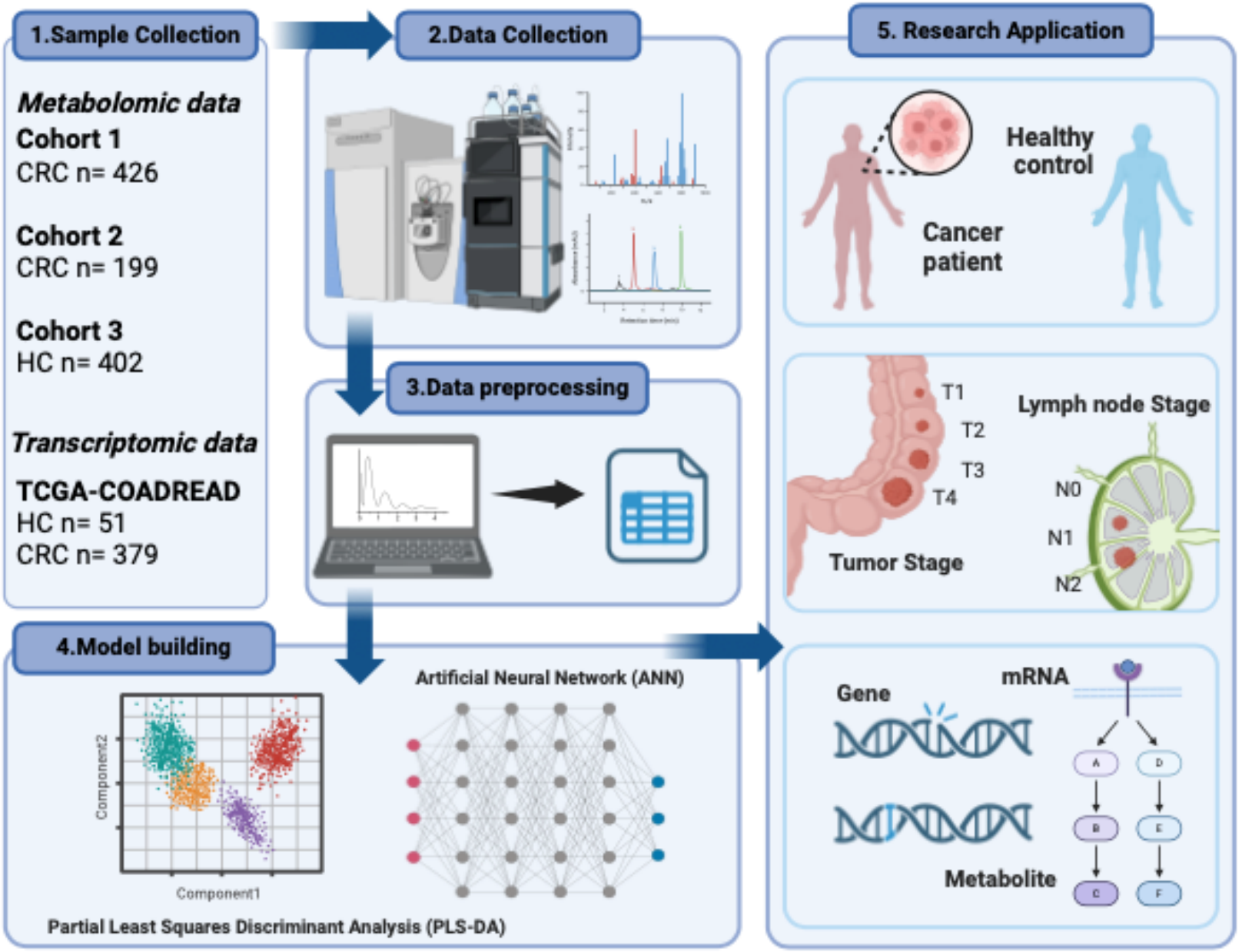


- **Incidence:** High incidence globally, especially in developed countries.
- **Mortality:** High mortality rates, while early diagnosis significantly reduces mortality rates.
- **Treatment:** Surgery, chemotherapy, radiation, and targeted therapy. Early diagnosis enables healthcare professionals to initiate appropriate treatment strategies and successful outcome.
- **Early Diagnosis Challenges:**
  - The absence of specific symptoms in the early stages.
  - Diagnostic methods like colonoscopies, are invasive and uncomfortable, leading to reluctance among individuals to undergo screening.

