Artificial Intelligence-Driven Anticancer Peptide Discovery

Junrui Wu[1#], Shuaiqi Ji[1#], Kashif Iqbal Sahibzada[2,3,4#], Mengxue Lou[1],
Feiyu An[1], Wenqian Li[1], Jiawei Guo[1], Taowei Zhang[1], Xinyi Zhang[1],
Yilin Chou[1], Henan Zhang[1], Hao Jin[5], Teng Ma[5], Weichi Liu[5], Begali Alikulov[6],
Natalia Alekseevna Golovneva[7], Hooi Ling Foo[8,9], Issayeva Kurala[10],
Zhihong Sun[5]*, Dongqing Wei[4]*, Rina Wu[1]*

1College of Food Science, Shenyang Agricultural University,
National Agricultural Environmental Microbial Germplasm Resource Bank,
Liaoning Engineering Research Center of Food Fermentation Technology,
Shenyang Key Laboratory of Microbial Fermentation Technology Innovation, Shenyang 110866, PR China
2College of Biological Engineering, Henan University of Technology, Zhengzhou 450001, PR China
3Department of Health Professional Technologies, Faculty of Allied Health Sciences,
The University of Lahore, Lahore 54570, Pakistan
4State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of
Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology,
Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, PR China
5Key Laboratory of Dairy Biotechnology and Engineering, Ministry of Education,
Inner Mongolia Agricultural University, Hohhot 010018, PR China
6Department of Genetics and Biotechnology, Samarkand State University,
140104, Samarkand, Uzbekistan
7Institute of Microbiology of the National Academy of Sciences of Belarus,
ul. Kuprevicha, 2, Minsk, 220141, Belarus
8Department of Bioprocess Technology, Faculty of Biotechnology and Biomolecular Sciences,
Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia
9Lactic Acid Bacteria Biota Technology Research Program,
Research Laboratory of Probiotics and Cancer Therapeutics,
UPM-MAKNA Cancer Research Laboratory (CANRES),
Institute of Bioscience, Universiti Putra Malaysia,
43400 UPM Serdang, Selangor, Malaysia
10Department of Biotechnology, Toraighyrov University, Pavlodar, 140008, Kazakhstan

Junrui Wu, Shuaiqi Ji, Kashif Iqbal Sahibzada, Mengxue Lou, Feiyu An, Wenqian Li, Jiawei Guo, et al. 2025.
Artificial Intelligence-Driven Anticancer Peptide Discovery. ***iMetaOmics*** 2: e70063. https://doi.org/10.1002/imo2.70063

