Transcriptomics in Early Safety Screening

Case studies from QSTAR Project

Bie Verbist

janssen | PHARMACEUTICAL COMPANIES OF Johnson&Johnson

Overview

Introduction

Case study 1: lead selection

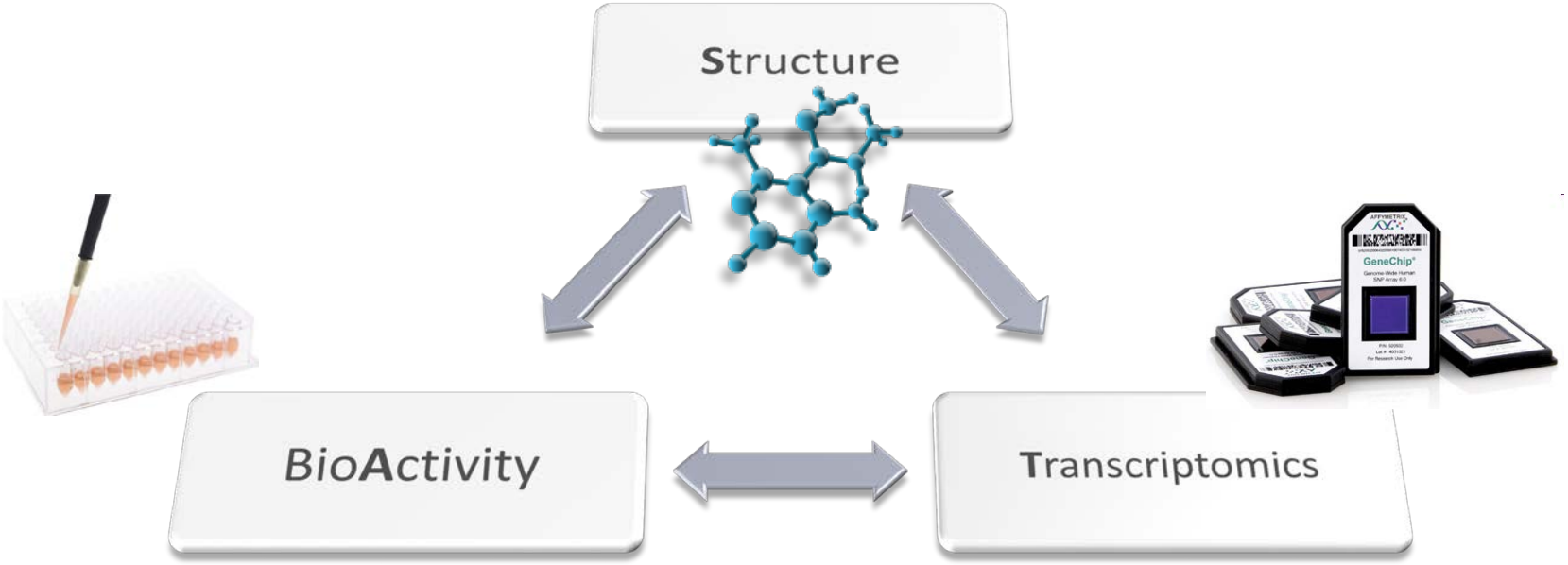
Case study 2: lead optimization

Wrap up

janssen | PHARMACEUTICAL COMPANIES OF Johnson & Johnson

QSTAR

Quantitative **S**tructure-**T**ranscriptomics-**A**ctivity Relationships



50% government sponsored 2-year research project involving a multidisciplinary team of 30 Janssen employees (11 departments), 8 academic partners and 3 CROs.

