

MINIMALLY INVASIVE MULTI-OMICS BIOMARKER DISCOVERY

Best Practices and Pitfalls

Klemens Vierlinger

Workshop on Omics Approaches for Biomarker Discovery

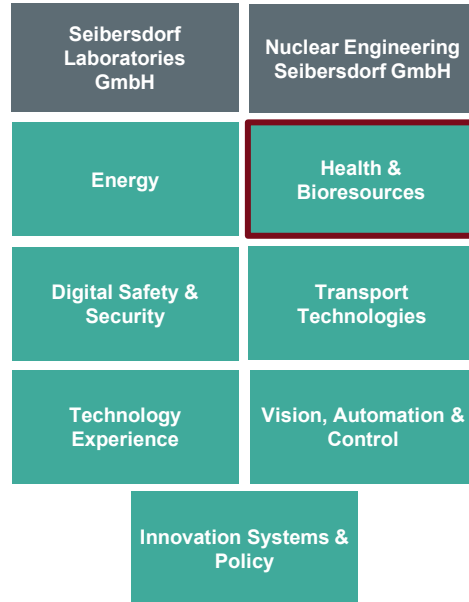
24 September 2026

Institute of Biochemistry and Molecular Genetics

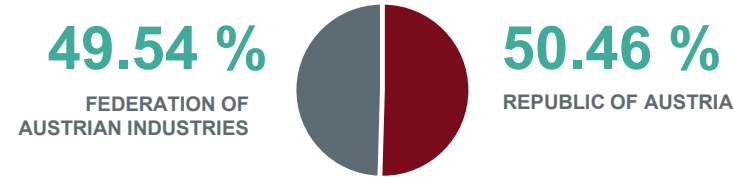
Faculty of Medicine, University of Ljubljana, Slovenia



AIT AUSTRIAN INSTITUTE OF TECHNOLOGY



OWNERSHIP STRUCTURE



TOTAL REVENUES

~ 218 m EUR

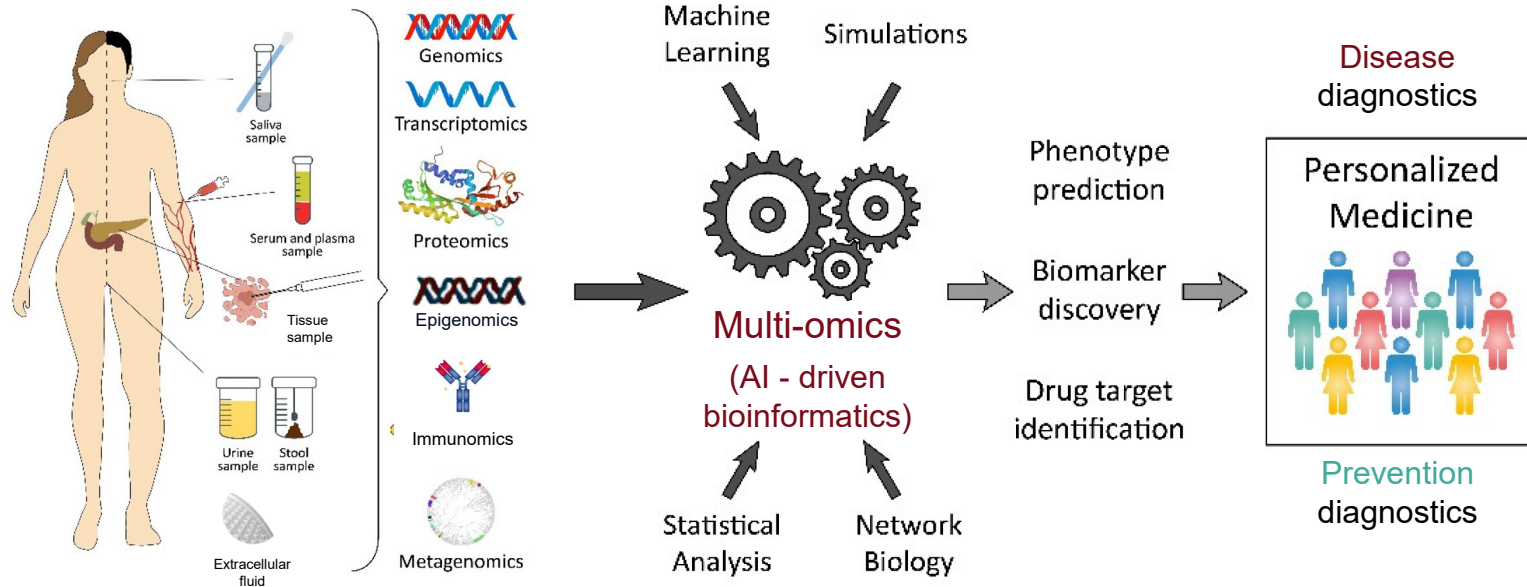
EMPLOYEES

> 1.650
employees



MOLECULAR DIAGNOSTICS at AIT

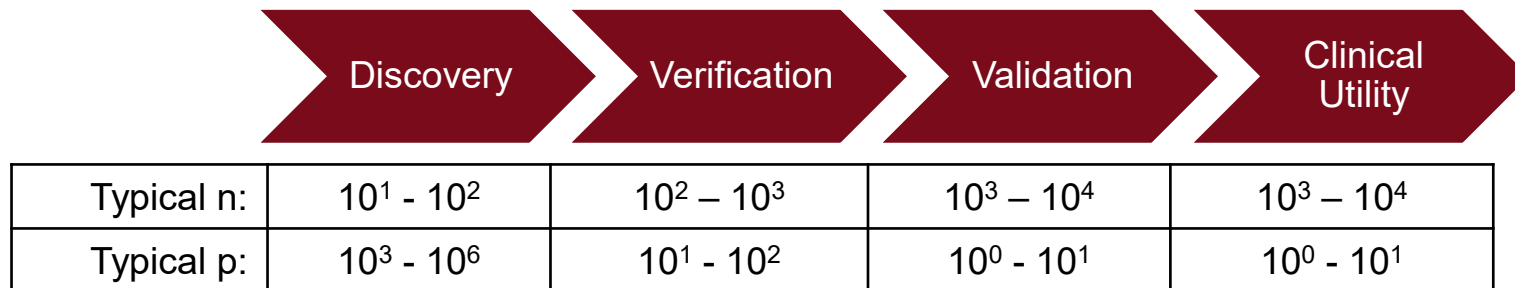
Minimal or non-invasive molecular diagnostic / Liquid biopsy



AGENDA

- 1) Pitfalls in Biomarker Discovery
- 2) Minimally invasive Biomarker Discovery
 - 1) Antibodies as Biomarkers
 - 2) Circulating RNAs as Biomarkers

BIOMARKER DEVELOPMENT



Only ~10% of biomarkers advance from discovery to validation, with ~90% failing due to technical or design flaws.