Introduction

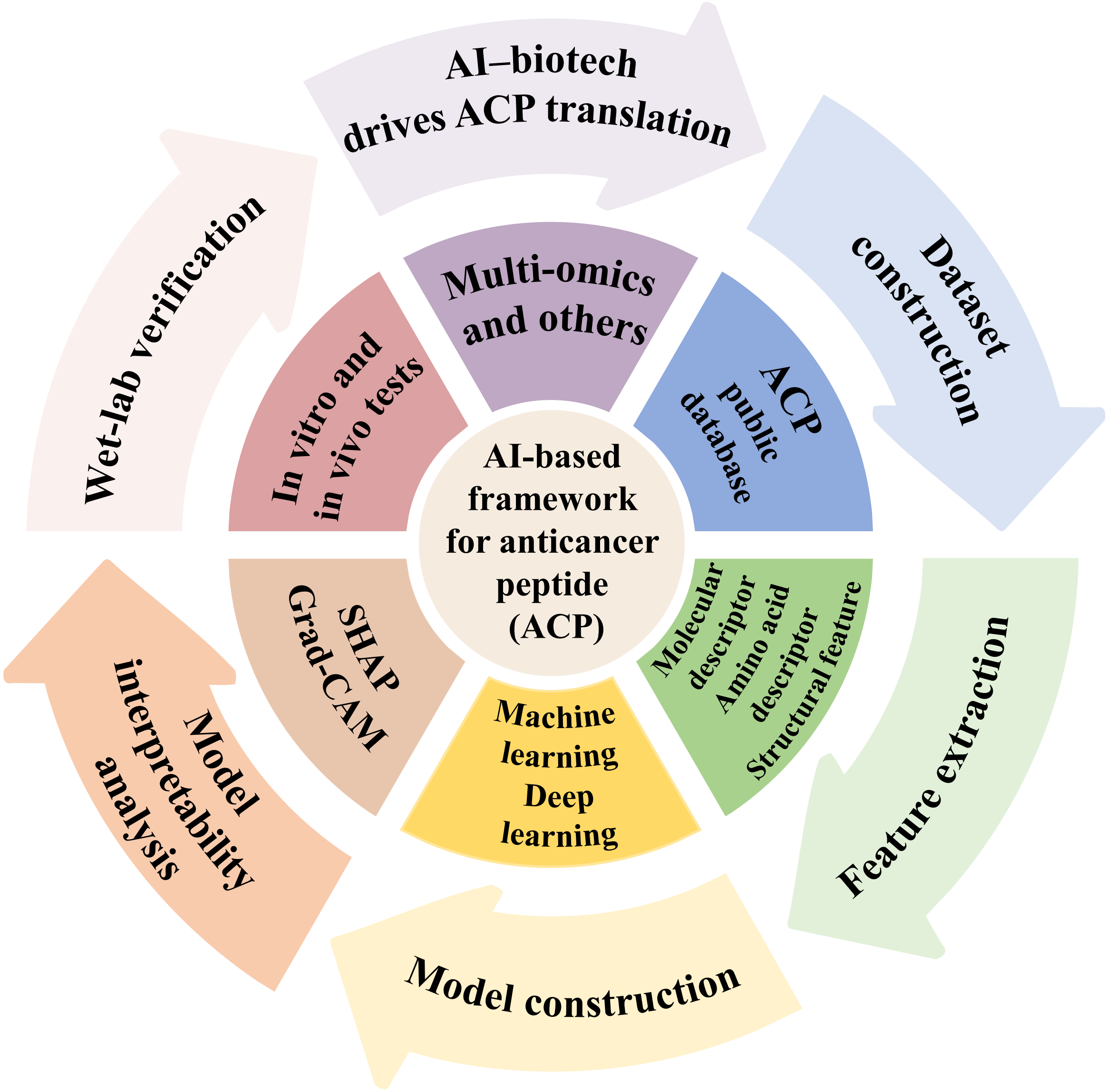


Table S1. Summary of 68 ACP prediction models

Model name	Dataset size	Feature extraction method	Model	Interpretability
-	Positive: 650 Negative: 650	one-hot encoding physicochemical properties	2D CNN	-
G-ACP	Positive: 468 Negative: 600	PAAC	Extra Trees (ET)	-
-	Positive: 1151 Negative: 1139	Extended Dipeptide Composition (EDPC)	SVM Decision Tree RF KNN	-
ACP_DA	Positive: 859 Negative: 859	Ordinal Encoding (OE) AAC DPC Grouped Tripeptide Encoding	Light Gradient Boosting Machine (LGBM)	SHAP
ACPScanner	Positive: 551 Negative: 150	SVMProt-188D NetSurfP-3.0 ESM-1b PEP-FOLD3	LGBM-GAT	SHAP

