

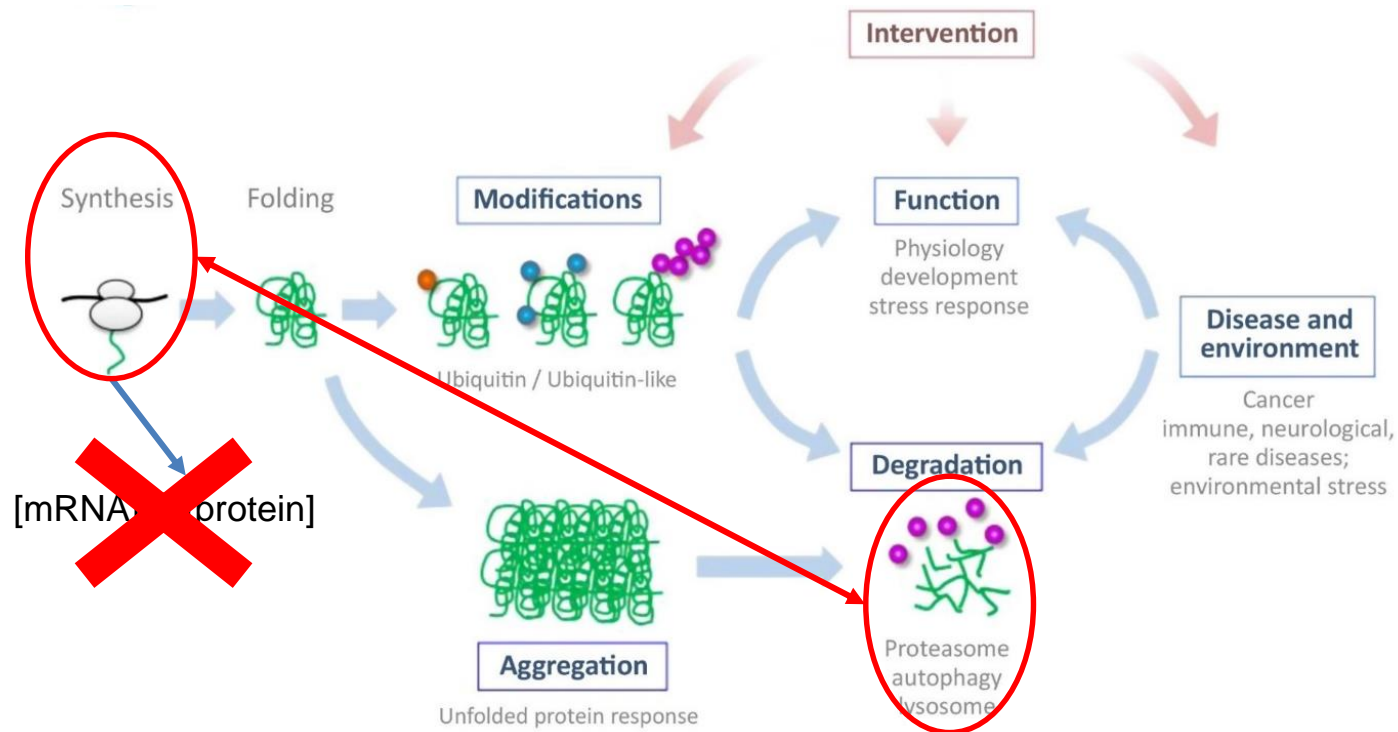
Génétique
Quantitative
et Évolution
Le Moulon



Success and bottlenecks of large scale plant protein turnover measurement: a tomato fruit maturation case study

Willy Bienvenut, IDEEV, GQE, Gif/Yvette, France

Proteostasis and large scale protein turnover rate determination



Dissmeyer *et al.* Trends Biochem. Sci. 2019

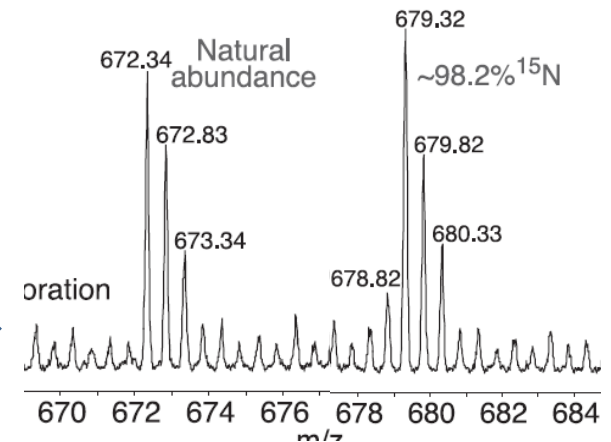
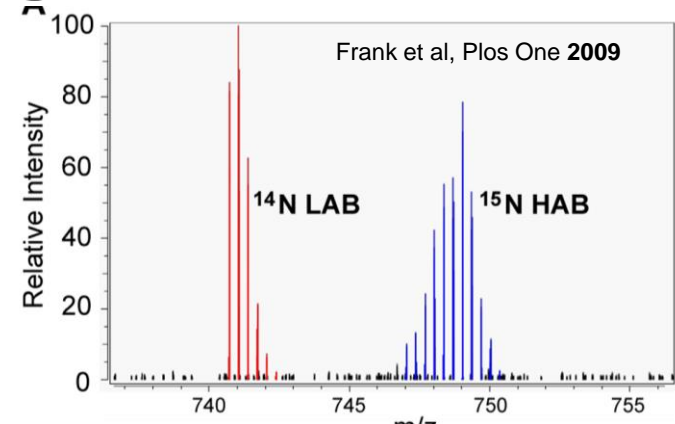
Influence of **protein Half-life and Turnover rate**
in **protein abundance variability**

Protein turnover in phototrophic organisms...