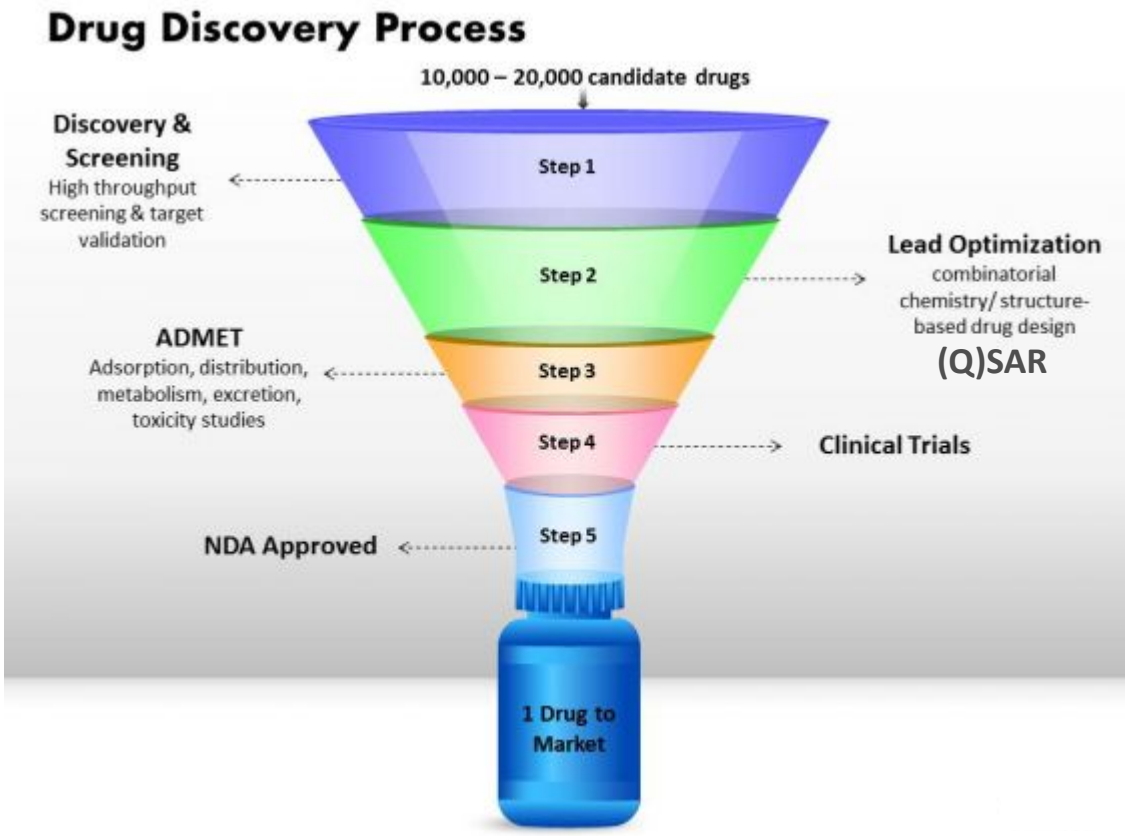
Goal

Increase the productivity of the R&D process, by reducing the risks of failure during the expensive late stages in drug development.

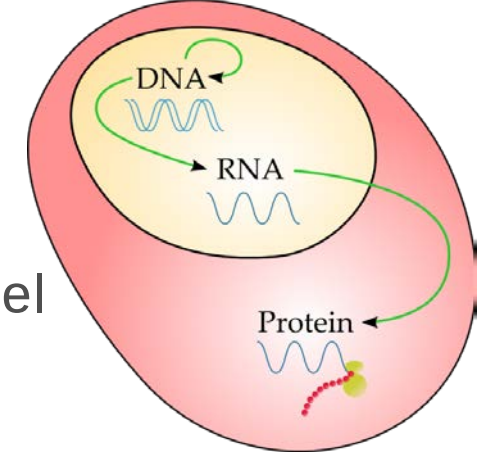
Compounds fail often in step 3 or 4 due to adverse effects.

Can **transcriptomics** help in guiding the SAR ?

-> (Q)STAR



Transcriptomics



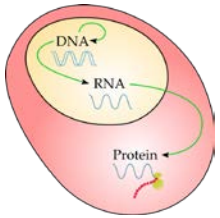
T*ranscriptomics* examines the expression level of mRNAs in a given cell population.

It includes *all mRNA* transcripts in the cell, hence it studies the genes that are being actively expressed at any given time under any given condition (e.g. after compound treatment):

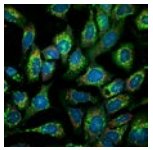
- **Multi-dimensional** assay: measures many biological effects simultaneously.
- **High-troughput** techniques: e.g. Microarray, RNA seq

	Gene 1	Gene 2	Gene 3	....
123	8.4	2.3	3.5	...
148	2.3	3.1	3.4	...
...	...	...	...	...