Highlights

1. Summarized 68 anticancer peptide (ACP) models.

2. Artificial Intelligence (AI) interpretability and wet-laboratory validation enhance ACP mining trustworthiness.

3. AI combined with multi-omics and biotechnologies drives ACP clinical translation.

4. Summarized a comprehensive AI-driven ACP screening framework.

The establishment process of the ACP AI model

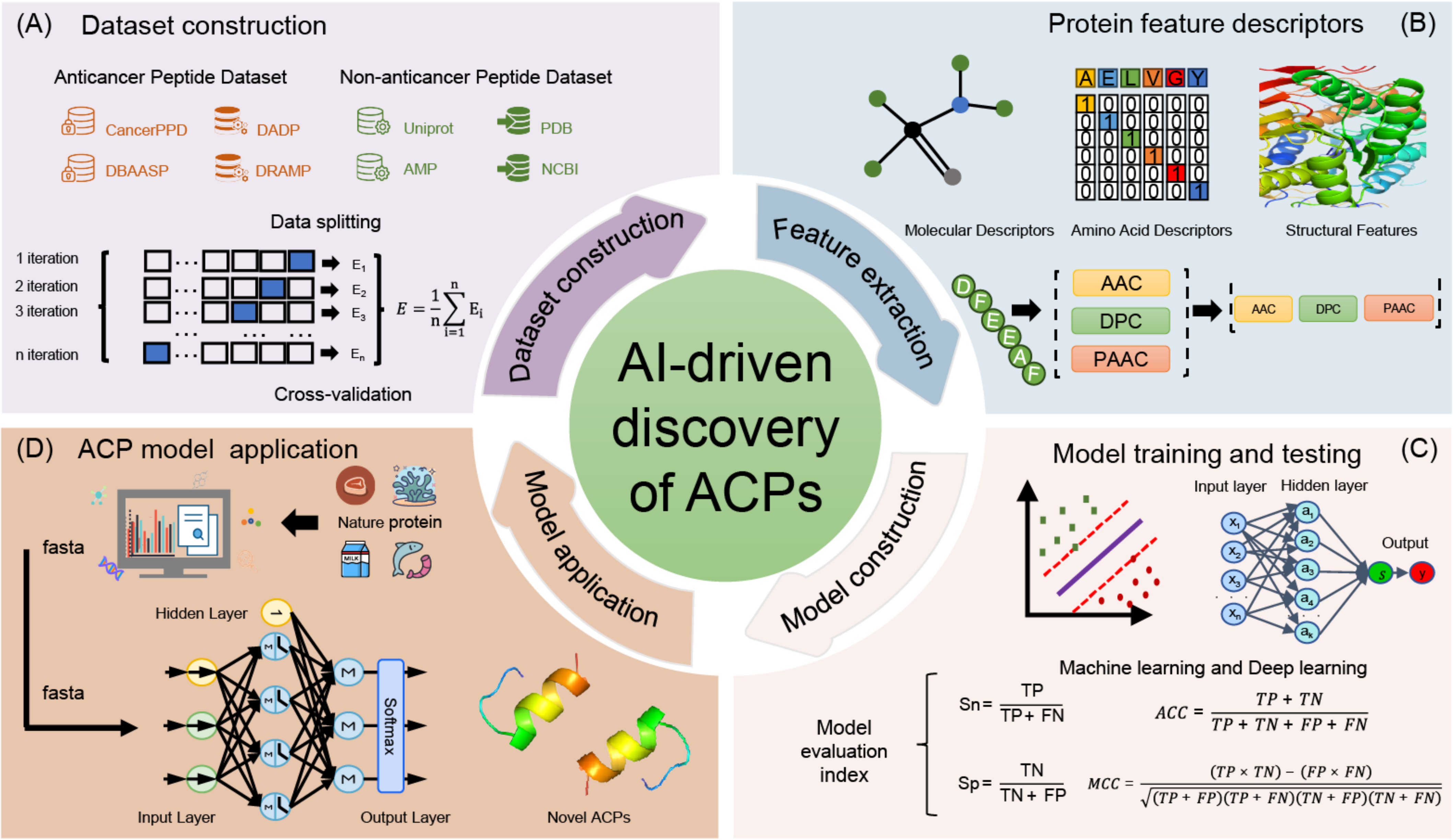


Figure 1. The establishment process of the ACP AI model.