S. Lycopersicum
Var. *Moneymaker*
hydroponic culture
In mineral substrate

Water, CO₂,
fertilizers,
trace
elements...



¹²C ¹³C (CO₂)

¹⁴N ¹⁵N (NH₄ / NO₃)

Huttlin et al. *Mol. Cell. Proteomics* 2017

Usual K_d/K_s equations

Very few technical approaches proposed for plants...

$$K_D = - \frac{\ln(FCP \cdot (1 - LPF))}{t}$$

$$\frac{K_S}{A} = \frac{FCP - e^{-K_D \cdot t}}{1 - e^{-K_D \cdot t}} \cdot K_D$$

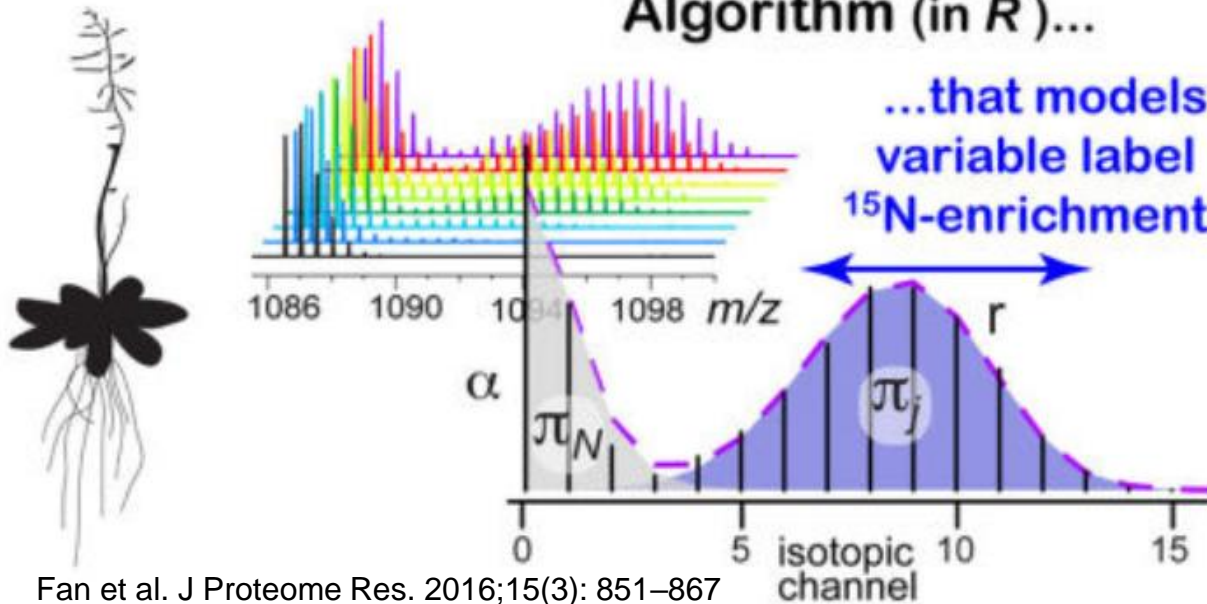
Determination of the:

FCP: Fold Change in Protein

LPF: Labelled Protein Fraction

Is that simple...?

A Proteome-Scale Protein Turnover Algorithm (in R)...



- High data processing
- Some scripts not are available
- N labelled peptides are not compatible with MS characterisation:
 - Most of the acquired data are useless

An alternative approach... ?

- Adapted to the processing tools available locally
 - i2MassChroQ (X!Tandem identification tool and inference filtering interface)
 - MassChroq (peptide quantitation)
 - MCQR (Protein quantitation & statistics R package)
- Providing a dedicated pipeline to perform large scale protein turnover measurement
- Reducing data complexity (MS) and cost (^{15}N) ...

Determination of the FCP coefficient
(Fold Change in Protein)



Label-free quantitation
MCQR tool

Adapted version compatible with ^{15}N labelling

Determination of the LPF
(Labelled Protein Fraction)



Based on the isotopic distribution
**MassChroQ &
dedicated R scripts**

Isotopic distribution: a new way to explore...

