Biomarkers discovery in rare diseases and implications for therapy
Alessandra Ferlini

Disclosure of Interest

EU Project BIO-NMD (Coordinator)
FINGER, NMD-CHIP, NEUROMICS EU PROJECTS (Full Partner)

PI, PROSENSA DUCHENNE TRIAL PRO044
PI GSK (PRO051/DRISAPERSEN GSK2402968) DUCHENNE TRIAL

TREAT-NMD Alliance (Member)
European NeuroMuscularCentre Scientific Committee (Member)
• **TOPICS**

• Biomarkers: definition, types and use
• Biomarkers in therapies: robustness, pros & cons
• Biomarkers in RDs: needs and expectations
• Biomarkers in drug repositioning
• The EU funded BIO-NMD project
• BIO-NMD outputs: novel biomarkers
• Conclusions: biomarkers policy and future perspectives
**BIOMARKER DEFINITION**

- **Biomarker definition**
  
  *Definition:* EMA

- A *measurable* (DNA and/or RNA if genomic biomarker) *biological* characteristic, that is an indicator of normal biologic processes, pathogenic processes, and/or response to therapeutic or other interventions.

- **Types:**
  - Diagnostic
  - Prognostic
  - Therapeutic
    - Pharmacogenetics (profiling the genotypic response to a drug)
    - Pharmacogenomics (identifying the gene(s) target of a drug)
    - Pharmacokinetic (what the drug does to the body)
    - Pharmacodynamic (what the body does to the drug)
Why biomarkers are crucial in rare diseases therapies?

- Hereditary rare diseases are chronic and often degenerative disorders and the chronicity often hampers using clinical outcome measures (COMs)

- Novel trials are often short and not always impacting on clinical phenotype

- Biomarkers can flank COMs offering early endpoints, surrogating efficacy measures, mirroring specific clinical features

- Novel trials are often short and not always impacting on clinical phenotype

- Biomarkers can be surrogate (instead of-) and/or mirroring (accompanying/anticipating-) COMs being early, sensible and specific marker of therapy efficacy/efficiency

- Biomarkers can be used in non-expensive neonatal screening (appropriate when therapy is available)

- Biomarkers can facilitate drug discovery and development, speeding up drug designation, they be diagnostics when identified a specific cause of the disease

- Early diagnosis at neonatal stage may maximise the efficacy of novel therapies

- RDs may have a long process from identifying novel drugs, followed by regulatory activities
Biomarkers: strength and weakness

**REPEATABILITY**
- fluids/ tissue /cells

**SPECIFICITY**
- Disease, pathways, genotype

**REPEATABILITY**
- reproducibility with different techniques, in different labs

**ROBUSTNESS !**

**ENVIRONMENT**

IT INFLUENCES RNA AND PROTEIN (but not DNA) BIOMARKERS REPEATABILITY

(dynamic, different sampling procedures, metabolic, circadian, age-dependent, dietary variations...)

RNA/PROTEIN FLUCTUATIONS! (lessons from Cancer...)

“ensemble feature selection” :
multiple feature selections are combined in order to increase the robustness of the final set of selected markers
Biomarkers: strength and weakness

DNA BIOMARKERS (SNPs-CNVs)
- High demanding task (discovery)
- high scale validation (1000>10000 controls)
- Functional meaning (development)
- Stable-specific-repeatable-easy accessible

RNA BIOMARKERS
- Tissue-related (discovery)
- Repeatability & specificity: low (development)
- Tissue/function-specific
- Not easy-accessible
- Finely-tuned: stability medium

PROTEOMIC BIOMARKERS
Fluids: (discovery & development)
Repeatability & stability: low
Specificity: organ/tissue dependent
- Discovery screening-based
- Cost/benefit: good ratio
- Easy accessible

High ROBUSTNESS
Medium/High ROBUSTNESS!
Low ROBUSTNESS!
Proteomic and transcriptomic biomarkers are influenced by POP (Personal Omics Profiling)

- extensive dynamic changes in molecular components and biological pathways depending on ages, conditions, environment, diseases, etc
- need of a longitudinal Integrated POP (iPOP) using all omics analyses to interpret healthy and diseases states

-Key Point: connecting genomic information with dynamic omics activities
Drug repositioning and biomarkers

Biomarkers can provide significant shortcut for drug repositioning strategies

Serological biomarkers (as neo-epitopes)

#Retrospective
Collecting biomarkers data from failed trials to redirect drug purpose