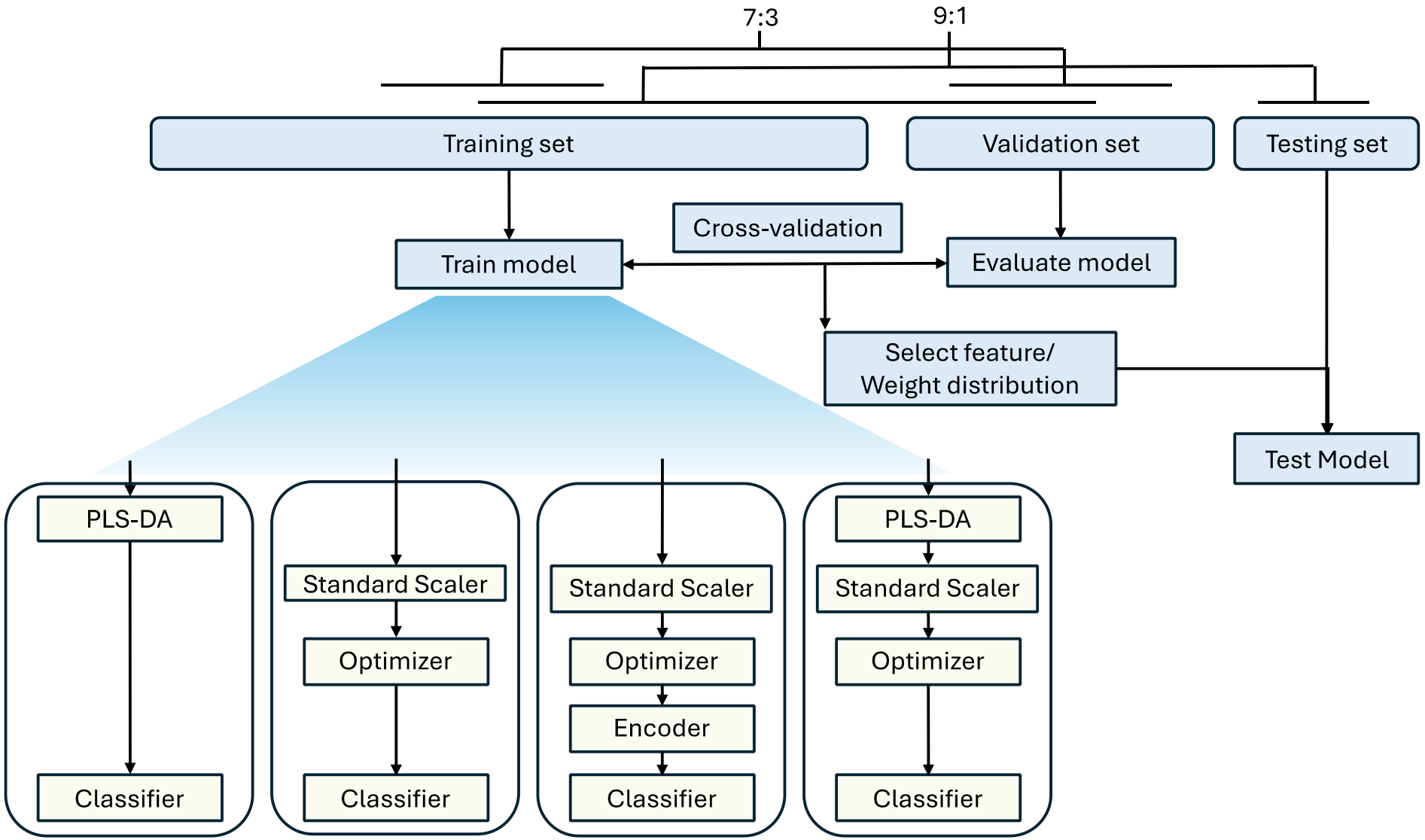
Highlights

- Significant upregulation of key metabolic pathways, including the TCA cycle, purine metabolism, and amino acid metabolism, was observed in CRC cases compared to healthy controls.
- Metabolic shifts correlated with tumor (T) and lymph node (N) stages, including increased pyruvic acid levels and decreased phenol levels in metastatic cases.
- Purine metabolism showed distinct patterns, with upregulation in CRC but downregulation in advanced tumor stages, linked to oncogenic signaling and nutrient deprivation.
- The PANDA pipeline effectively integrated metabolomic and transcriptomic data, enhancing robustness in CRC biomarker discovery and progression monitoring.