Examples of commonly used traditional ML algorithms for ACP models

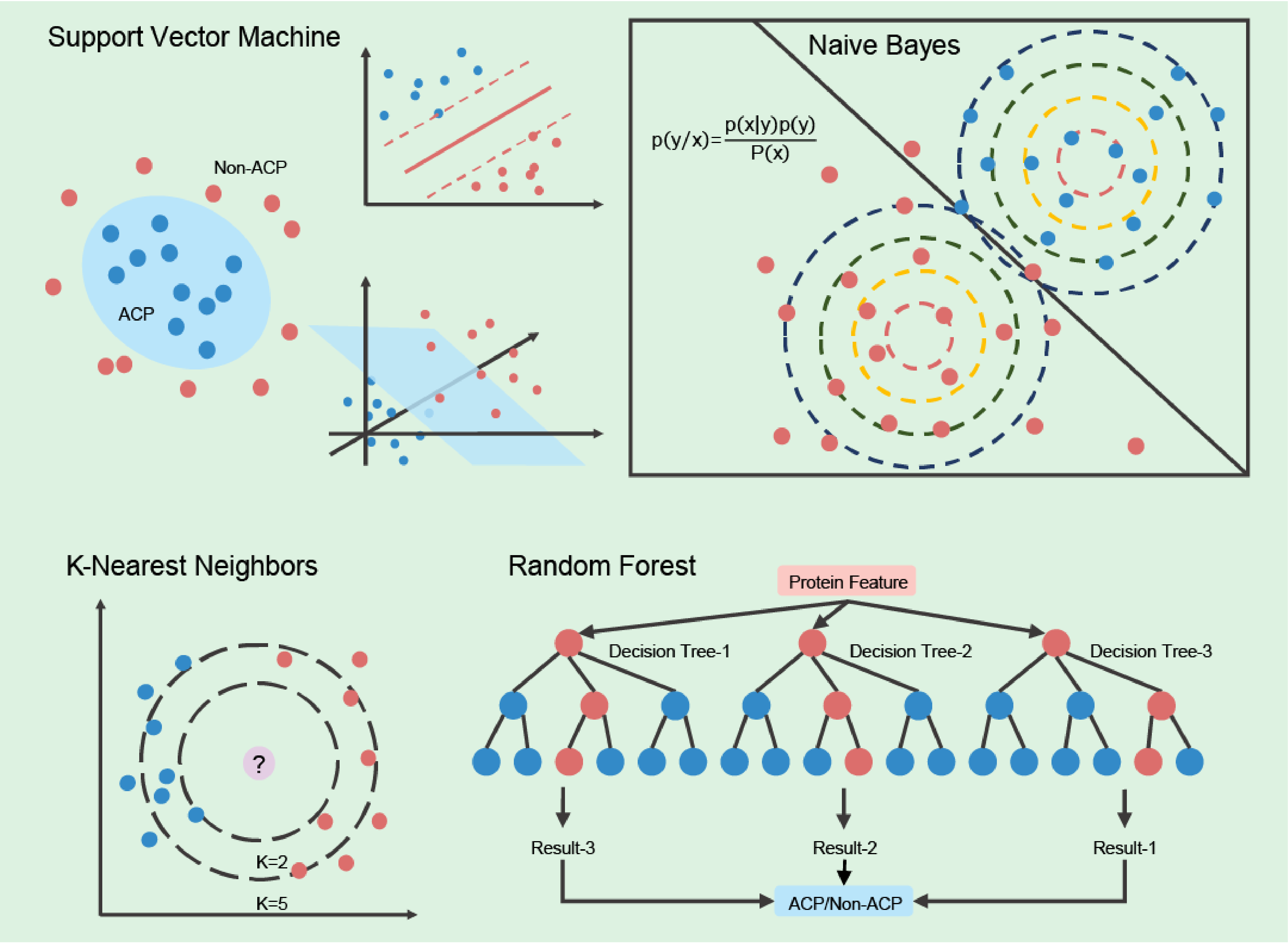


Figure 2. Examples of commonly used traditional ML algorithms for ACP models

Examples of commonly used deep learning algorithms for ACP models

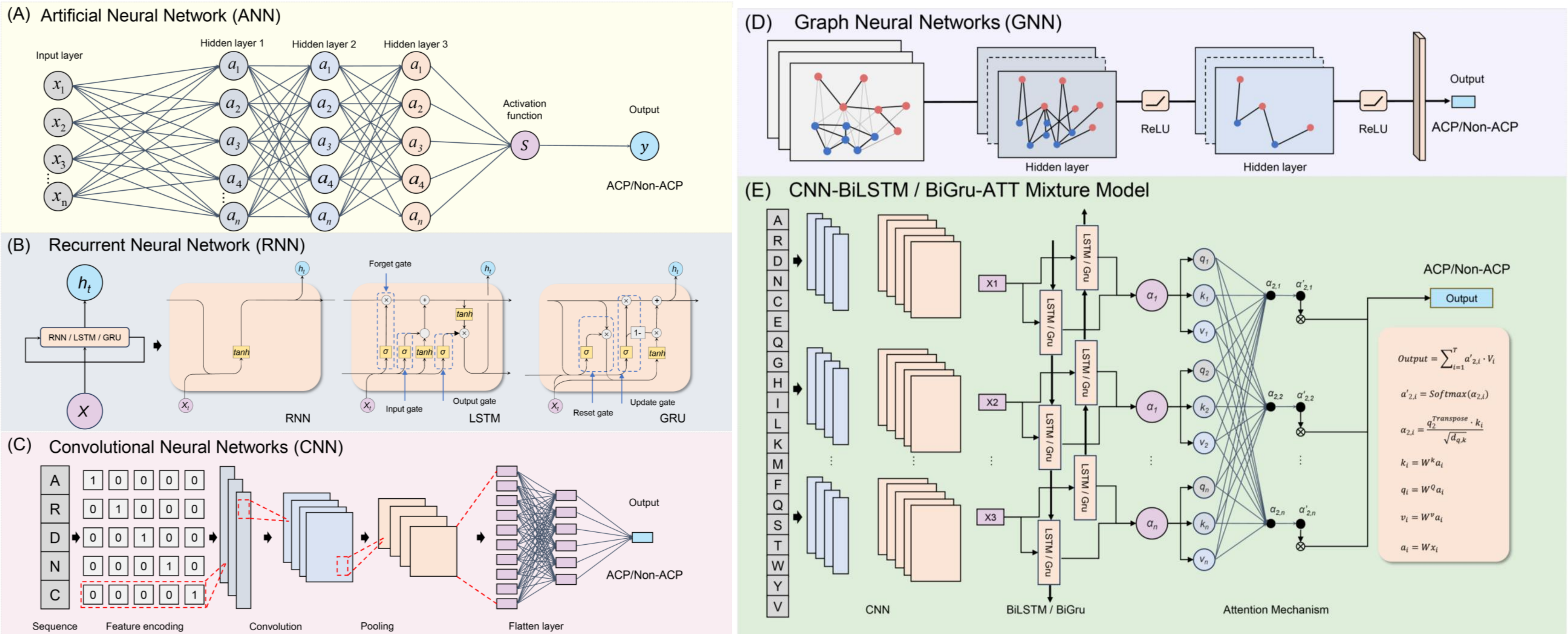