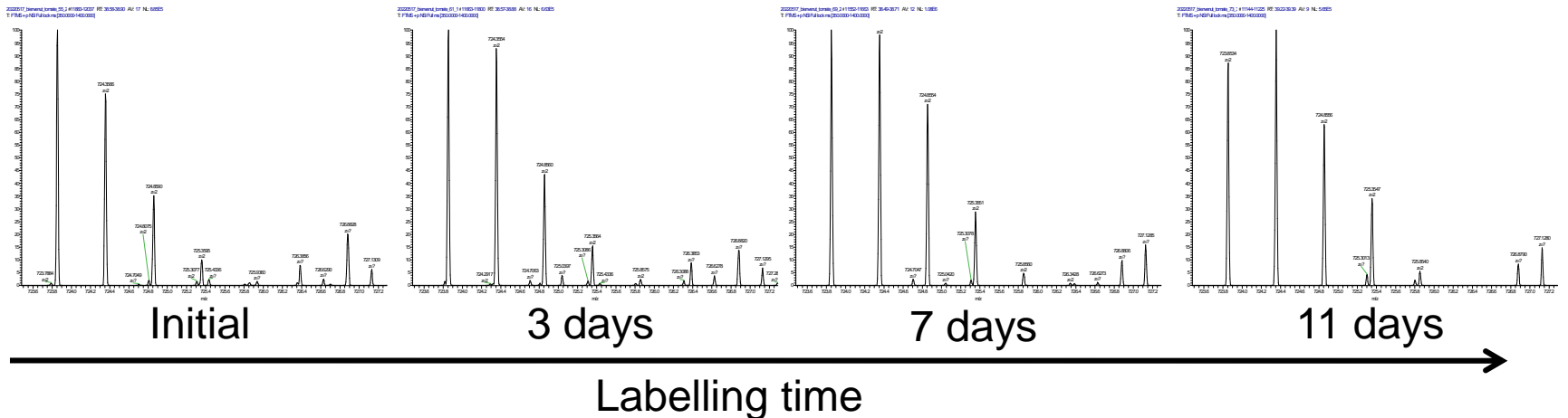
S. Lycopersicum
Var. money Maker
hydroponic culture
In coconuts fiber



5-10% ^{15}N
enriched media
→

Standard proteomics
sample processing

Metabolic
labelling pulse



How to estimate the LPF ?

$$\text{LPF} = 1 - \frac{R_1^{\text{Exp}} - R_1^{15\text{N}}}{R_1^{\text{Nat}} - R_1^{15\text{N}}}$$

$$R_1 = \frac{I_1}{I_0}$$

Exp: experimental measurement

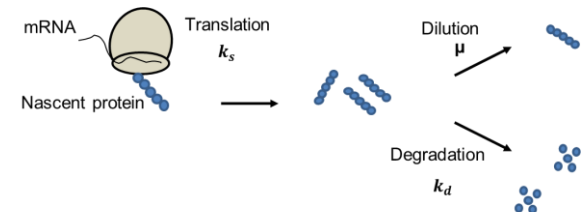
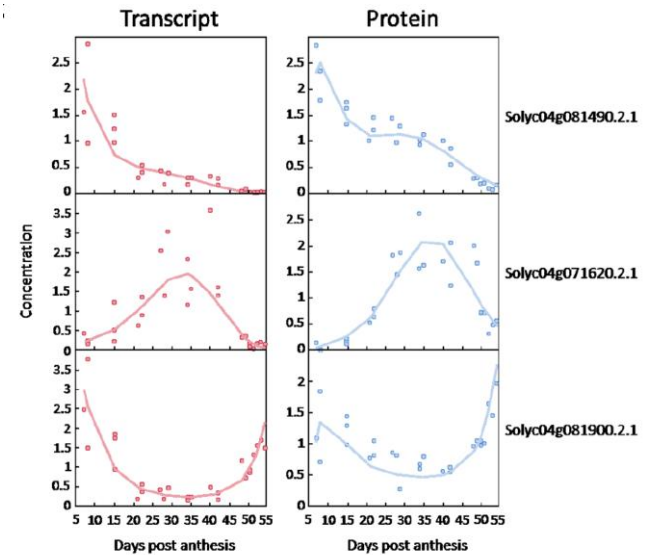
¹⁵N: Max ¹⁵N incorporation at t_∞ (Theo/Calc)

Nat: Natural isotopic distribution (Theo/Calc)

The PROOFER project:

exPeRimental validatiOn OF protEIn tuRnover

- Protein turnover model developed on tomato fruit
 - Belouah *et al.* 2019, J. Proteomics
- Protein turnover model using:
 - Transcriptomics data:
 - Absolute quantitation of 22877 transcripts with spikes in the experimental RNAseq
 - Proteomics data:
 - XIC absolute quantification of 2490 proteins using peptide intensity modelling



$$\frac{dp(t)}{dt} = k_s r(t) - (k_d + \mu(t))p(t)$$

Samples preparation and proteins ID



- * All samples collected at 12 days post anthesis
- * Labelling time was variable (0- 12 days)