… and other techniques



- Proteomics: looking at entire set of proteins that are produced or modified by a system. Different level of understanding compared with transcriptomics:
  - Amount of mRNA is rough estimate of level of protein, because degradation or inefficient translation.
  - Many transcripts (mRNA) give rise to more than one protein, through alternative slicing or post-translational modifications.
  - ….

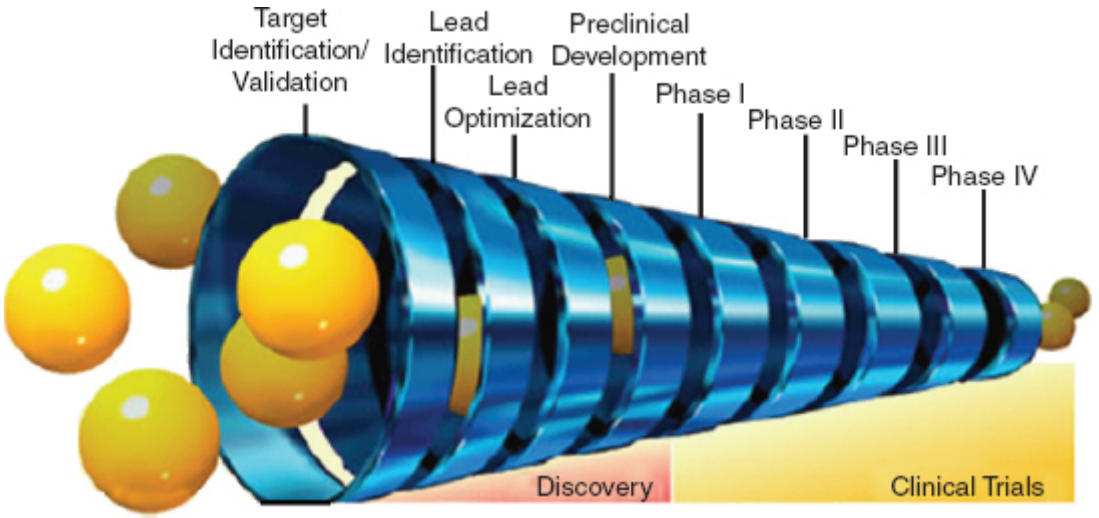


- High content imaging techniques: quantify simultaneously multiple phenotypic and/or functional parameters in biological systems

- …

Can **high-troughput techniques** provide insights in target- and compound-related **risk factors** during **drug discovery**?

When discussions on novel drug targets take place and compound series are identified and optimized?



Case study 1: lead identification

Transcriptomics guiding lead identification

Background of project:

- Inhibition of proto-oncogene tyrosine protein kinase ROS
- Lack of selectivity for compounds of this target class, based on historical information.
- Five chemotypes with some activity identified from HTS.

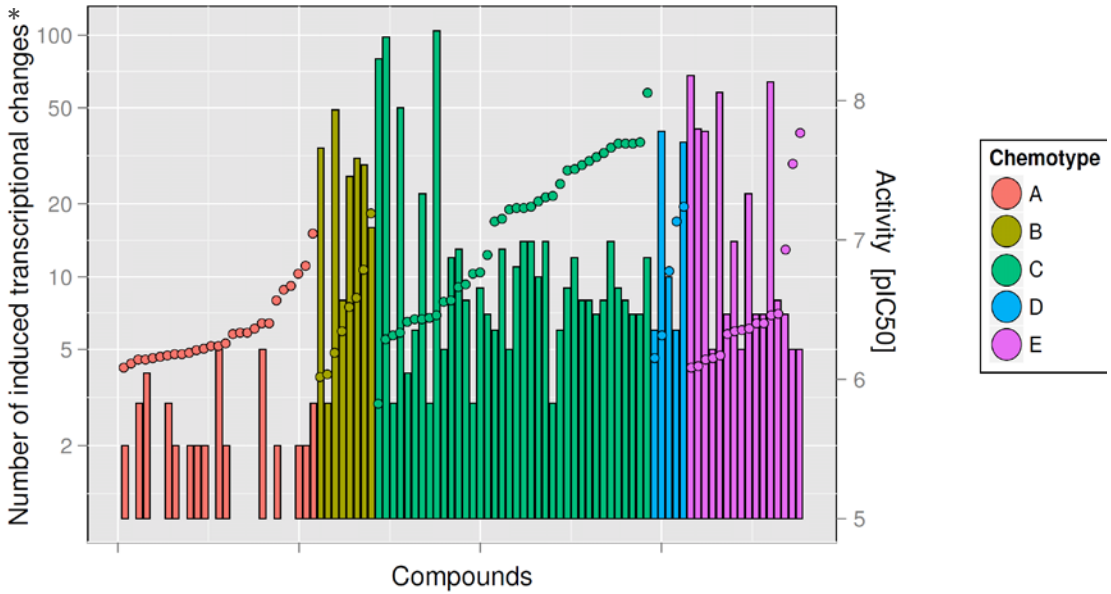
Transcriptomics guiding lead optimization