Figure 3. Examples of commonly used deep learning algorithms for ACP models

Schematic of AI interpretability analysis

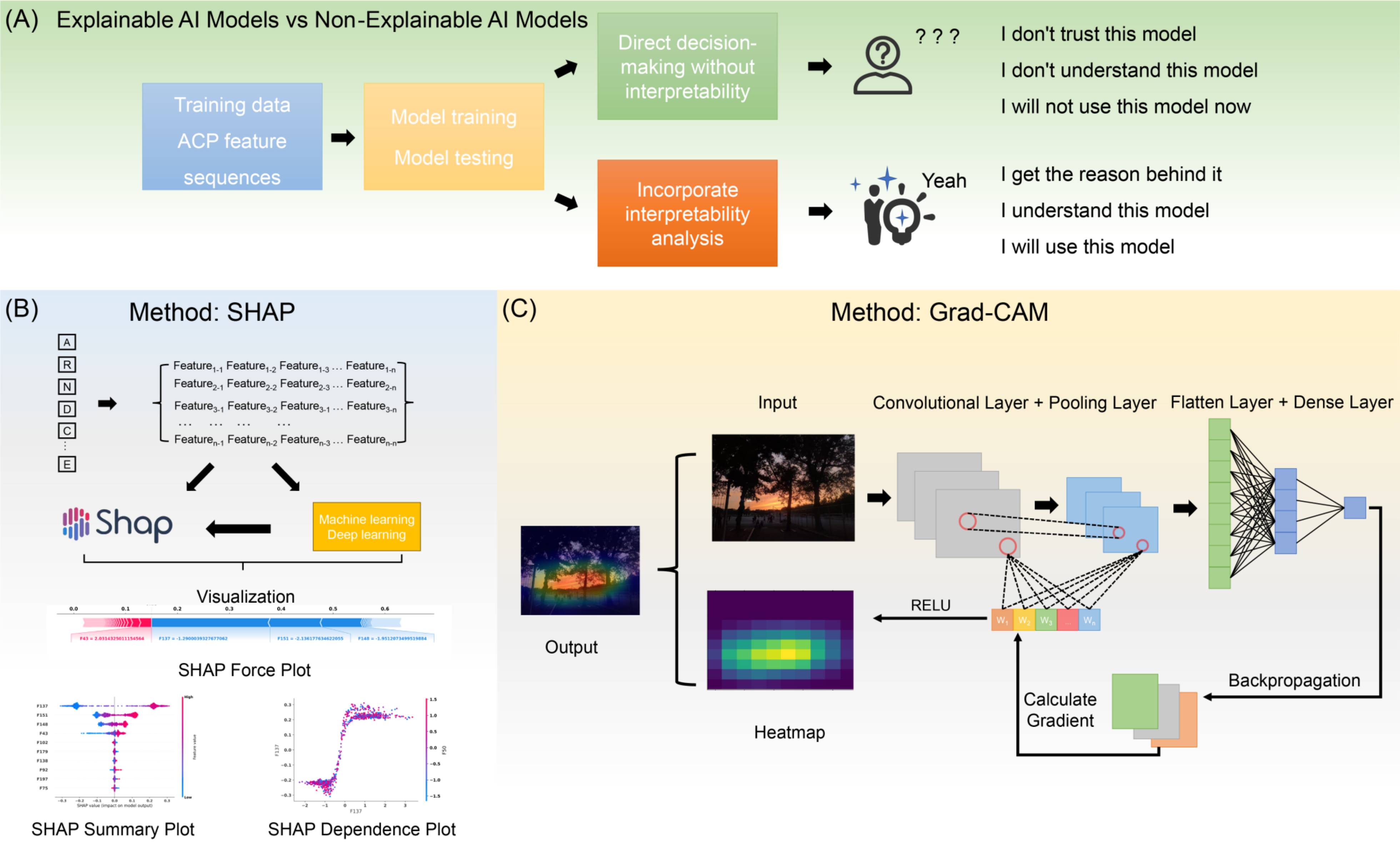


Figure 4. Schematic of AI interpretability analysis

Integrating AI-selected peptides with multi-omics technologies and other biotechnologies