Overview of Our Study



The PANDA Workflows for Metabolite Biomarker Selection



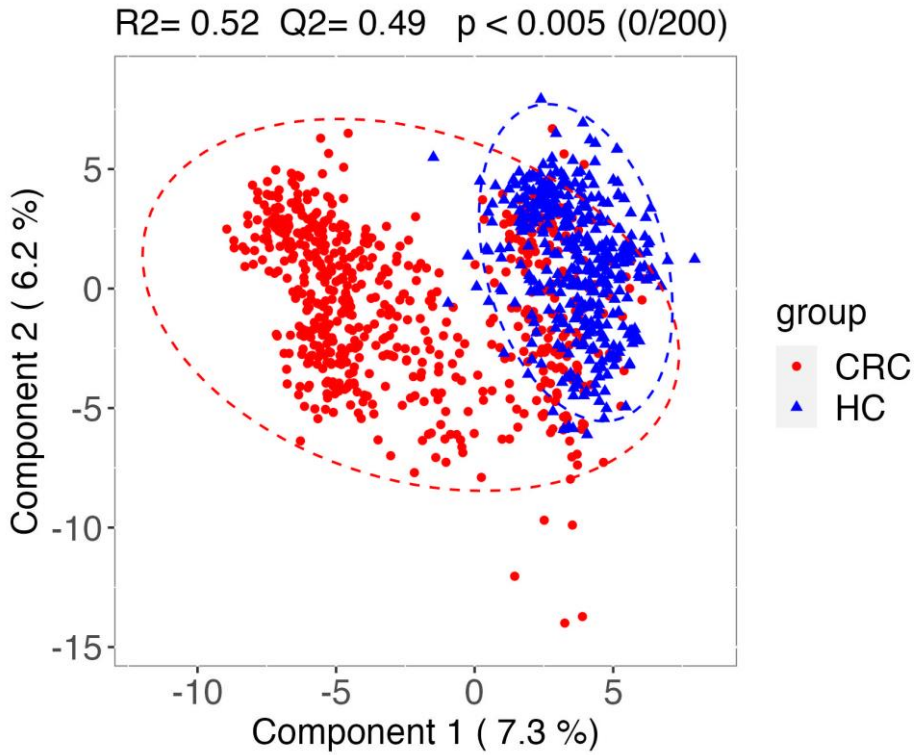
Metabolomic and Transcriptomic Analyses Distinguish CRC vs. Healthy Control