Tomato developmental stages

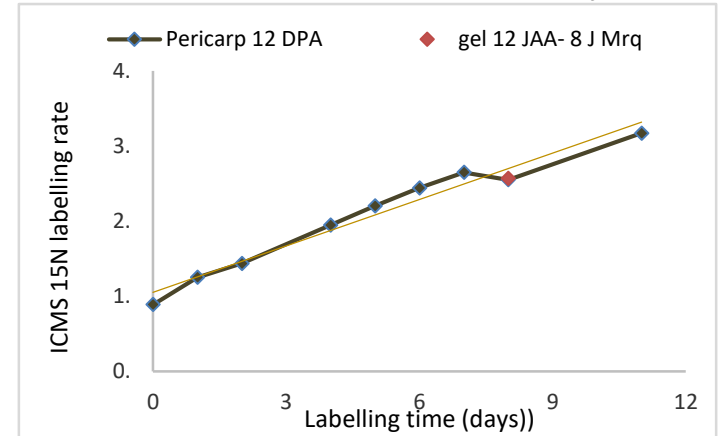


Samples collected at multiple ¹⁵N labelling time points

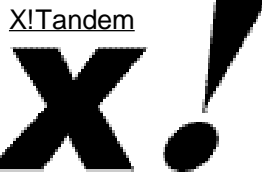


Standard proteomics sample preparation

Evolution of the ¹⁵N rate by ICMS

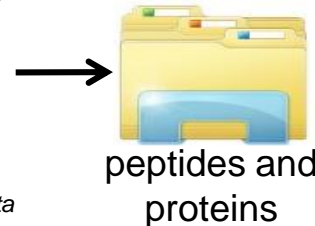


i2MassChroQ



REFERENCE-DB.protein.fasta

List of identified



MassChroQ
Label free quantitation
& Isotopic distribution exports

Isotopic

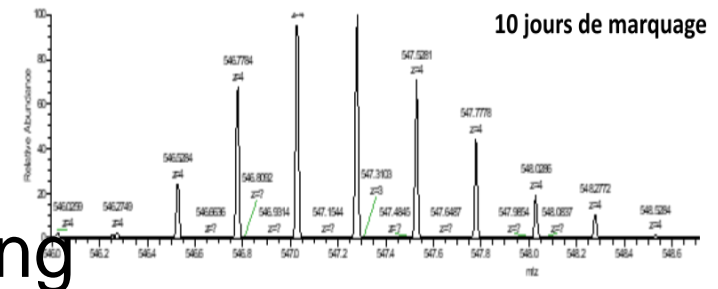
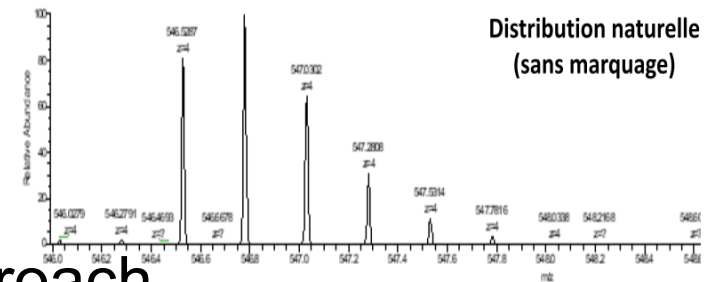


MCQR: R package for quantitation

How to determine the Fold Change in Protein (FCP)

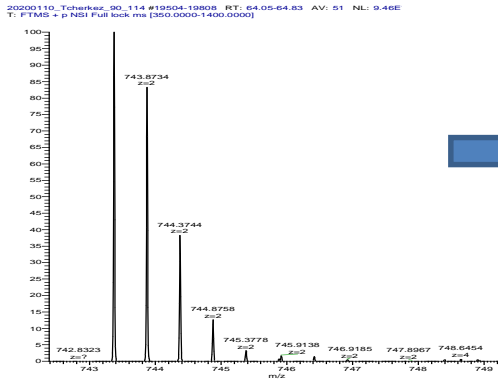
- Drawbacks:
 - Based on 1st or most intense isotope
 - ¹⁵N metabolic labelling modify the isotopic distribution
- Modified processes:
 - Intensity normalisation
 - Alternative “missing value” approach
 -
- New **IsoProt_MCQR** script
 - Compatible with ¹⁵N labelling