BEST (Biomarkers, EndpointS, and other Tools)
Resource, FDA-NIH Biomarker Working Group

Critical Steps:

- **Discovery**
- Tech transfer: NGS -> qPCR, MS -> ELISA
- Matrix transfer: tissue -> minimally invasive matrices
- Replication in Independent Validation Cohorts

PITFALL #1: INADEQUATE SAMPLE SIZE

-> Leads to false positives / false negatives -> failure in independent validation

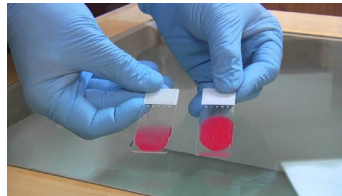
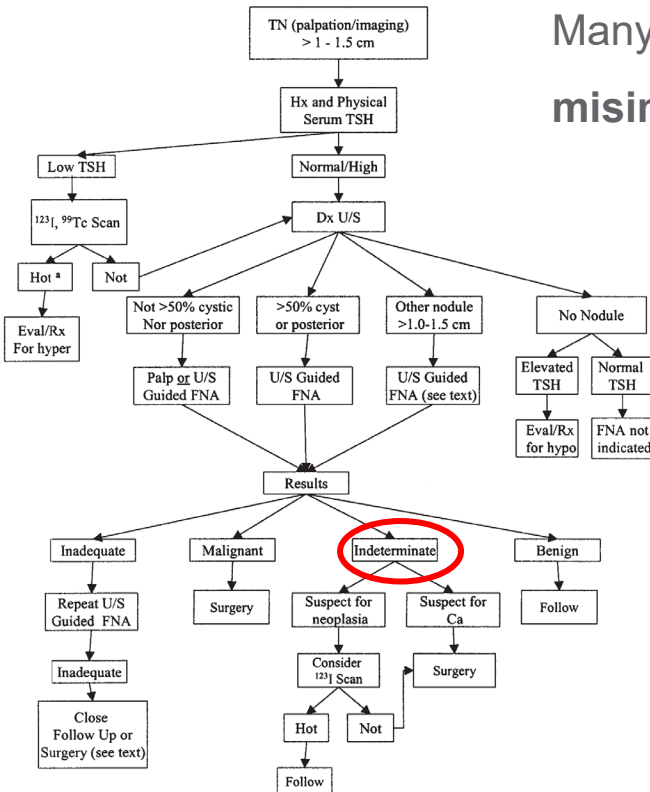
- Pilot data often missing / scarce
- True effect sizes often overestimated
- Disease heterogeneity often underestimated
- Budget constraints
- Samples availability
- Particularly relevant when feature space is large (e.g. DNA methylation)



- Use open data as pilot data (with caution)
- Two-stage design (pilot -> discovery cohort)
- Large consortia often necessary
- Use suitable, platform-specific sample size calculation tools:
 - OCplus
 - RnaSeqSampleSize
 - RNASeqPower
 - OLINK Insight sample size calculator

PITFALL #2: INADEQUATE DEFINITION OF USE CASE / TARGET POPULATION

Many studies fail to **clearly define** their target population, leading to **misinterpretation, poor reproducibility, and limited clinical utility.**



Bethesda Class	Diagnostic Category	ROM [%]	Management
I	Nondiagnostic or Unsatisfactory	1-4	Repeat FNA with ultrasound guidance
II	Benign	0-3	Clinical follow-up
III	AUS/FLUS	5-15	Repeat FNA
IV	Follicular Neoplasm	15-30	Surgical lobectomy
V	Suspicious for Malignancy	60-75	Near-total thyroidectomy or surgical lobectomy
VI	Malignant	97-99	Near-total thyroidectomy

Cibas et.al. 2009

PITFALL #2: INADEQUATE DEFINITION OF USE CASE / TARGET POPULATION

Considerations for designing a discovery cohort

- **Be specific:** E.g., patients with stage XX disease, which are eligible for treatment YY according to guidelines ZZ instead of "patients with cancer"
- **Prevalence** of positive cases in target population?
- Are the **results actionable**?
- Which sample **matrices** are **available** in a routine clinical setting?
- **Define clinical context:** consider comorbidities, prior treatments, disease progression.
- Define suitable **control group**
- **Consider demographic factors:** Age, sex, ethnicity, and genetic background are often overlooked.

PITFALL #3: MATRIX TRANSFER BIAS

Problem: Tissue biomarkers don't translate well into circulation

- **Low Abundance in Circulation:** Tissue biomarkers are **diluted** in blood ($1:10^6$ – $1:10^9$ ratio).
- **Lack of Specificity** Circulating biomarkers may **originate from multiple tissues**, not just the target disease.
- **Biological Barriers** Tissue barriers (e.g., blood-brain barrier, extracellular matrix) prevent biomarker release.



- Avoid matrix transfer
- Select highly specific signals (e.g. tumor mutation)
- Use highly sensitive detection methods (e.g. ddPCR, PEA)

PITFALL #3: TECHNOLOGY TRANSFER

Failure rate in Translation from Discovery technology to Biomarker assay

- Preamanalytical factors
- Antibody availability (Proteomics)
- qPCR assay design (e.g. in GC-rich regions like CpG islands, short RNA fragments)
- Selection of suitable control / standardisation assays