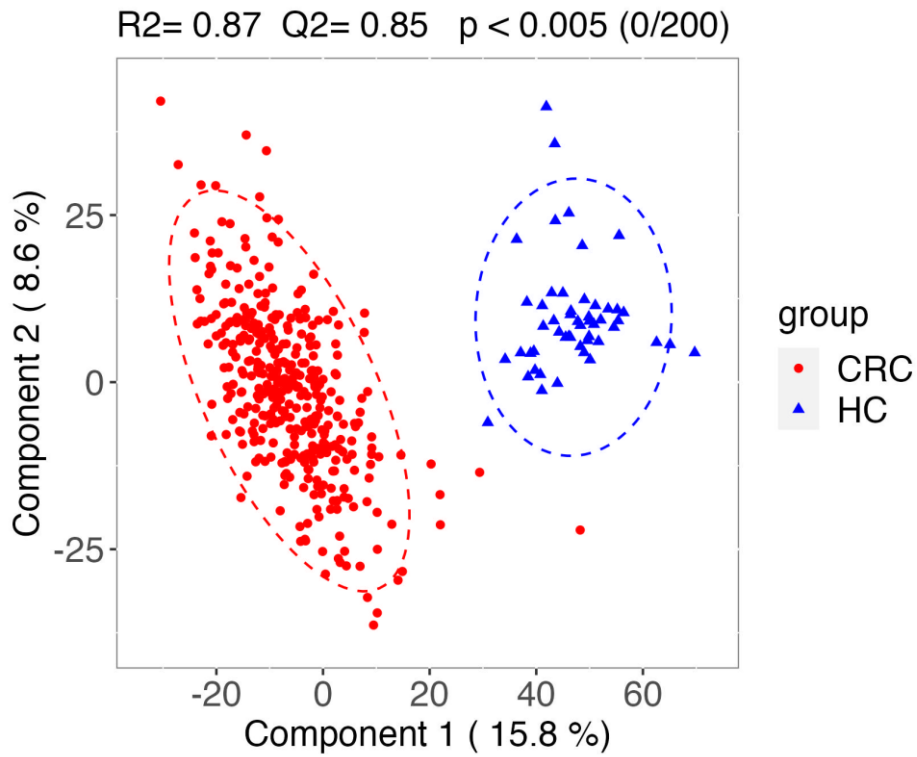
Metabolites N=240

	CRC	HC
Training	393	253
Validation	169	109
Testing	63	40

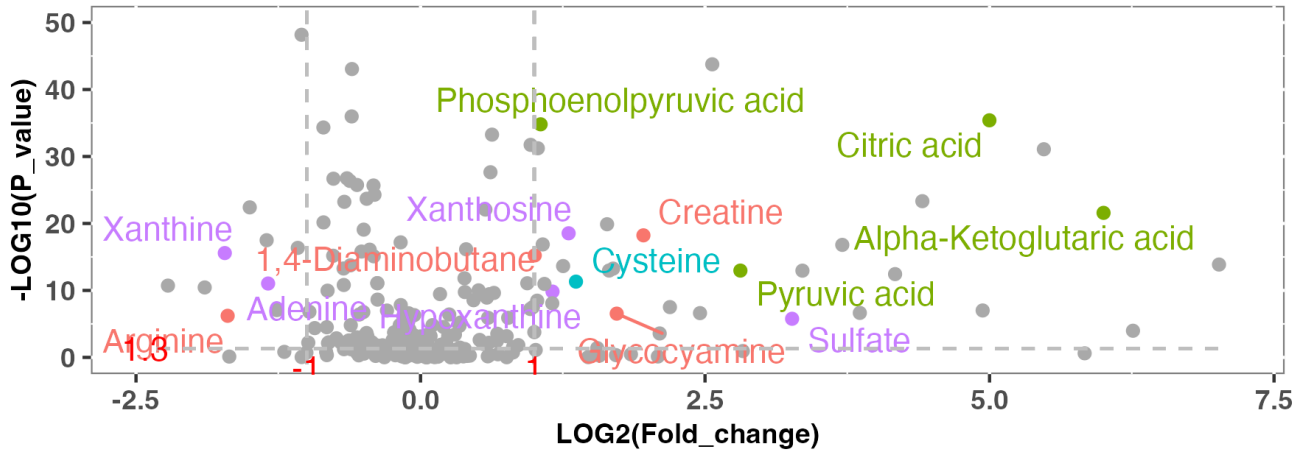
mRNAs N=20503

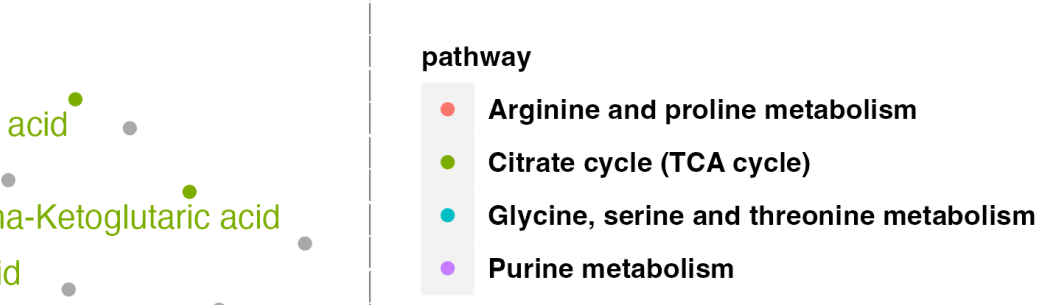
	CRC	HC
Training	239	32
Validation	102	14
Testing	38	5