#Prospective
novel screening in samples of already conducted trials

Some examples:
- CTX-I
- Osteoporosis>short bowel Syndrome
- HbA1c
- Type 2 Diabetes>congenital steatohepatitis
Final View: Pros/Cons of Biomarkers

• Pros:
  - Objective
  - Change more rapidly than other endpoints
  - Faster detection of potential drug candidates
  - Predict earlier and faster drug response than clinical outcome measures
  - May reduce drug development costs
  - Accelerate development of novel drugs
  - May speed time to market

• Cons:
  - Validation - high demanding (new assays, new models) and costly
  - Qualification: required, robust evidences, validated pathways, repeatability, must be provided
The BIO-NMD project

“Identifying and validating pre-clinical biomarkers for diagnostics and therapeutics of Neuromuscular Disorders”

Type of funding scheme*
Collaborative Project
(Small or medium-scale focused research project)

Work programme topics addressed
HEALTH-2009-2.4.4-1: Rare neurological diseases.

Name of the coordinating person:
ALESSANDRA FERLINI

TOWARD A BIOMARKERS APPLICATION IN
- Personalized medicine (“right drug for right patient”)
- surmount critical barriers to speed up the clinical applications
- identifying new surrogate non-invasive endpoints
Who is involved?

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>Participant organisation name</th>
<th>Participant Name</th>
<th>Short</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Università degli Studi di Ferrara</td>
<td>UNIFE</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Academisch Ziekenhuis Leiden - Leiden University Medical Center</td>
<td>LUMC</td>
<td>The Netherlands</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>University of Newcastle upon Tyne</td>
<td>UNEW</td>
<td>United Kingdom</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>University of Padova</td>
<td>UNIPD</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Institute of Child Health, University College London</td>
<td>UCL</td>
<td>United Kingdom</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>University of Rome Tor Vergata</td>
<td>UNIROMA</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Institut National de la Santé Et de la Recherche Médicale</td>
<td>INSERM</td>
<td>France</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Università degli Studi di Milano</td>
<td>UNIMI</td>
<td>Italy</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Kungliga Tekniska Högskolan</td>
<td>KTH</td>
<td>Sweden</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Ariadne Genomics Inc.</td>
<td>ARIADNE</td>
<td>US</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>AppliedBiosystems</td>
<td>AB</td>
<td>Germany</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>ACIES SAS</td>
<td>ACIES</td>
<td>France</td>
<td></td>
</tr>
</tbody>
</table>

*With company headquarters based in the USA*
NMDs: definition according to EU

Neuromuscular hereditary disorders (NMDs) belong to rare neurological disorders

- include more than 350 different phenotypes
- all together affect 30/10,000 people in EU


http://www.eurordis.org/article.php3?id_article=839

-heterogeneous phenotypes
-recently promising novel PERSONALISED treatments (antisense oligonucleotides, stop codon reversion, stem cells, drugs) on trials (Duchenne muscular dystrophy, Spinal muscle atrophy)
-intense effort and support from Patients Association
-excellent network with EU SMEs in discovering and developing orphan drugs
The BIO-NMD approach for discovery task

- candidates genes approach
- OMIC approaches: genomics, transcriptomics, proteomics
- Fine tuning of selected phenotypes

• KEY WORDS OF BIO-NMD
  • BIOBANKS
  • DATABASE AND BIOINFORMATICS
  • SAMPLE/TISSUE/PHENOMICS SHARING
  • PARALLELING – integrated OMICS PROFILING (i-OP)
    – DIFFERENT BIO-samples IN THE SAME PATIENTS
    – SAME BIO-samples IN DIFFERENT PATIENTS (CATEGORIES)
    – COMPARE PATIENTS CATEGORIES (MULTIPLE FEATURES, SUB-PHENOTYPES)
      • To avoid normal control analysis
    • >>> genome inflation, control fluctuations
### APPROACHES USING PATIENTS’ MATERIAL

<table>
<thead>
<tr>
<th>Parallel OMICS techniques</th>
<th>Targeted gene techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole exome sequencing</td>
<td>Genes/pathways identified by Ariadne MEDSCAN software (~200 genes)</td>
</tr>
<tr>
<td>Targeted exon sequencing</td>
<td>Genes/pathways identified in animal models</td>
</tr>
<tr>
<td>Targeted SNP capturing</td>
<td>Low-invasive cell search</td>
</tr>
<tr>
<td>Whole RNA sequencing</td>
<td>Multiplex ELISA</td>
</tr>
<tr>
<td>MyomiRNA sequencing</td>
<td>chIP immunoassay</td>
</tr>
<tr>
<td>2-DIGE mass spectrometry</td>
<td>SNP study</td>
</tr>
<tr>
<td>Liquid chromatography-mass spectrometry</td>
<td>Low/non-invasive cell studies</td>
</tr>
</tbody>
</table>