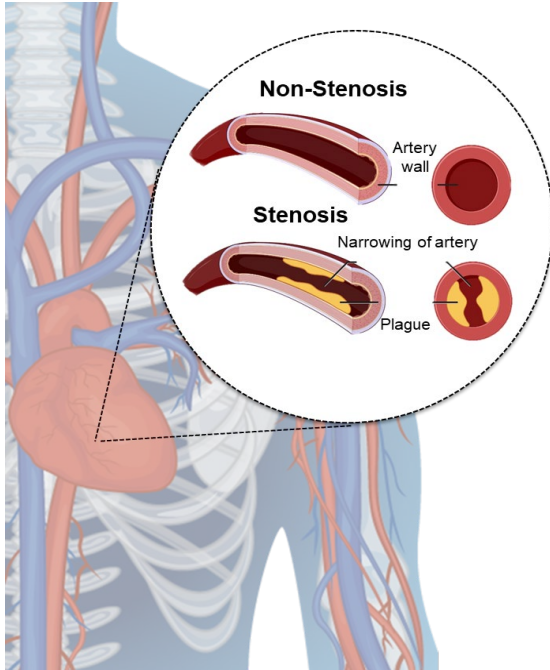
SUMMARY: RECOMMENDATIONS FOR BIOMARKER DISCOVERY

- Allow for adequate sample size - Be realistic in your expectations
- Choose your discovery matrix and omics layer wisely
 - Dimensionality
 - Need for matrix transfer
 - Affordability
 - Biological plausibility
- Select redundant features from discovery
- Involve industry early
- Consider regulatory aspects early (e.g. risk classification)

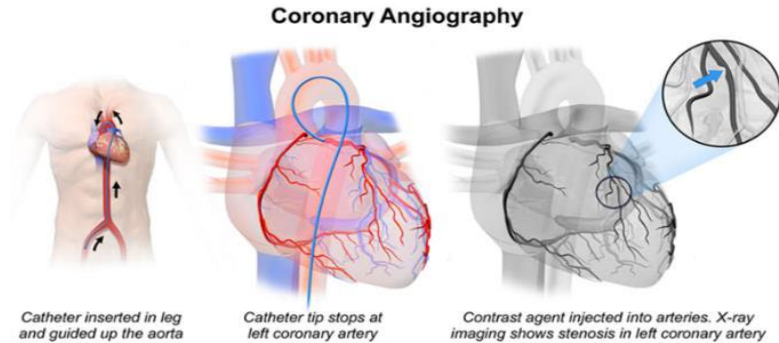
MINIMALLY INVASIVE BIOMARKER DISCOVERY

For risk stratification in cardiovascular disease





- Coronary artery disease (CAD) is a complex chronic inflammatory disease caused by lipid build up in arteries = atherosclerosis
- Gold standard diagnostic test = Coronary Angiography (CA)

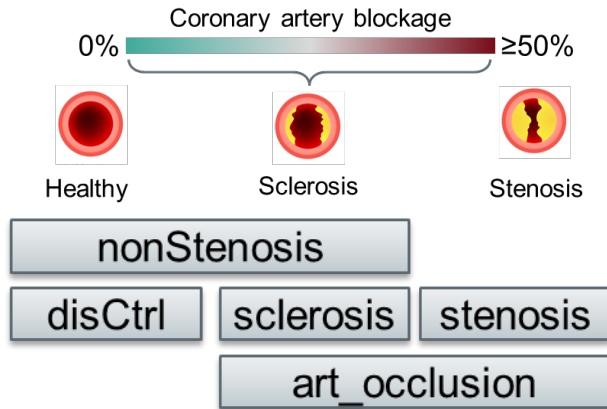


- 40% of patients undergoing CA due to indicative symptoms show normal-appearing coronary arteries (no stenosis)

Need for non-invasive, radiation-exposure free diagnostic test

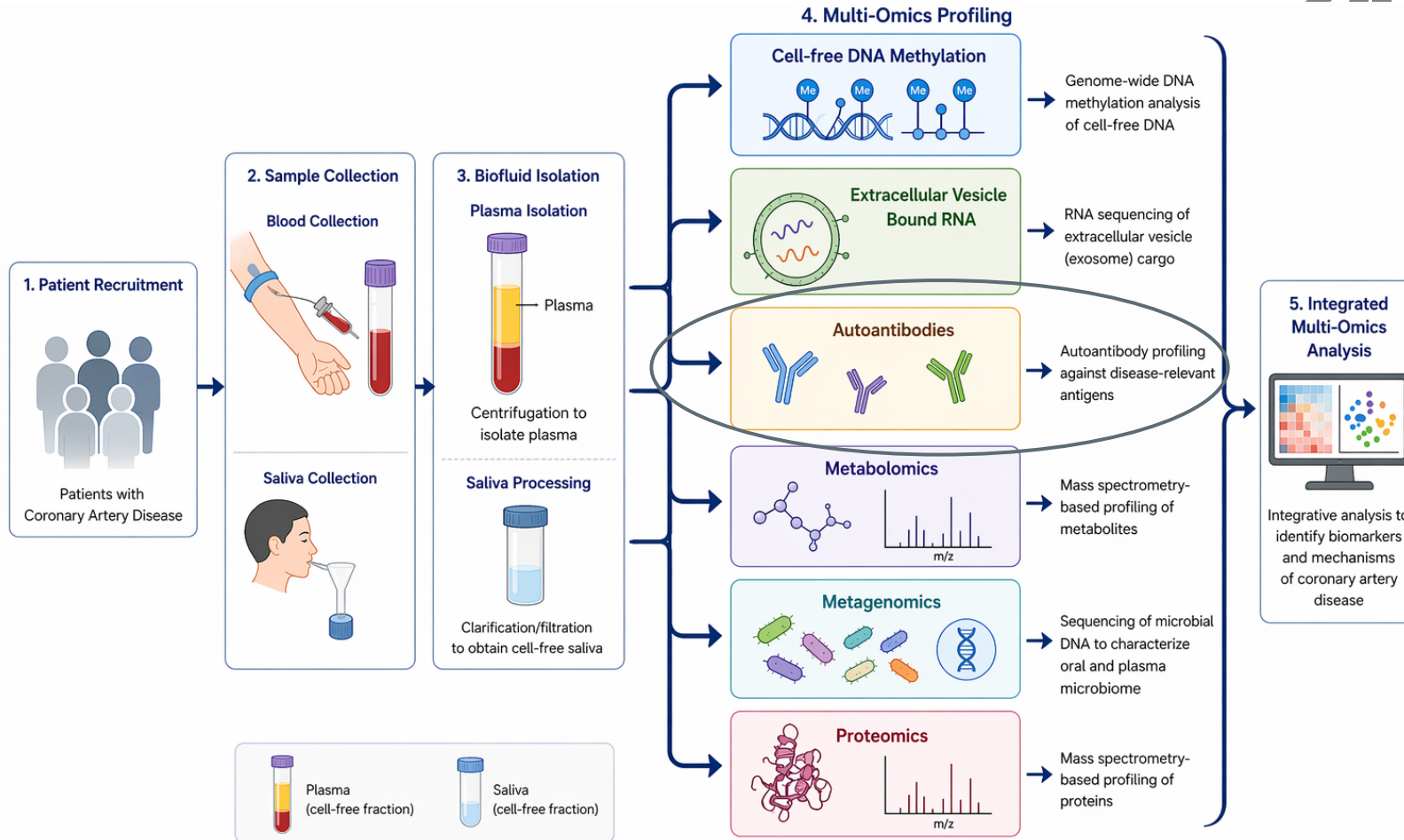
STENOSIS COHORT

- Pseudoprospective (sampling before angiography)
- 09/2020 – 09/2021
- Rich data
 - Personal
 - Clinical
 - Pharmacological
 - Phenotypical