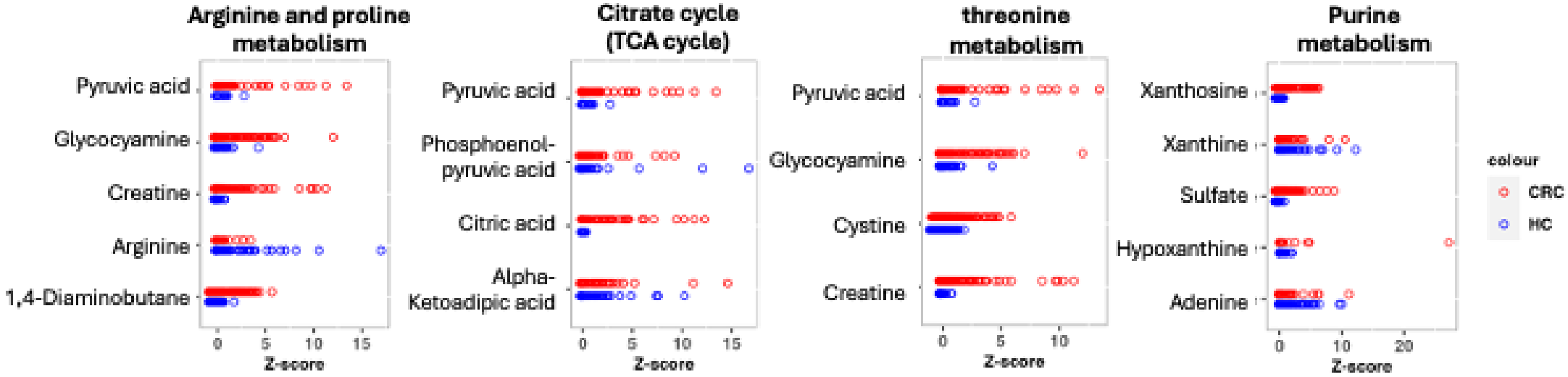




CRC Induce Metabolic Reprograming in Multiple Metabolic Pathways

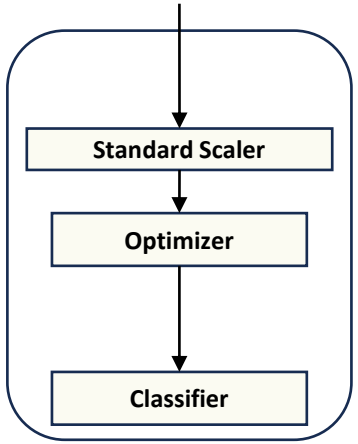




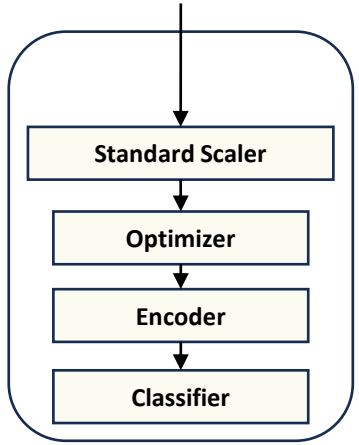


PANDA Supported Cancer Stage Classification

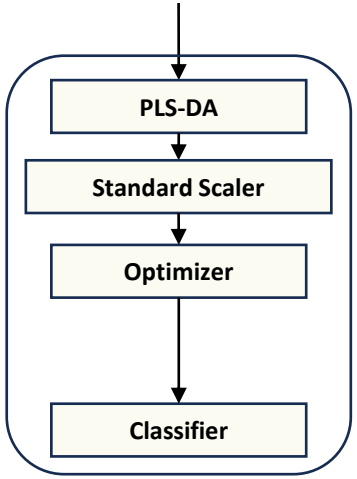
Tumor stage	T1	T2	T3	T4
Training	27	41	117	35
Validation	11	18	59	15
Testing	4	6	19	6



