Reproducibility Of 2D Gel-based Proteomics Experiments

David Bramwell [1], Mary Caponite Hurley [2], Alamgir Khan [3], Gereon Poschmann [4], Jules A. Westbrook [5], Sjouke Hoving [6]

[1] Nonlinear Dynamics, Newcastle upon Tyne NE1 2ET, UK [2] Michigan Proteome Consortium, Ann Arbor, MI 48109-0404, USA [3] Australian Proteome Analysis Facility (APAF), Macquarie University, Sydney NSW 2109 Australia [4] Medical Proteom-Center, Ruhr-University of Bochum, D-44801 Bochum, Germany [5] Proteome Research Centre, UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Ireland [6] Novartis Institutes for BioMedical Research, WSJ-88.801, CH-4002 Basel, Switzerland

2D gels: the good sides and the real issues

- Limited sample prep required, denaturing conditions → sample preservation
- Inherently reproducible subset of proteome: always the 2'000 most abundant protein species in the sample (in special gel systems even more)
- Instant quantitation with good dynamic range using fluorescent dyes
- Straightforward ID and (partial) characterization of proteins, since all peptides in one spot
- Excellent resolution of most isoforms
- Parallel sample processing

First dimension

- IEF unforgiving towards any ionic "contaminant"
- Proteins are considered least soluble at iso-electric point, especially at the high concentrations achieved in IEF
- Poor coverage of "extreme" proteins:
  - TM-domain proteins
  - < 8 kDa, > 250 kDa, extreme p*I*
- General – poorly understood – issue with IEF of basic proteins

Second dimension

- "simple" SDS-PAGE, no real issues (???)

Protocol reminder

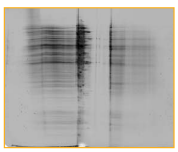
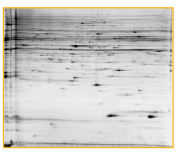
In the original invite the protocols were stated as:

"Each participating lab should run the two different, but related, samples on 2D gels according to the supplied protocol, then perform an analysis between the two groups, using the software provided to identify what are in their opinion the top 200 significantly changing spots (this would be up to 200 if the lab believes that there are less than 200 significantly changing spots, or over 200 spots if this result is obtained). This of course includes newly appearing spots.

In principle, no spot editing should be required, but if deemed needed any rejected spots should have comments added as to why they were rejected or edited (this is easily done in the supplied workflow).

The full analysis should then be archived and uploaded to the specified site for comparison between labs and further analysis."

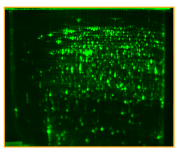
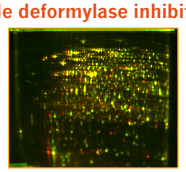
Suboptimal experiences

as shown by the GE Healthcare Discussion Board
(> 10'000 issues reported)

The two samples

H. influenzae — **treated with actinonin (peptide deformylase inhibitor)**

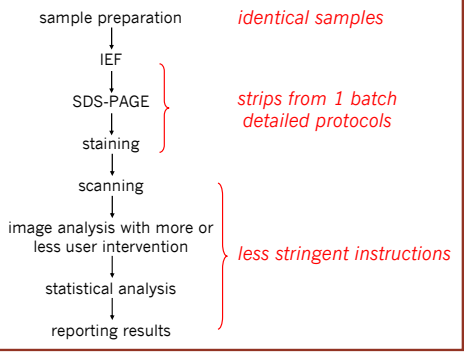
hundreds of differences to challenge the matching and alignment process

Challenges in comparing

- Coverage of the proteome
  - if proteome coverage is partial in each experiment, overlap between them is inherently limited
- Technical challenges, e.g. sampling in LC-MS experiments
  - "inherent" lack of reproducibility: dependent on sample complexity
- 2D-PAGE
  - no sampling problem – inherent reproducibility
  - many technical challenges
- Goal of this experiment
  - short-term: assess sources of variability
  - long-term: provide reference materials and protocols to proteomics community

Validating 2D PAGE in practice

sources of variability

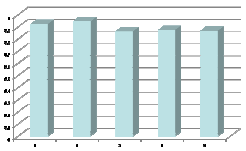


http://www.fixingproteomics.org

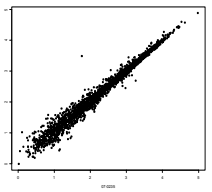
We support the fixing Proteomics campaign

Results

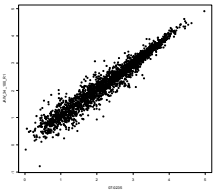
Reproducibility whole image correlation



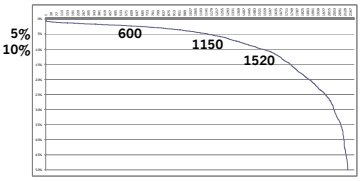
Volume ratio linearity within lab



Volume ratio linearity across lab



Variability of multiple runs
(1 lab, 3 operators 4 week timeframe)



20 gels each of treated and non-treated samples
High reproducibility does not require internal standards

Next steps

- Expand the experiment beyond the original five labs (ongoing)
- Submit study results for publication
- Provide reference sample to community
- Further options under discussion
  - central image repository with matching function for web-based QC of gels with reference sample

Conclusions

- Protein expression analysis based on 2D PAGE is reproducible across labs (and highly reproducible within a lab)
  - labs could easily share a common reference and hence the linked ID's
- Main sources of variability
  - user manipulation and interpretation of images (this study)
  - sample prep (in general)??
- With some constraints, experimental procedures appear to be remarkably robust
  - most similar images were generated with three different IEF instruments
  - variability in 2nd dimension, which affects direct image alignment, but not differential analysis

Acknowledgements

- Teams in Bochum, Sydney, Dublin, Ann Arbor and Basel for producing the data
- Team at Nonlinear Dynamics in Newcastle for a large effort in coordinating and performing the analysis
- Bio-Rad and CILBiotech for supporting the next phase with samples and logistics
- Jan van Oostrum and Will Dracup, the motivators behind (and sometimes at the forefront of) the scenes