	disCtrl (N=37)	sclerosis (N=32)	stenosis (N=77)	Total (N=146)	p value	
age					0.299	
N-Miss	0	0	2	2		
Mean (SD)	62.757 (15.394)	67.031 (10.751)	65.160 (9.320)	64.958 (11.485)		
Range	0.000 - 82.000	41.000 - 81.000	41.000 - 83.000	0.000 - 83.000		
sex					< 0.001	←
m	13 (35.1%)	21 (65.6%)	58 (75.3%)	92 (63.0%)		
w	24 (64.9%)	11 (34.4%)	19 (24.7%)	54 (37.0%)		
family_history					0.063	←
N-Miss	5	1	6	12		
Mean (SD)	0.250 (0.440)	0.194 (0.402)	0.408 (0.495)	0.321 (0.469)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
BMI					0.266	
N-Miss	9	7	20	36		
Mean (SD)	30.847 (6.498)	29.834 (6.697)	28.783 (4.396)	29.547 (5.570)		
Range	21.338 - 44.922	23.147 - 53.146	20.797 - 38.874	20.797 - 53.146		
smoking					0.049	←
active	5 (13.5%)	7 (21.9%)	19 (24.7%)	31 (21.2%)		
former	5 (13.5%)	7 (21.9%)	25 (32.5%)	37 (25.3%)		
no	27 (73.0%)	18 (56.2%)	33 (42.9%)	78 (53.4%)		
typical_symptoms					0.035	←
Mean (SD)	0.270 (0.450)	0.250 (0.440)	0.468 (0.502)	0.370 (0.484)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
atypical_symptoms					0.275	
Mean (SD)	0.243 (0.435)	0.281 (0.457)	0.156 (0.365)	0.205 (0.405)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
diabetes					0.320	
Mean (SD)	0.216 (0.417)	0.188 (0.397)	0.312 (0.466)	0.260 (0.440)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
hyperlipidemia					0.762	
Mean (SD)	0.541 (0.505)	0.625 (0.492)	0.597 (0.494)	0.589 (0.494)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
hypertension					0.560	
Mean (SD)	0.730 (0.450)	0.781 (0.420)	0.818 (0.388)	0.788 (0.410)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
obesity					0.233	
Mean (SD)	0.351 (0.484)	0.219 (0.420)	0.208 (0.408)	0.247 (0.433)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
vorhofflimmern					0.617	
Mean (SD)	0.108 (0.315)	0.094 (0.296)	0.156 (0.365)	0.130 (0.338)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		
copd					0.076	←
Mean (SD)	0.081 (0.277)	0.188 (0.397)	0.052 (0.223)	0.089 (0.286)		
Range	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000	0.000 - 1.000		

OMICS LAYERS IN STENOSIS COHORT



Autoantibodies: Immune system proteins that mistakenly target the body's own tissues.

Biomarker potential: Autoantibodies can serve as early, non-invasive indicators of cardiovascular disease (CVD).

Why they matter:

- Detectable years before clinical symptoms.
- Reflect underlying pathological processes (e.g., cancer, inflammation, tissue damage).
- Stable in circulation, making them ideal for blood-based tests.

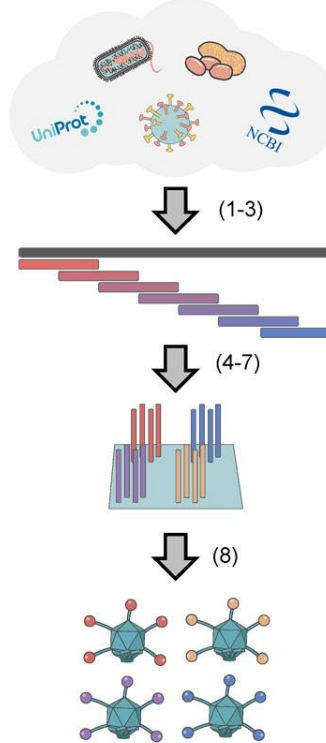
Examples in CVD:

- Anti-Phospholipid (aPL): Atherosclerosis – early detection
- Cardiac troponin I (cTnI): Myocardial Infarction (MI) risk stratification

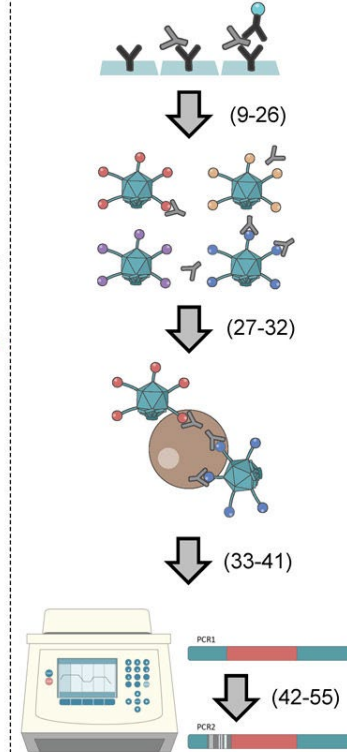
Phage Immuno Precipitation Sequencing

- Minimally invasive: serum, plasma, saliva, ...
- Small sample input: only a few ul needed
- Cost-effective for large cohorts
- Good translatability
- Limitation: only linear epitopes