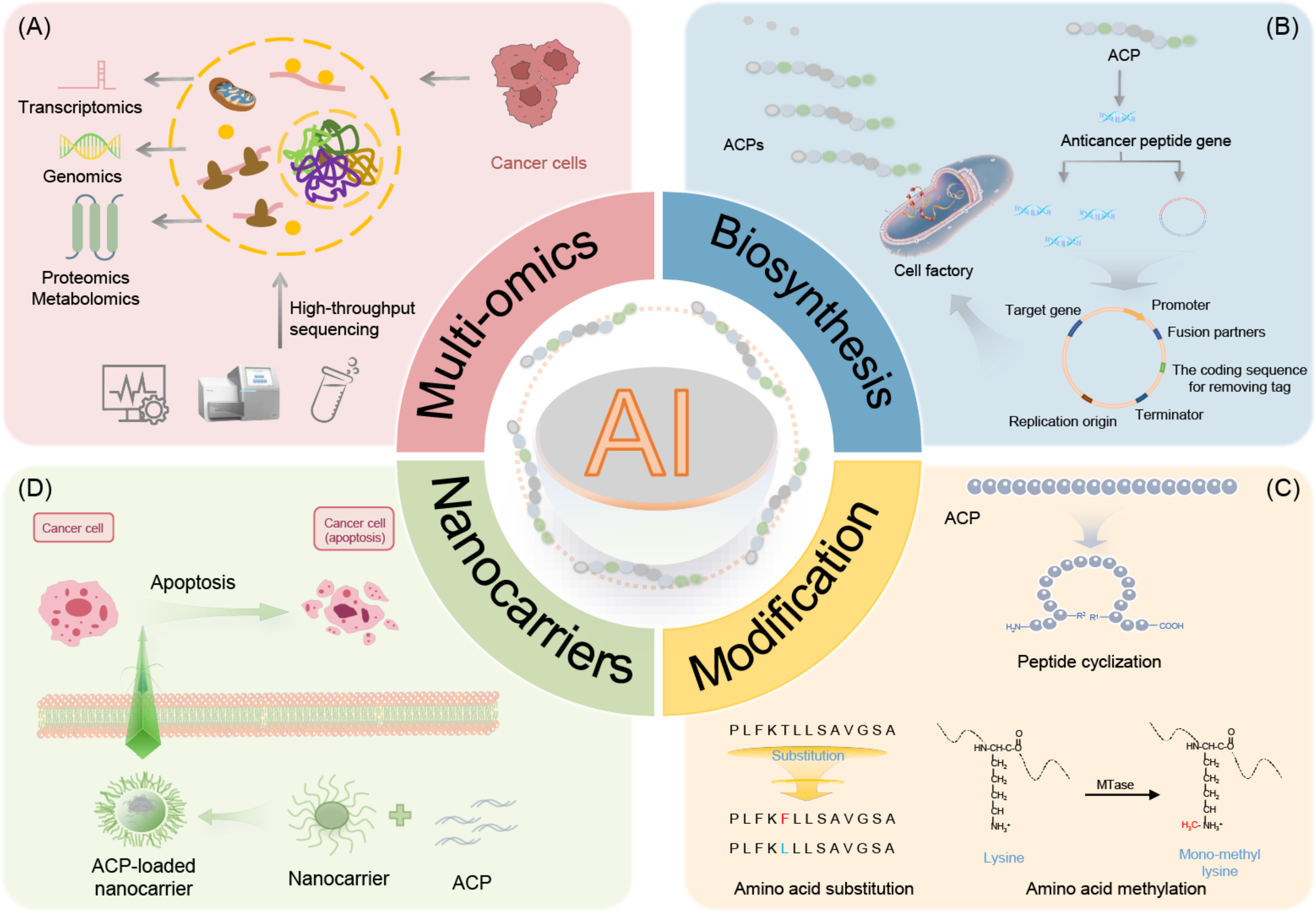


Figure 5. Integrating AI-selected peptides with multi-omics technologies and other biotechnologies