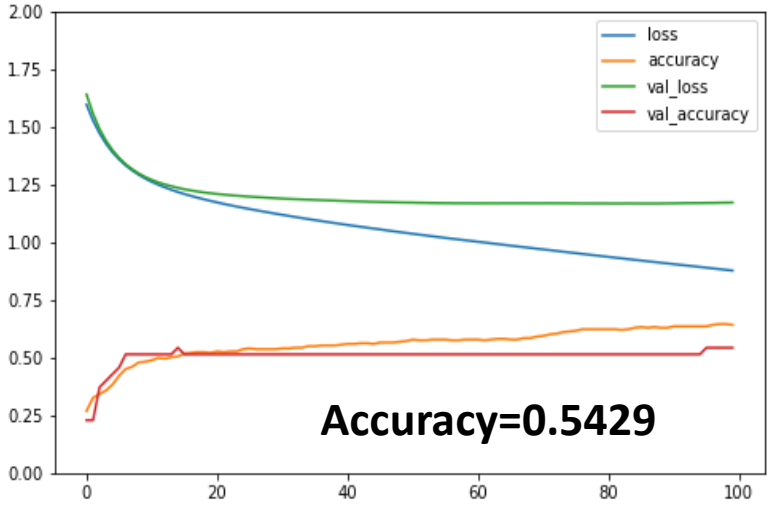
Metabolites N=240

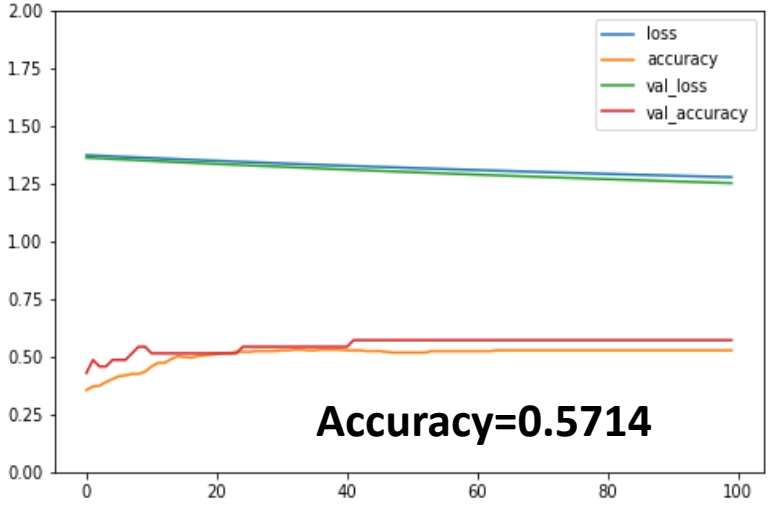


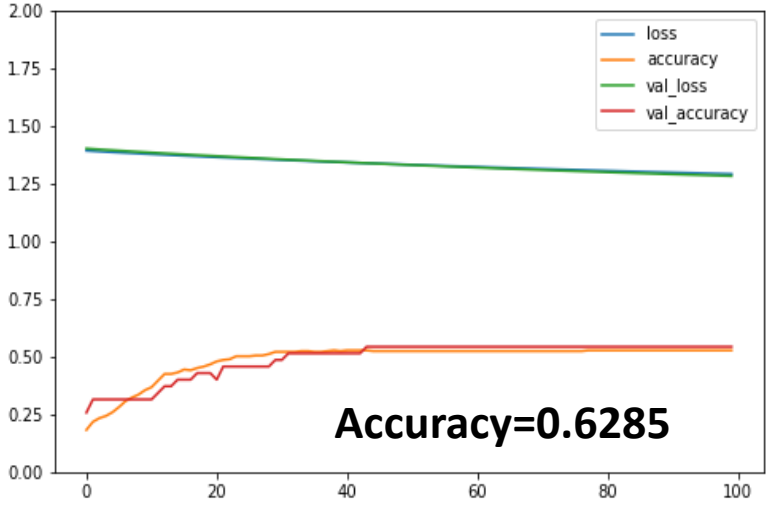
Encoded features N=3



Components 1-8 from PLS-DA N=8

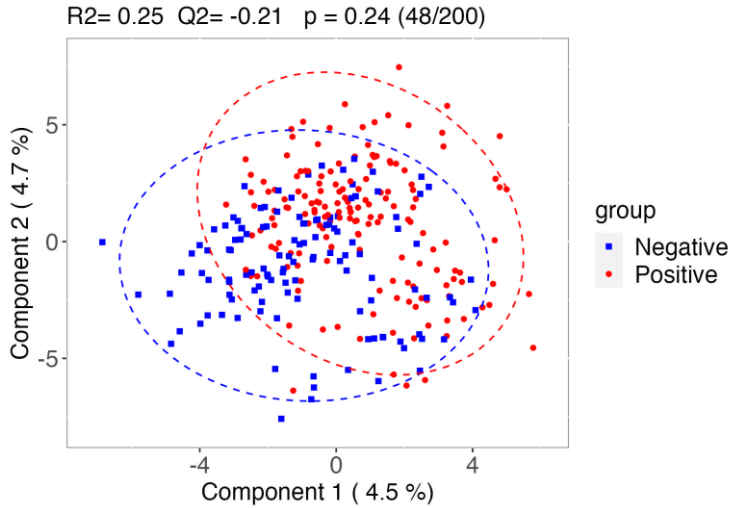


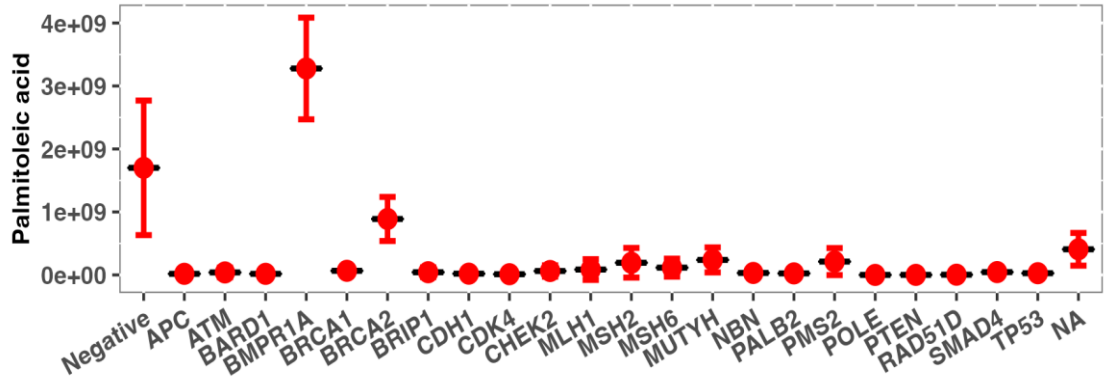


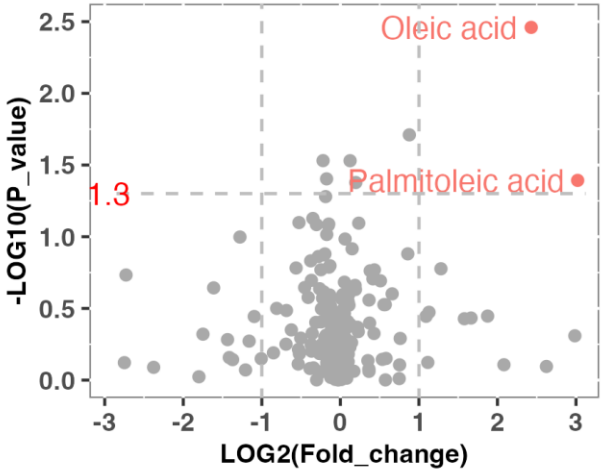


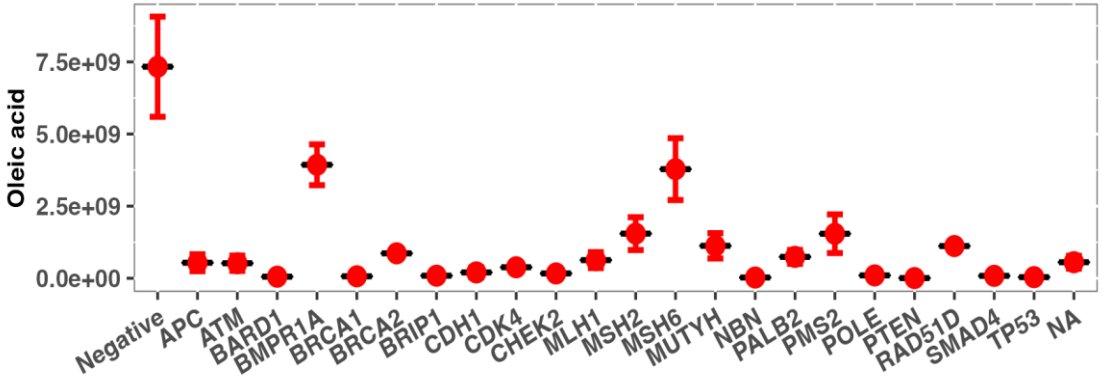
Gene Mutation Associated Metabolic Changes in CRC

Genes	Metabolomic	Transcriptomic
Negative	123	Na
APC	7	161
ATM	15	29
BARD1	3	17
BMPR1A	4	61
BRCA1	1	0
BRCA2	2	1
BRIP1	4	3
CDH1	5	32
CDK4	1	1
CDKN2A	5	26
CHEK2	8	1
MLH1	24	42
MSH2	30	4
MSH6	24	4
MUTYH	6	5
NBN	1	3
PALB2	2	3
PMS2	17	0
POLE	1	3
PTEN	1	61
RAD51D	1	6
SMAD4	1	115
TP53	1	62









Summary

- Utilized two distinct statistical methodologies, PLS-DA and ANN, we developed a novel pipeline of the combination of PLS-DA and ANN to analyze metabolic profiles associated with CRC diagnosis and prognosis.