séquence du peptide : AGVKEYKLYTYTPEYETK



Display of the isotopic ratios

Peptide experimental isotopic distribution



$$\rightarrow R_i = \frac{I_i}{I_0}$$

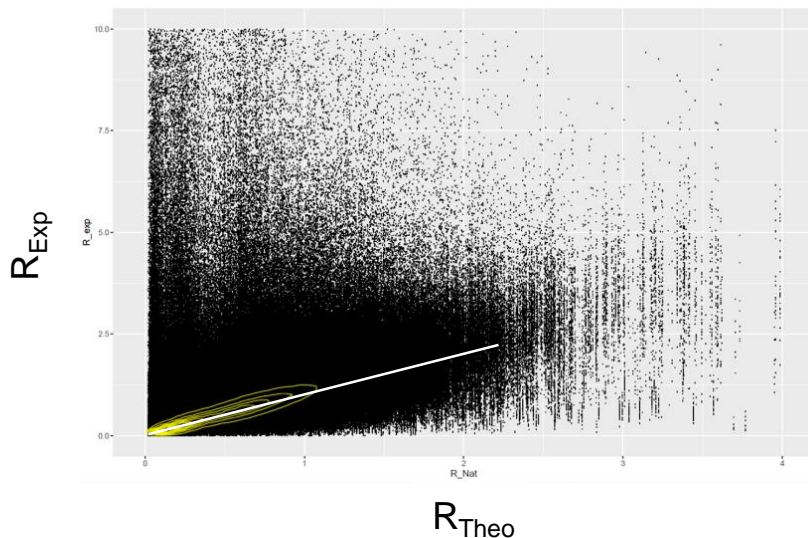
Experimental Intensity

Extracted by MassChroQ
Converted in R_i by R script

Theoretical R_i calculation

IsoSpec python library by M. Lacki;
adapted by P. Rusconi

Example for the Ctrl sample at $t=0$



5.190.167 R_i
106.219 peptides profiles
29,840 nr pepIDs

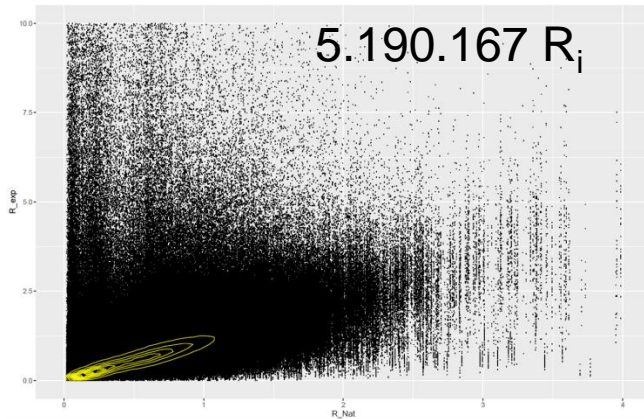
$$R_{\text{exp}} = 0.76 * R_{\text{theo}} + 0.39, R^2 = 0.02, \text{IFS} = 0.01$$

Filtering the raw data

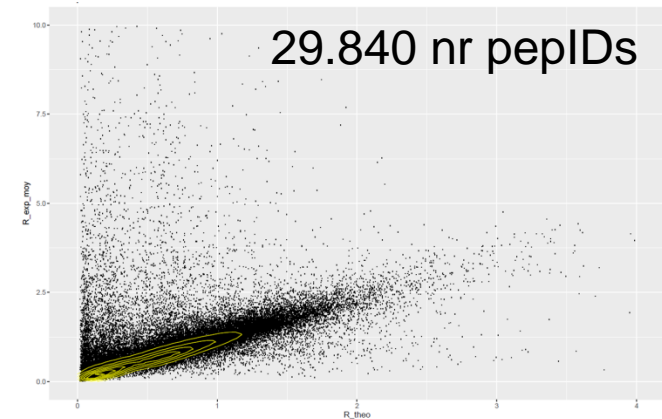
- Minimum number of isotopes:
 - Usually 4 isotops min (I_0 - I_3)
- Maximum ^{15}N incorporation rate:
 - based on theo. Max. ^{15}N rate isotopic distribution
- R_i dispersion:
 - $A_v \pm 1$ SD range retained (Theo. 70% retained)
- **Minimum data redundancy:**
 - Min 2/3 Anal. Rep.,
 - >4 isotopic profiles / Rep,
 - >3 Isotopes / isotopic profil
 - ...

Filtering the raw data (initial vs. final)

$$R_{\text{exp}} = 0.76 * R_{\text{theo}} + 0.39, R^2 = 0.02, \text{IFS} = 0.01$$

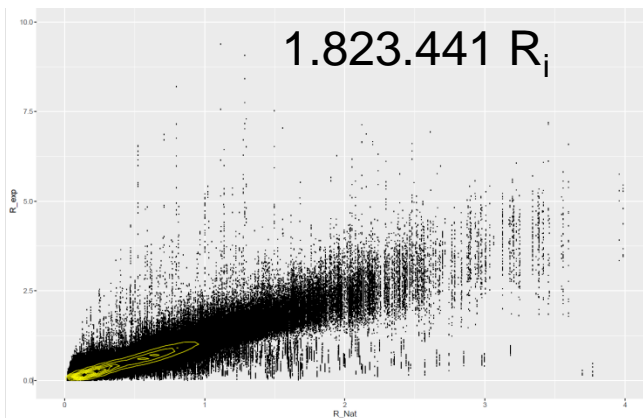


$$R_{\text{exp}} = 0.75 * R_{\text{theo}} + 0.43, R^2 = 0.07, \text{IFS} = 0.03$$

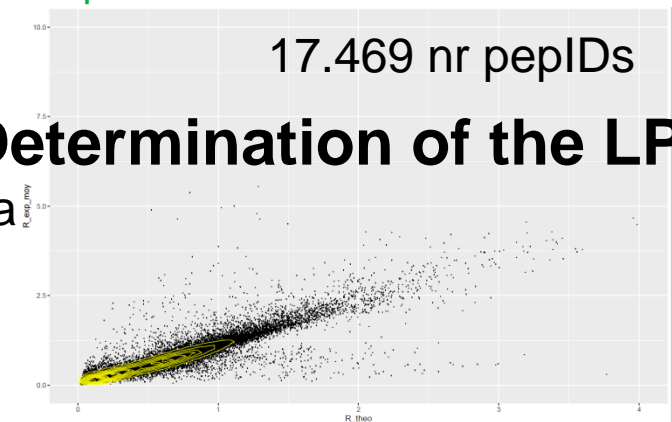


Initial data

$$R_{\text{exp}} = 1.01 * R_{\text{theo}} + 0.05, R^2 = 0.95, \text{IFS} = 1.27$$