Comprehensive AI-based ACP screening framework

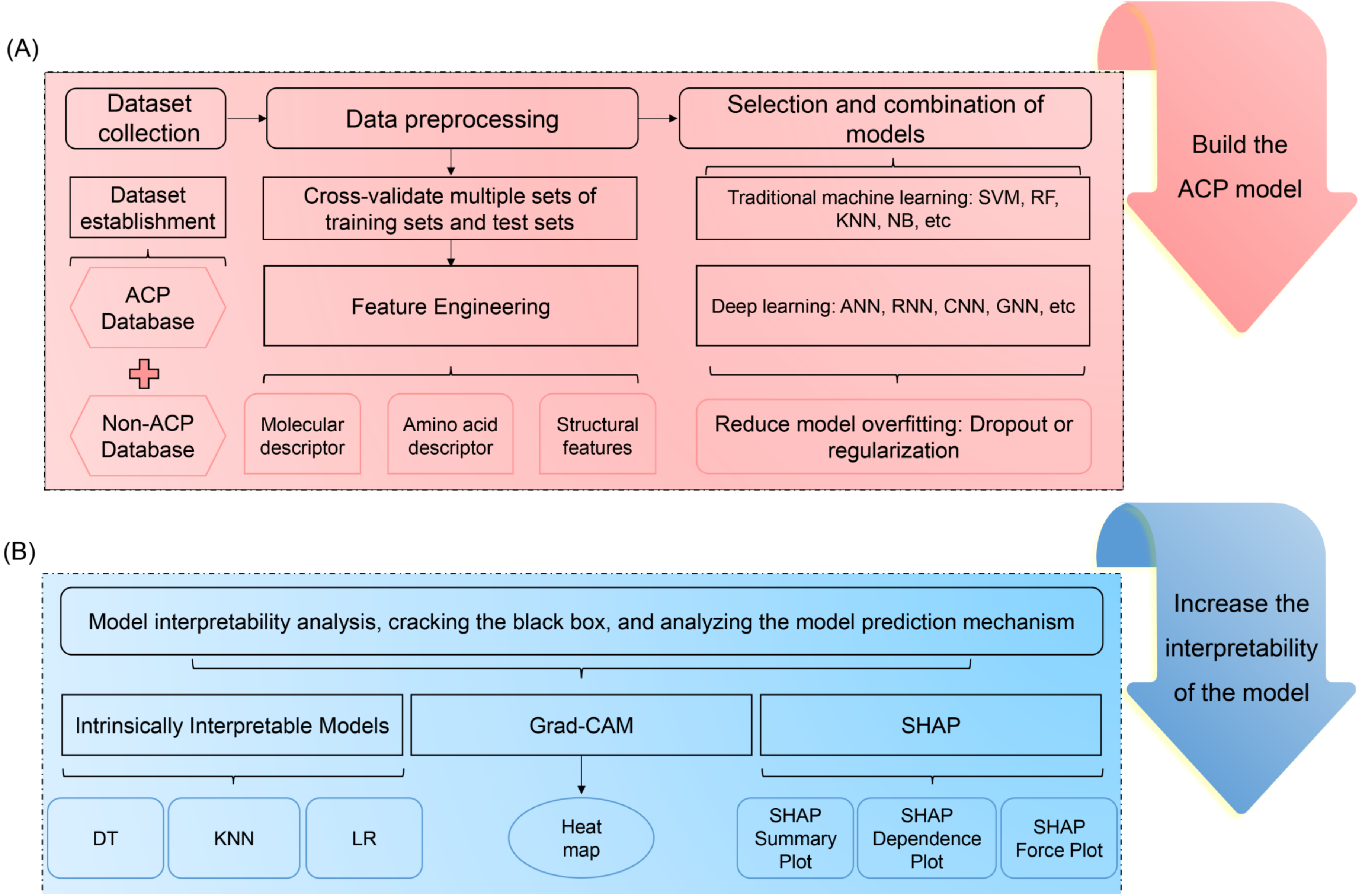


Figure 6. Illustrative diagram of the AI-driven comprehensive framework for ACP screening