### PATIENT CATEGORIES DEFINED

- **DMD UNDER STEROIDS**
  - **RESPONDERS VS NON-RESPONDERS**
    - **Muscle**: longRNAseq, myomiRNA seq, 2-DIGE
    - **DNA**: whole exome, SNP arrays, candidate genes
    - **Cells**: steroid influence on exon skipping
  - **DMD SUPER SURVIVORS**
    - **Muscle**: long RNAseq, myomiRNA seq, 2-DIGE
    - **DNA**: whole exome, SNP arrays, candidate genes, targeted exome sequencing
  - **DMD EARLY vs LATE LOSS OF AMBULATION**
    - **Muscle**: longRNA seq, myomiRNA seq, 2-DIGE
    - **DNA**: whole exome, SNP arrays, candidate genes, targeted exome sequencing

- **DMD vs BMD**
  - **Muscle**: longRNAseq, myomiRNA seq, 2-DIGE
  - **DNA**: whole exome, SNP arrays, candidate genes, targeted exome sequencing

- **COLVI patients before and after cyclosporine A treatment**
  - **Muscle**: longRNAseq, myomiRNA seq, 2-DIGE
  - **DNA**: whole exome, SNP arrays, candidate genes

- **SYMPTOMATIC FEMALE CARRIERS OF DYSTROPHIN MUTATIONS**
  - **Muscle**: longRNAseq, myomiRNA seq, 2-DIGE
  - **DNA**: whole exome, SNP arrays, candidate genes

- **UCMD vs BM PATIENTS**
  - **Muscle**: 2-DIGE

- **DMD AND BMD DISEASE SEVERITY, PROGRESSION AND DRUG (STEROID) RESPONSE**
  - **Plasma/serum**: multiplex ELISA, chip immune-assays
Overall omics strategies for biomarker discovery

Proteomic Biomarkers Discovery:
- Protein microarrays
- LC/MS/MS
- FTICR
- 2D-DIGE

Genomic Biomarkers discovery:
- Whole exome sequencing
- Whole genome CNVs
- RNAseq and miRNAseq

Biomarker validation in animal models:
- mdx
- mdx Utrn -/-
- Col6a1 -/+

Cross-checking data
Select optimal candidate biomarkers

Biomarkers Validation - Genomic/proteomic
- Methodological: different laboratories/different platforms

TREAT-NMD
Ready to Clinical trials

Predclinical biomarkers validated for Dystrophinopathies and COLVI myopathies

EMEA regulations

TREAT-NMD Neuromuscular Network

Clinicians: Phenotypes correlation Collecting patients non-invasive samples

Bioinformatics "ad hoc" NMD dedicated software

UNEW (EuroBioBank)
Tissues (80)*
Cells (80)*

UNEW (EuroBioBank):
DNA (400)
Tissues (80)*
Cells (80)*

NMD Patients

ARIADNE:
BIOMARKERS AND TOOLS DEVELOPED

Tools designed during the course of the BIO-NMD project

Designed novel arrays for targeted sequencing and SNP detection
Designed and validated chIP-on-chip array for dystrophin chromatin profiling
Designed novel enrichment methods for targeted sequencing and copy number variation (CNV) detection
Designed novel bioinformatics tools for neuromuscular gene pathways analysis
Designed novel bioinformatics tools for OMIC data description, collection and interpretation
Designed novel immunoassay arrays for protein profiling in plasma/serum (using a collection of more than 500 antibodies)

Biomarkers identified during the course of the BIO-NMD project

**Exploratory (human)**
Identified more than 400 susceptibility genes as diagnostic or pathophysiological biomarkers
Identified 3 SNPs as drug response and disease severity biomarkers
Identified 15 ncRNAs as pathophysiological biomarkers
Identified 3 novel candidate genes for COLVI myopathies
Identified 864 candidate biomarkers to be explored in body fluids

**Exploratory (animal models)**
Identified specific gene pathways and expression profiles (autophagy, circadian clock, apoptosis) involved in animal models
EXAMPLE OF INTEGRATED OMIC PROFILING FOR BIOMARKERS DISCOVERY IN DUCHENNE MUSCULAR DYSTROPHY