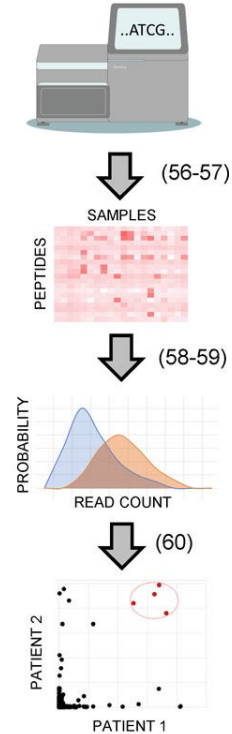
A. Create library



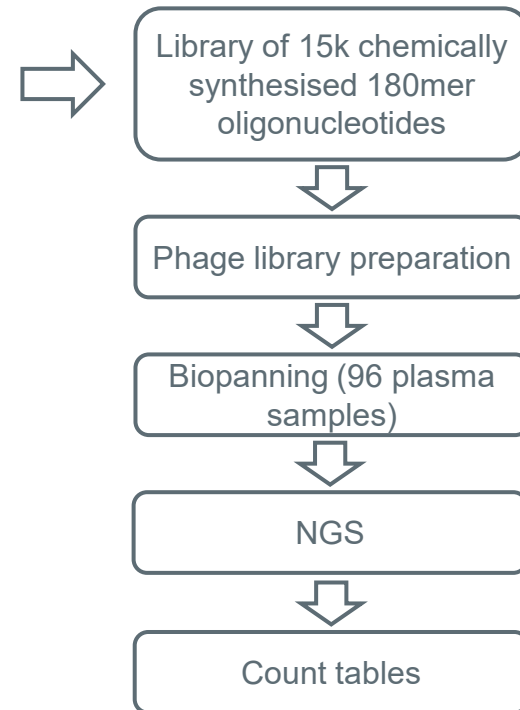
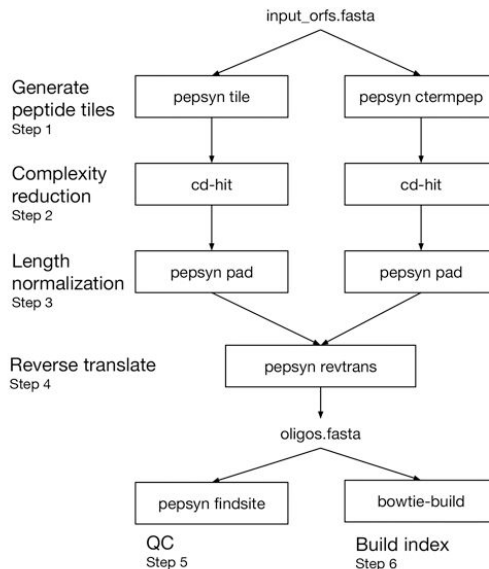
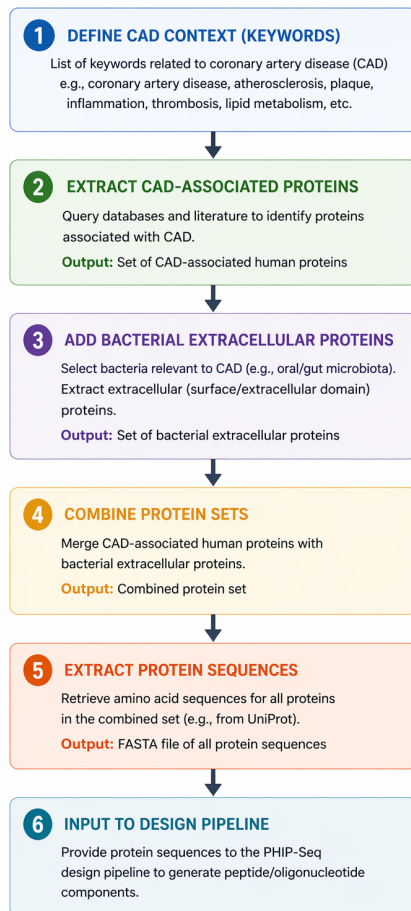
B. Screen serum



C. Analyze data

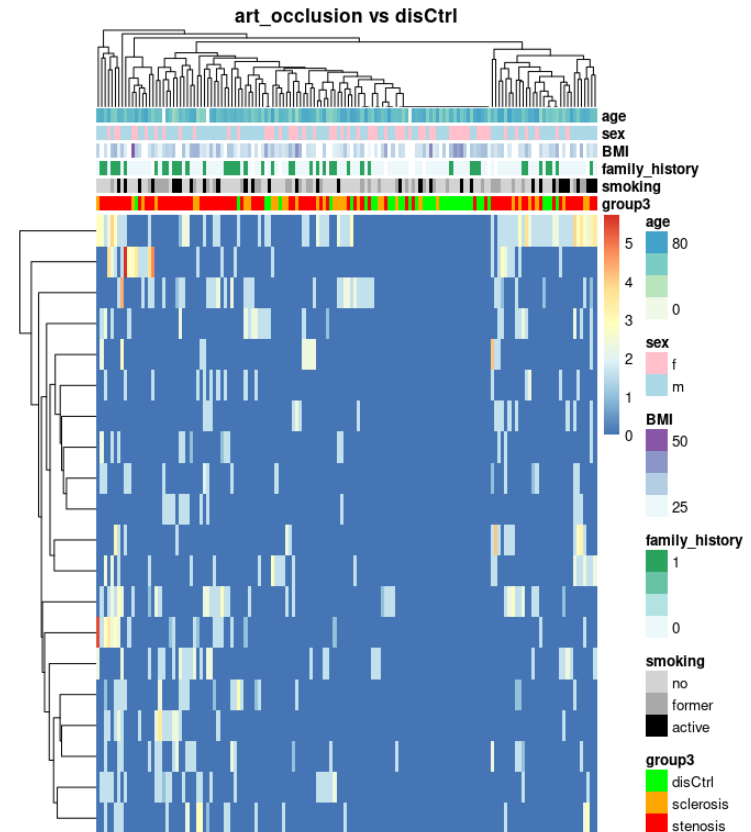
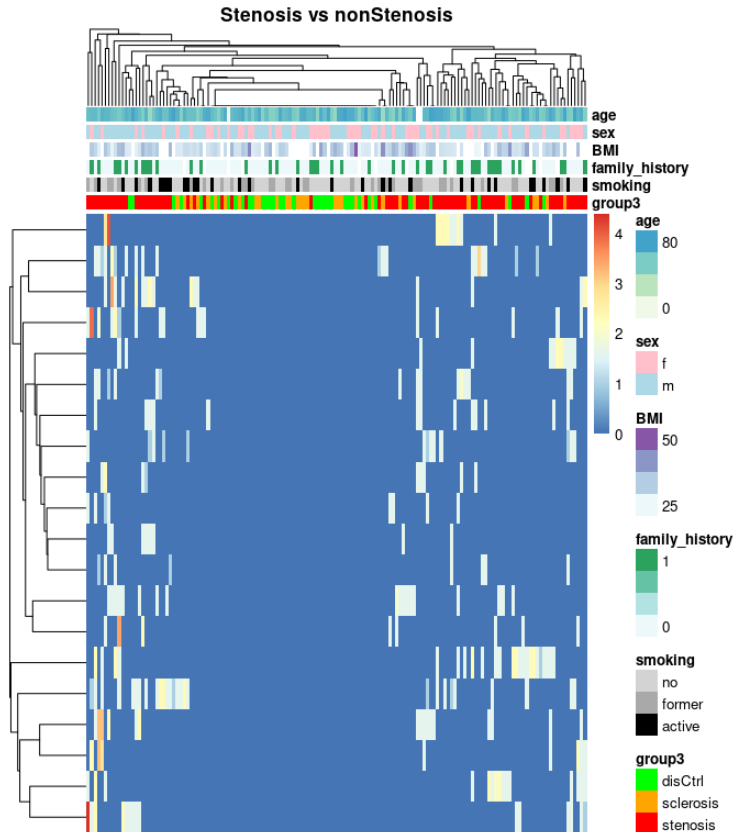