Comprehensive AI-based ACP screening framework

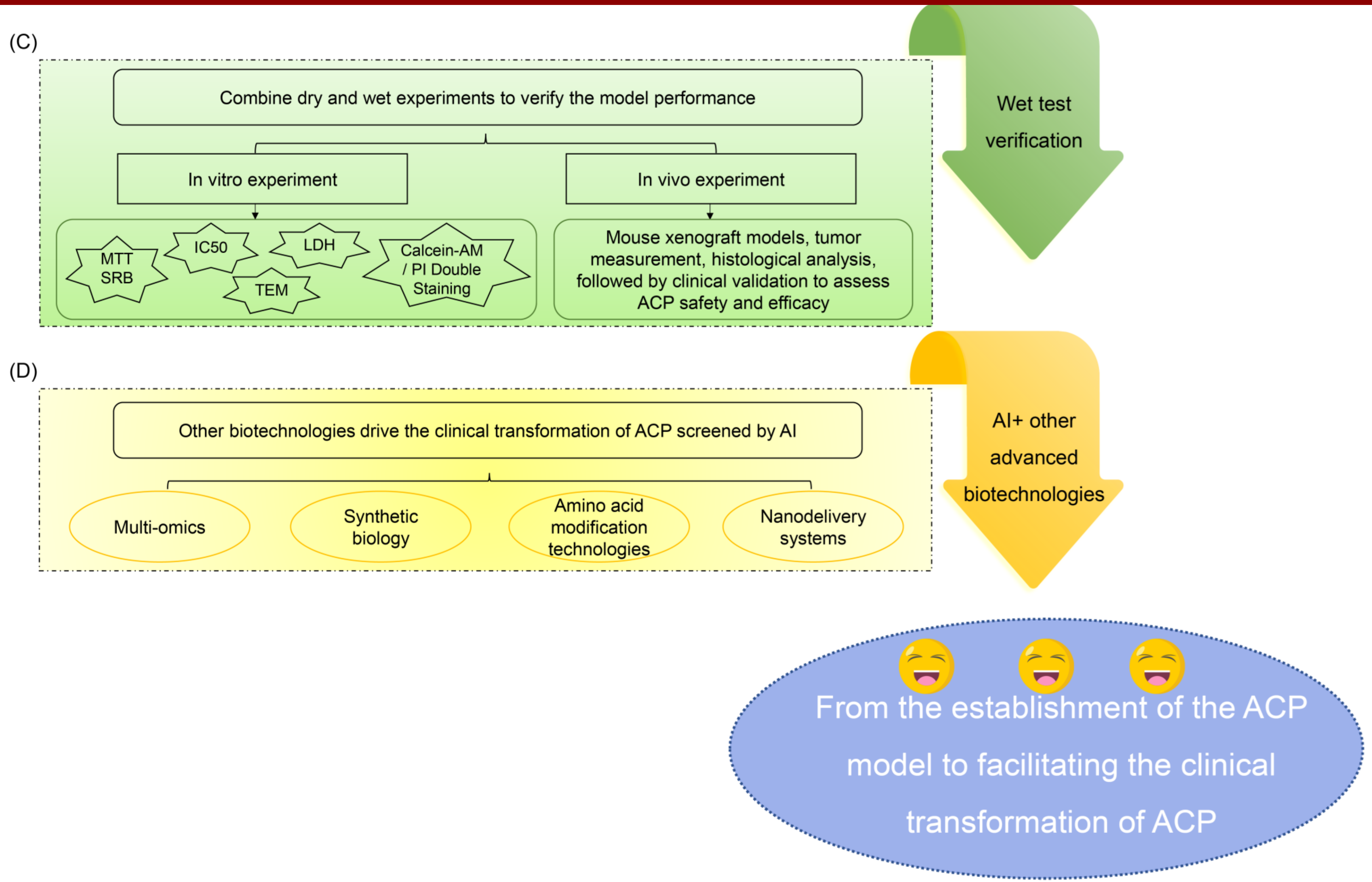


Figure 6. Illustrative diagram of the AI-driven comprehensive framework for ACP screening

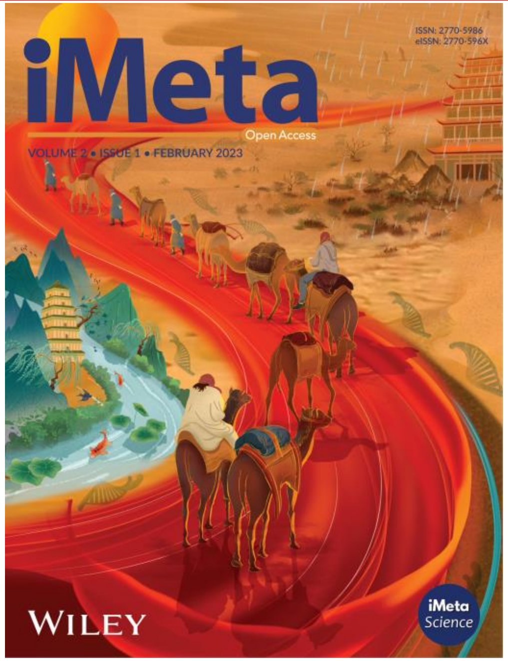
Summary

- To outline a complete AI-based ACP screening framework, we detailed the summary of 68 ACP models. Based on the training processes and limitations of these 68 models, we presented a comprehensive ACP screening framework that includes dataset collection and organization, feature extraction, model training and optimization, interpretability analysis, and wet-laboratory validation.
- Additionally, this framework integrates multi-omics, synthetic biology, amino acid modification technologies, and nanocarrier technologies to facilitate the process from model development to clinical application. These technologies should be integrated to increase the biological activity, structural stability, and manufacturability of AI-selected ACPs.
- In conclusion, AI-driven ACP discovery holds great potential. It can expedite peptide screening and functional prediction, significantly enhancing the efficiency and accuracy of novel peptide-based anticancer drug development, thereby offering more diverse and effective treatment options for cancer.

Junrui Wu, Shuaiqi Ji, Kashif Iqbal Sahibzada, Mengxue Lou, Feiyu An, Wenqian Li, Jiawei Guo, et al. 2025.
Artificial Intelligence-Driven Anticancer Peptide Discovery. ***iMetaOmics*** 2: e70063. https://doi.org/10.1002/imo2.70063

iMeta: To be top journals in biology and medicine

WILEY



"***iMeta***" launched in 2022 by iMeta Science Society, impact factor (IF) **33.2**, ranking top 65/22249 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 *Cell*,** *Nature Biotechnology/Methods/Microbiology/Medicine/Food*. Its unique features include video abstract, bilingual publication, and social media with 600,000 followers. Indexed by **SCIE/ESI**, **PubMed**, **Google Scholar** etc.

"***iMetaOmics***" launched in 2024, 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.

"***iMetaMed***" launched in 2025, with a **target IF>15, similar to *Med, Cell Reports Medicine***, *eBioMedicine, eClinicalMedicine* etc.

Society: http://www.imeta.science

Publisher: https://wileyonlinelibrary.com/journal/imeta

Submission:
- iMeta: https://wiley.atyponrex.com/journal/IMT2
- iMetaOmics: https://wiley.atyponrex.com/journal/IMO2
- iMetaMed: https://wiley.atyponrex.com/journal/IMM3

iMetaScience

iMetaScience

office@imeta.science

imetaomics@imeta.science

Promotion Video

Update 2025/7/6