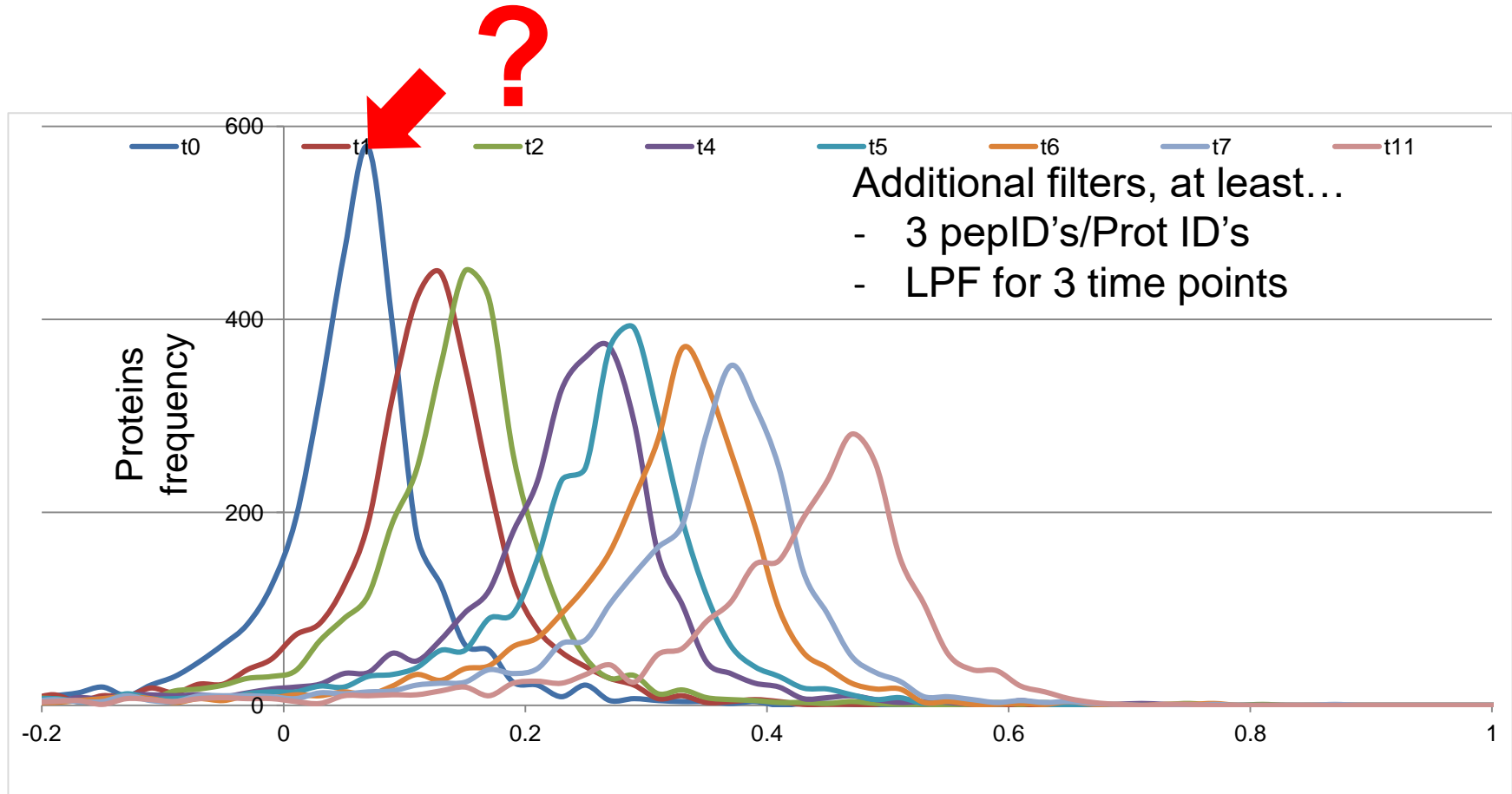
$$R_{\text{exp}} = 1.00 * R_{\text{theo}} + 0.06, R^2 = 0.94, \text{IFS} = 1.21$$



Filtered data

Determination of the LPF

Protein LPFs distribution

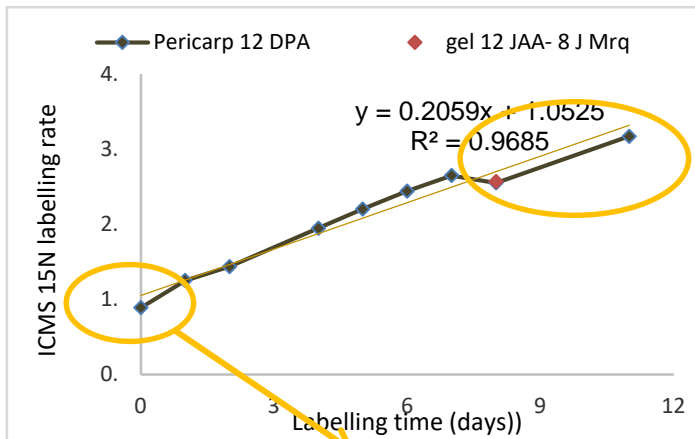


$\langle \text{LPF}(t_0) \rangle$ Expected at 0

Exp $\langle \text{LPF}(t_0) \rangle = 0.0517 \pm 0.0947$

Experimental drawbacks...