Combining Phenomics, Next generation Target Capture (SOLiD™4) and RNAseq
Aim: discovery of pharmacogenetic (SNP) biomarkers for steroid response
STEROID THERAPY IN DMD

• Randomised controlled studies show that glucocorticoid corticosteroid therapy in DMD improves muscle strength in the short-term (six months to two years).
• the regime is 0.75 mg/kg/day, daily.
• adverse effects are common but not clinically severe (short term).
• adverse effects are severe in long-term treatment.
• Patients can be high responders or low/non responders
• Identifying low/non responders will impact on therapy choices and clinical practice.
Duchenne Patients Selection (Phenomics)

21 Patients enrolled
Selection criteria:
Response to corticosteroids (Prednisone)

13 DMD Low responders:
  Loss
  Ambulation <10 y

8 DMD High responders:
  Loss
  Ambulation >15 y

COMPARING PATIENTS’ CATEGORIES
Targeted sequencing (SNP)
RNAseq
METHODS: gene prioritization

- INTERACTOME MAP using the Ariadne MedScan pathway and identification of 243 genes connected to DMD pathways
  - custom array enriched for capturing the 243 genes
  - 23 patients analysed

TOTAL CALLED: 1714 SNPs in 21 DMDs

FILTERS APPLIED:
- 590 SNPs not in SNPs databases
- 446 SNPs monomorphic (single allele)
- 59 SNPs « individuals »

ANALYSED 619 SNPs  376 synonymous  243 missense
DISCRIMINANT ANALYSIS

- Novel statistical tool - Discriminant analysis (DA) - used to determine which variables (SNPs) discriminate between two occurring groups (Responder/nonResponders).
- DA defines if a SNP “preferably” belongs to a specific category therefore discriminates individuals into pre-defined groups (explained variance measures statistical meaning).
CONCLUSIONS

ONLY MISSENSE SNPs (243) ALLOW AT GROUPING THE 2 DMD CATEGORIES (responders vs non responders)

39 SNPs ARE CONSISTENTLY DIFFERENTIALLY GROUPED IN DMD-R VS DMD-NR

Explained Variance: 53.25%
INTEGRATED OMICS PROFILING: SNPs MAY MATCH TRANSCRIPTS DIFFERENTIALLY EXPRESSED IN MUSCLE

DMD MUSCLES AVAILABLE

RNAseq ON 6 MUSCLES OF THE SAME DMD PATIENTS

3 RESPONDERS
3 NON RESPONDERS

Genes with different signs (normalizing effect):
- DMD vs Normal > 0, High resp vs Low resp < 0
- DMD vs Normal < 0, High resp vs Low resp > 0
RESPONDERS
25 genes resulted normalized after steroid-treatment

Positive effect of glucocorticoids (High vs Low resp)
Conclusions

Predicting response to glucocorticosteroids in DMD patients

• We identified **6 SNPs in 4 genes** that discriminate between DMD responders and non responders to corticosteroids response

• Further validation in DMD patients

• Validation in other NMDs (congenital myopathies)

• Validation in non-rare diseases (asthma & autoimmune diseases, other disorders with variable steroid response)
BIO-NMD DIRECTIONS

BIO-NMD: EMA briefing meeting to propose 3 top biomarkers to be considered for a validation project

BIO-NMD: ENMC WORKSHOP APPLICATION FOR A JOINT EU/USA MEETING FOR BIOMARKER DISCOVERY OPTIMISATION IN NMDs

BIO-NMD: LINK WITH RD-CONNECT AND NEUROMICS FOR DATA EXPLOITATION, UNDER THE IRDiRC POLICY

- Collecting non-invasive samples from other NMDs (reducing costs; improving benefit, abating painful procedures)
- Validation process in other NMDs
- Improving prognosis and disease course assessment
- Diagnostic screening in wide population (cost-effective tool)
Final reflections

• Biomarkers pros & cons
  – Very useful for facilitating drug development process
  – Criticisms raised for biomarkers as “sole” surrogate endpoint

• Biomarkers in RDs
  – Urgent need for biomarkers for patients stratification and disease progression (clinical trial enrollment)
  – Urgent need for novel drug re-tasking purposes, drug response monitoring, side effects anticipation

• Biomarkers policy (core elements)
  • BIOETHICAL ELEMENTS
  • DATA SHARING (essential for RDs)
    • PRIORITIZATION MODELS
    • BIOBANKS AND REGISTRIES
  • DATA RELEASE & ACCESSIBILITY
ACKNOWLEDGMENTS