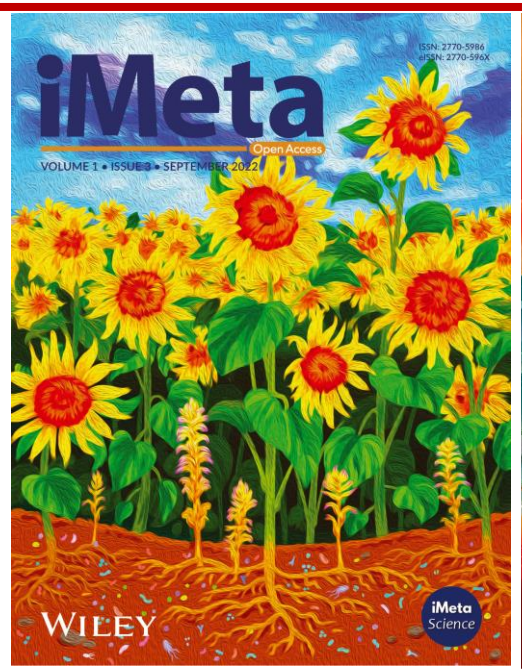
- We successfully identified significant metabolic alterations in CRC cases, including upregulation in pathways such as arginine and proline metabolism, citrate cycle, glycine, serine, and threonine metabolism, and purine metabolism

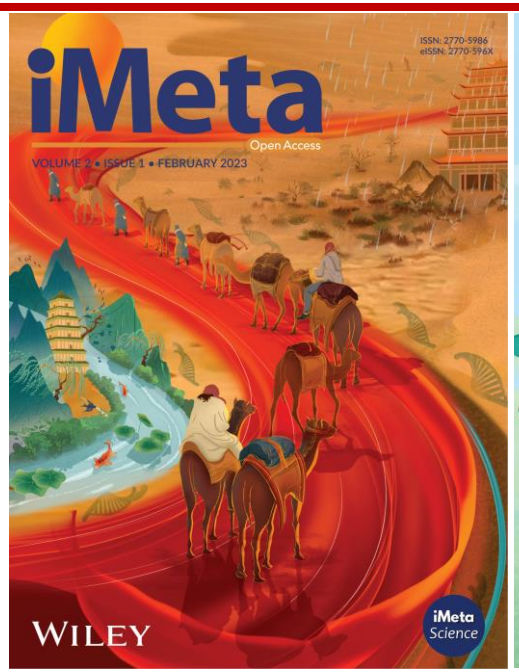
- We uncovered distinct metabolic changes associated with CRC prognosis and related genetic mutations, reinforcing findings through the integration of metabolomic and transcriptomic data.

Rui Xu, Hyein Jung, Fouad Choueiry, Shiqi Zhang, Rachel Pearlman, Heather Hampel, Ning Jin, Jieli Li, Jiangjiang Zhu. 2025. Novel machine-learning bioinformatics reveal distinct metabolic alterations for enhanced colorectal cancer diagnosis and monitoring. ***iMetaOmics*** 2: e70003. https://doi.org/10.1002/imo2.70003

iMeta: Integrated meta-omics to change the understanding of the biology and environment

WILEY





"***iMeta***" is a Wiley partner journal launched by iMeta Science Society in 2022, first impact factor (IF) **23.8** in 2024, ranking top 107/21973 in world and 2/161 in the microbiology. It aims to publish innovative and high-quality papers with broad and diverse audiences. **Its scope is similar to *Nature*, *Cell*, *Nature Biotechnology/Methods/Microbiology/Medicine/Food* etc**. Its unique features include video abstract, bilingual publication, and social media dissemination, with more than 600,000 followers. Indexed by **SCIE**, **PubMed**, **Google Scholar** etc.

"***iMetaOmics***" launched in 2024 and ***iMetaMed*** in 2025, with a **target IF>10, and its scope is similar to *Nature Communications***, ***Cell Reports***, *Microbiome, ISME J, Nucleic Acids Research, Briefings in Bioinformatics,* etc. All contributes are welcome!

Society: http://www.imeta.science
Publisher: https://wileyonlinelibrary.com/journal/imeta

Submission: https://wiley.atyponrex.com/journal/IMT2
https://wiley.atyponrex.com/journal/IMO2

office@imeta.science
imetaomics@imeta.science

Promotion Video

iMetaScience

iMetaScience