PHIPSEQ LIBRARY DESIGN



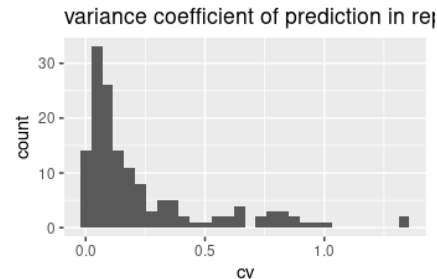
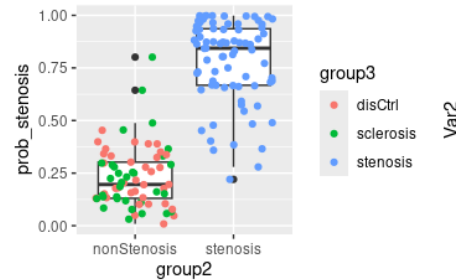
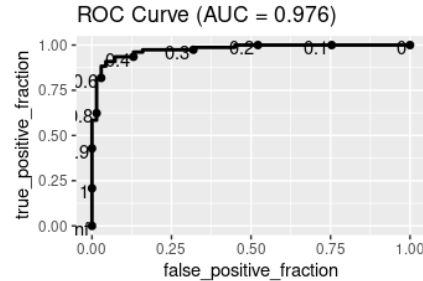
FEATURE SELECTION

By p value in Row-wise wilcoxon test (alternative="greater")