Evolution of the ^{15}N rate by ICMS



Natural ^{15}N rate = 0.364%



Technical problem related to automatic
plant watering system during the WE
Labelling time + 1 days

Tomato developmental stages



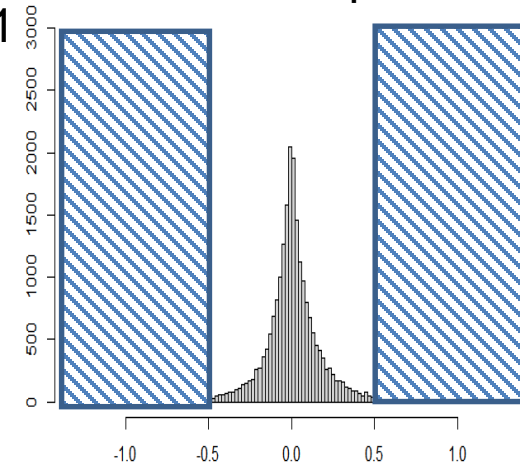
All samples were collected
at 12 days post anthesis

Labelling time was variable (0- 12 days)

By definition: 2511/2973 proteins
FCP = 1



Selection of
proteins with
no significant
variation in
quantity



Determination of the K_d/K_s

$$K_D = -\frac{\ln(FCP \cdot (1 - LPF))}{t}$$

Since $FCP = 1$

$$K_d(t) = \ln(1 - LPF)/t$$

$$\frac{K_s}{A} = \frac{FCP - e^{-K_d t}}{1 - e^{-K_d t}} \cdot K_D$$



$$K_s = K_d$$

Av. individual values

$$\langle K_d \rangle = 0.0528 \pm 0.0077 \text{ d}^{-1}$$

Classic linear reg.

$$K_d = 0.0536 \text{ d}^{-1}, R^2 = 0.987$$

Reg. line through the origin

$$K_d = 0.0538 \text{ d}^{-1}$$

Av. individual values

$$\langle K_d \rangle = 0.0756 \pm 0.0242 \text{ d}^{-1}$$

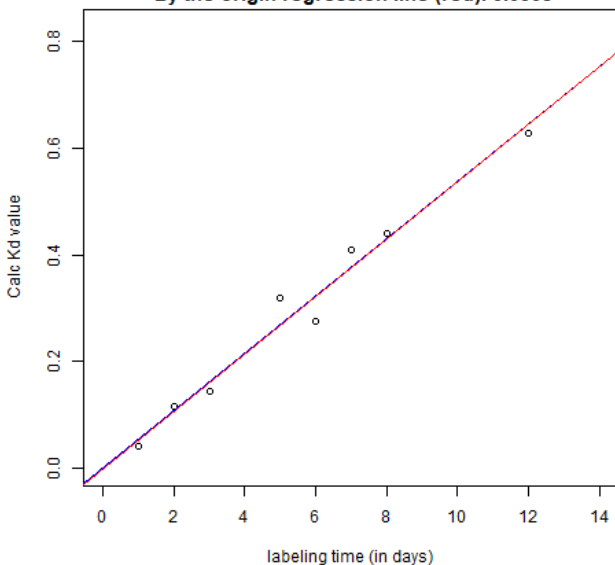
Classic linear reg.

$$K_d = 0.0469 \text{ d}^{-1}, R^2 = 0.839$$

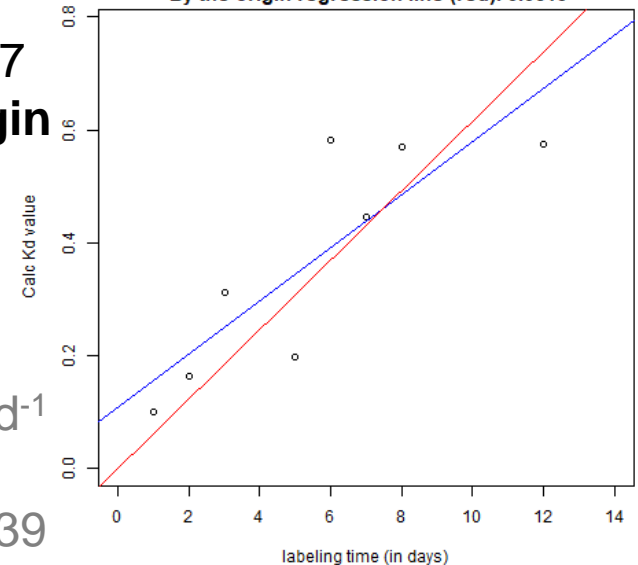
Reg. line through the origin

$$K_d = 0.0615 \text{ d}^{-1}$$

Solyc00g500057.1.1 : $\langle K_d \rangle = 0.0528 \pm 0.0077$
Regression line (blue): 0.0536 Corr. : 0.987
By the origin regression line (red): 0.0538