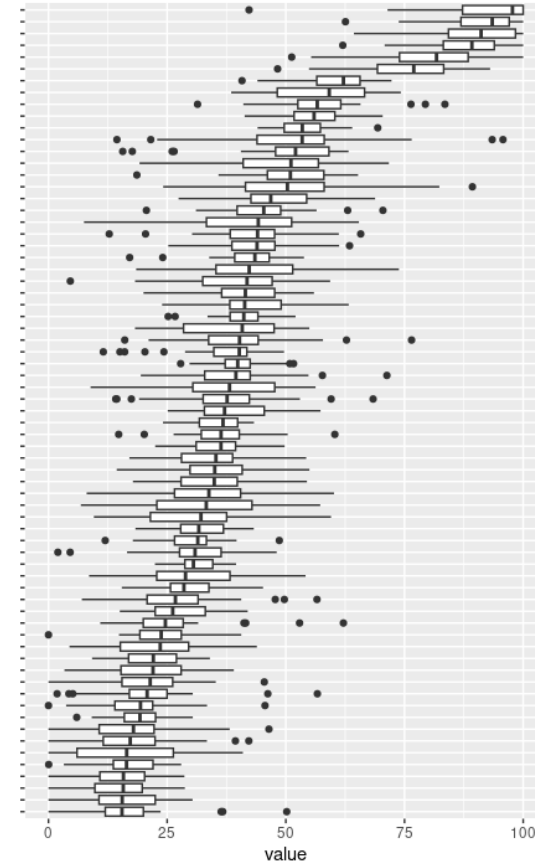


PREDICTION MODEL

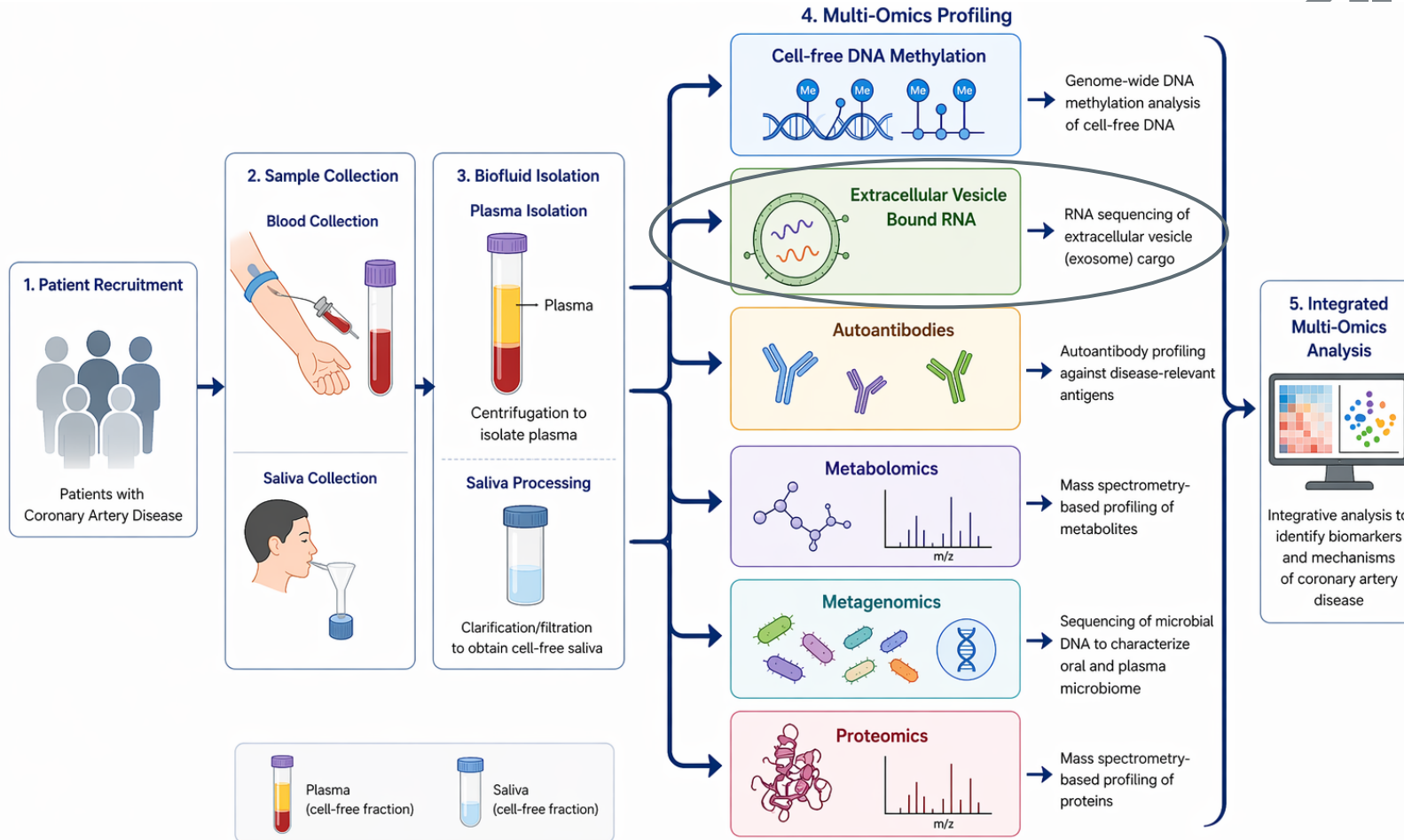
- Method: glmnet
- Nested crossvalidation:
 - train loop: 3rep 10xCV
 - test loop: 3rep 10xCV
- Selection metric: Accuracy
- Feature preselection: wilcoxon test ($p < 0.01$)



glmnet // wilcox features



OMICS LAYERS IN STENOSIS COHORT



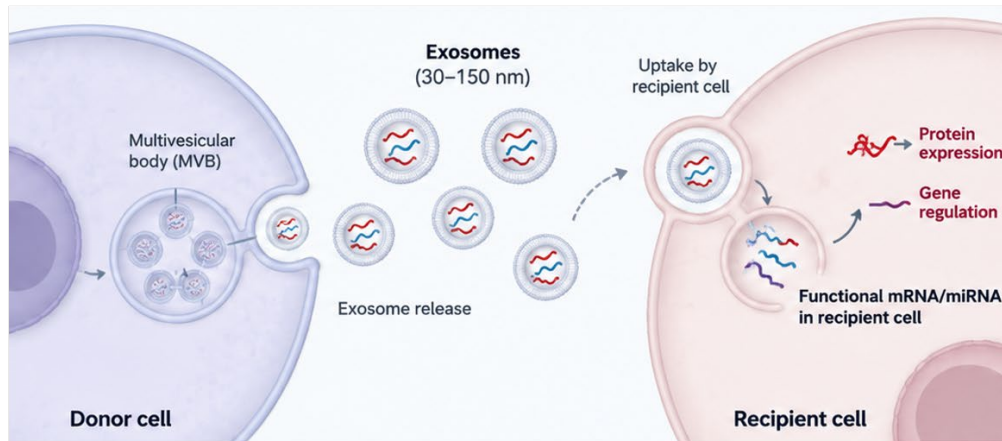
DIAGNOSTIC POTENTIAL OF EV-BOUND MIRNA

Letter | Published: 07 May 2007

Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells

Hadi Valadi, Karin Ekström, Apostolos Bossios, Margareta Sjöstrand, James J Lee & Jan O Lötvall 

Nature Cell Biology 9, 654–659 (2007) | [Cite this article](#)



1



Exosomes contain RNA and small RNA.

Exosomes isolated from mast cells contained mRNAs and miRNAs.

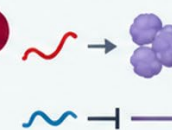
2



EV RNA can be transferred between cells.

Exosome-derived RNA was delivered to recipient cells and detected in their cytoplasm.

3



Transferred RNA can remain functional.

Delivered mRNAs were translated into protein and miRNAs could regulate gene expression in recipient cells.

EV / cfRNA provide a link between solid tissue and circulation

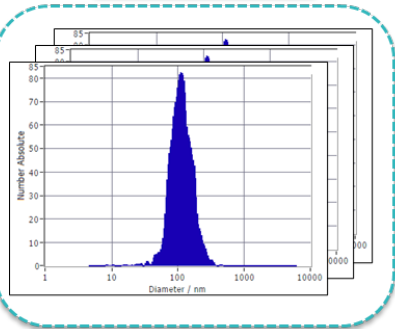
DIAGNOSTIC POTENTIAL OF EV-BOUND / CF RNA

STEP 1 EV ENRICHMENT



130–500 μL biofluid
 $\rightarrow \sim 10^9 - 10^{11}$ EVs

QUALITY CONTROL

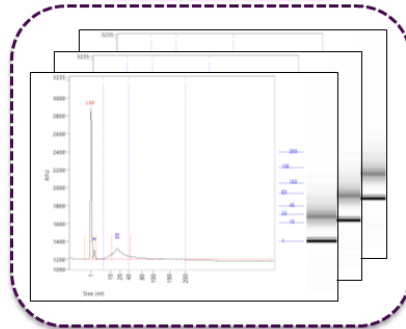


STEP 2 RNA ISOLATION

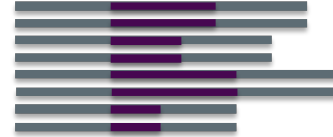


$\sim 10^9 - 10^{11}$ EVs
 $\rightarrow 1-5$ ng RNA

QUALITY CONTROL

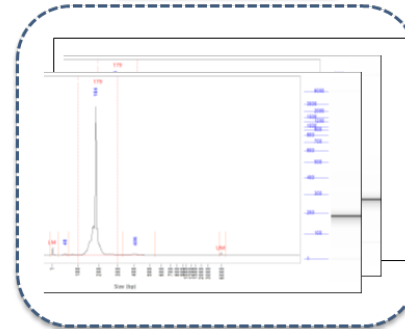


STEP 3 LIBRARY PREP



1–5 ng RNA
 $\rightarrow > 25$ ng DNA

QUALITY CONTROL

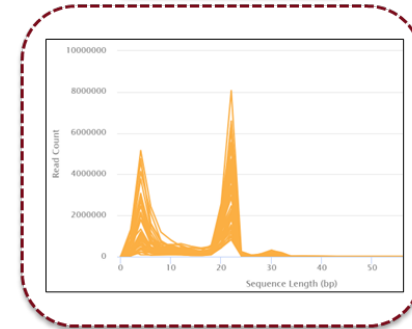


STEP 4 SEQUENCING

ACGTACGATC
CATCGACGAT
ACACACGACG

25 ng DNA
 $\rightarrow 5-20\text{M}$ reads

QUALITY CONTROL



WHOLE / UNFRACTIONATED BIOFLUID

soluble proteins
cell-free nucleic acids
metabolites
lipoproteins / protein complexes
cells + debris (fluid-dependent)

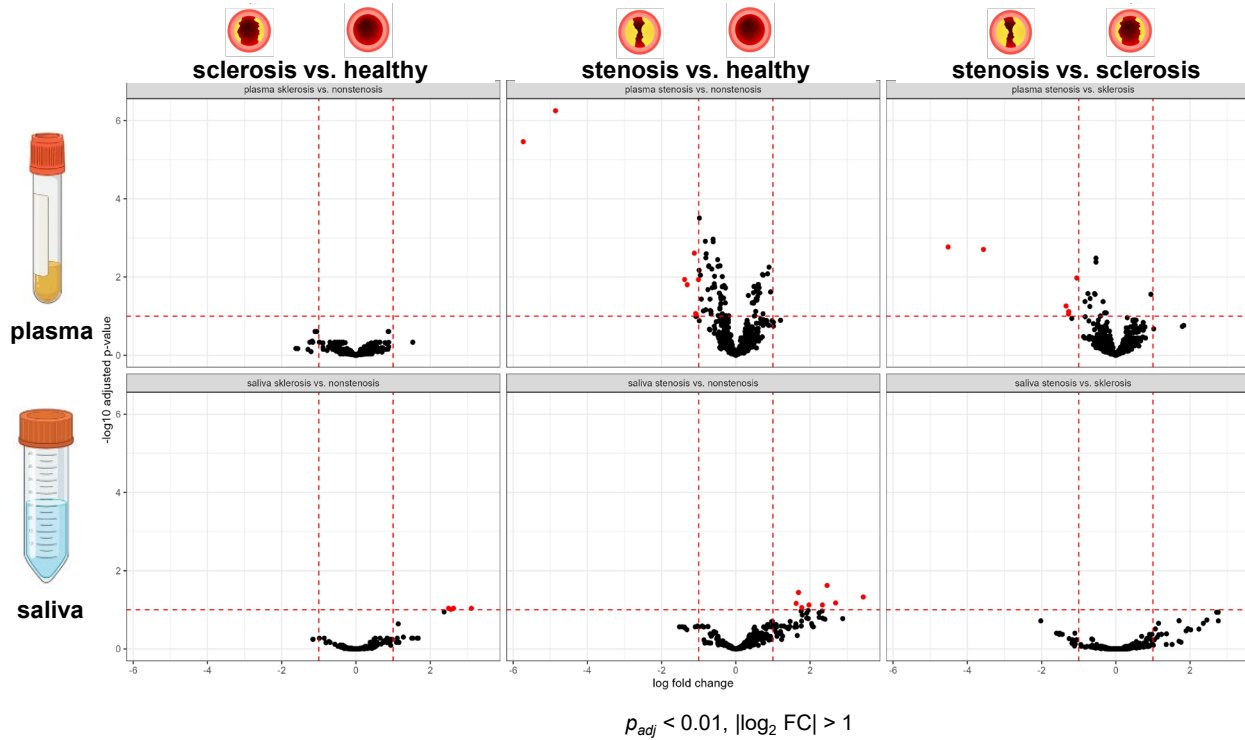
Usually simpler • potentially higher abundance
• **easier to standardize**

EV-ASSOCIATED SIGNAL

ENRICHMENT	concentrate a disease-associated molecular fraction
PROTECTION	cargo may be protected from degradation
MULTIPLEX CONTEXT	combine surface phenotype with internal cargo
CELLULAR CONTEXT?	potentially enrich EVs associated with particular cell types

Potential advantages which need to be demonstrated rather than automatic properties of every EV assay.

DIAGNOSTIC POTENTIAL OF EV-BOUND MIRNA



Iqra Yousaf



Michaela Hendling

- Both methods (PhIPseq & EV/cfRNA) offer hypothesis-free biomarker discovery
- Small amounts of serum / plasma required (50 - 500 ul)
- Both methods have potential to identify stenosis in a relevant clinical setting.
- Multi-omics analysis ongoing
- Methods are ready to be applied in other disease settings (e.g. endocrinology)

THANKS TO THE COLLEAGUES



Klemens Vierlinger
Sr Scientist
Bioinformatics



Gabriel Vignolle
Scientist
NGS Analysis



Christa Nöhammer
Sr Scientist
Biomarker development

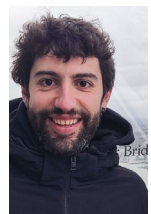
LANDESKLINIKUM 
WIENER NEUSTADT



Maximilian Tscharre



Franz Xaver Roithinger



Enrique Ozcariz
Jr Scientist
Bioinformatics



Michaela Hendling
Jr Scientist
Bioinformatics



Jasmin Huber
Jr Scientist
Biomarker development



Alice Limonciel



Guido Dallmann



EP PerMed
European Partnership
for Personalised Medicine

FWF Österreichischer
Wissenschaftsfonds

This project received funding from FWF under the frame of the European Partnership for Personalised Medicine, EP PerMed, (GA N° KIN 4017425 of the EU Horizon Europe Research and Innovation Programme).