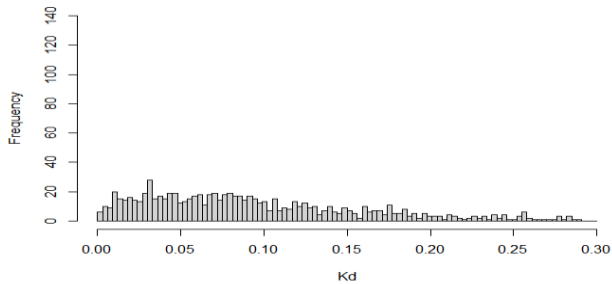
Solyc01g106360.3.1 : $\langle K_d \rangle = 0.0756 \pm 0.0242$
Regression line (blue): 0.0469 Corr. : 0.839
By the origin regression line (red): 0.0615



Results:

Predicted K_d (Belouah *et al.* model) vs. Exp K_d

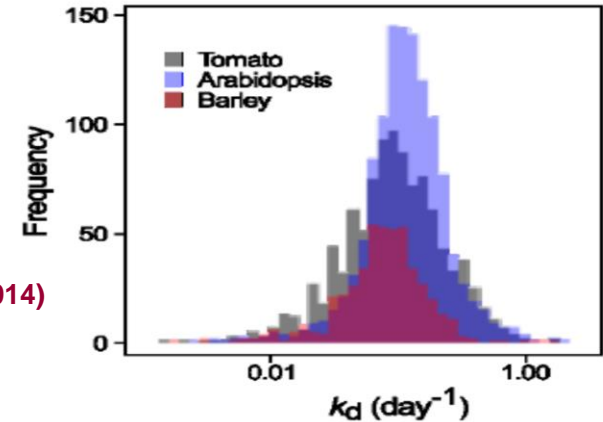
distribution des K_d (modèle)



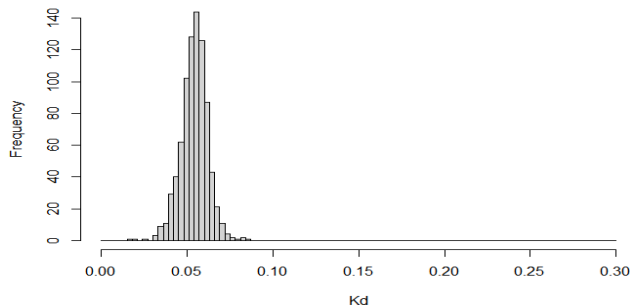
Distribution of 1103 predicted k_d of the tomato fruit : 0,113 day

Arabidopsis leaf median 0.113 day⁻¹,
~ 9 days (1228 k_d from Li *et al.*, 2017)

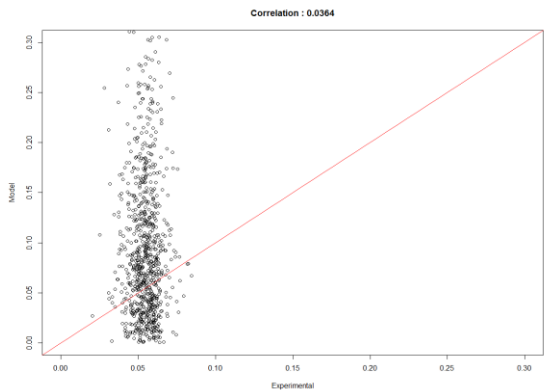
Barley leaf median 0.076 day⁻¹,
~13 days (505 k_d from Nelson *et al.*, 2014)



distribution des K_d (expérimentaux)



Averaged individual values
 $\langle K_d \rangle = 0.0601 \pm 0.0121 \text{ d}^{-1}$



Similar $\langle K_d \rangle$ value but with a much more restricted range
Strong discrepancies between the model and Exp. values

Conclusions

- Partial ^{15}N metabolic labelling:
 - compliant with LPF estimation
 - compatible with ^{15}N labelling
- Calculation of $K_d(K_s)$ is possible
- Restricted K_d range may be related to:
 - The sampling plan (too short, 1-10 days)
 - Probable influence of the metabolic lag time (28h)

Questioning...

- Some additional investigations are required to understand the results discrepancies...
 - What about a model using mRNA & Protein quantitative values to determine protein turnover...
 - Some major bias in Millar's processing pipelines (related to the FCP)...
 - Too few reliable studies in plant for an accurate comparison of the results

Acknowledgement's

Liam Le Goffic
Thomas Welfelé
Olivier Langella
Thierry Balliau
Filippo Rusconi,
Marlène Davanture,
Mélisande Blein-Nicolas



Collaboration with

C. Masclau-Daubresse & A. Marmagne
(IJPB, Versailles)

S. Colombié & Y. Gibon
(BFP, Bordeaux)

PROOFER grant (BAP department INRAE)