Transcriptional profiling on:

- HCC78 (non-small cell lung carcinoma) overexpressing ROS1
- Equimolar concentrations (10µM) for 8 hours.
- Microarray hybridizations
- Filtering on informative genes (I/NI calls)

Transcriptomics guiding lead identification

Number of gene expression changes used as a proxy for selectivity of a compound.

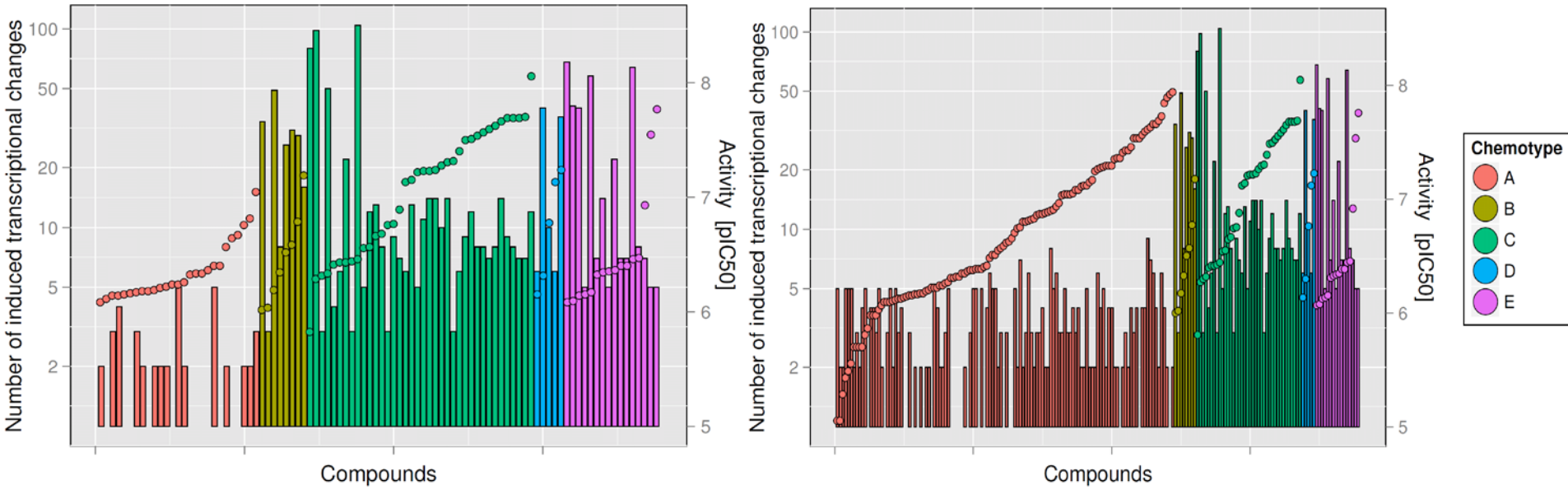


Chemotype A, least number of transcriptional changes but also lower in activity.

*Absolute log fold change > 1, compared to controls

Transcriptomics guiding lead identification

Can inhibitory effect of chemotype A be improved, while keeping the selectivity?



Number of transcriptional effects remain low, while activity improves.

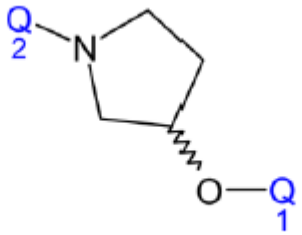
Case study 2: lead optimization

janssen | PHARMACEUTICAL COMPANIES OF Johnson & Johnson

Transcriptomics guiding lead optimization

Background of project:

- Inhibition of PDE10A, novel target for antipsychotics.
- 58 compounds within 1,3-alkoxy-substituted pyrrolidines



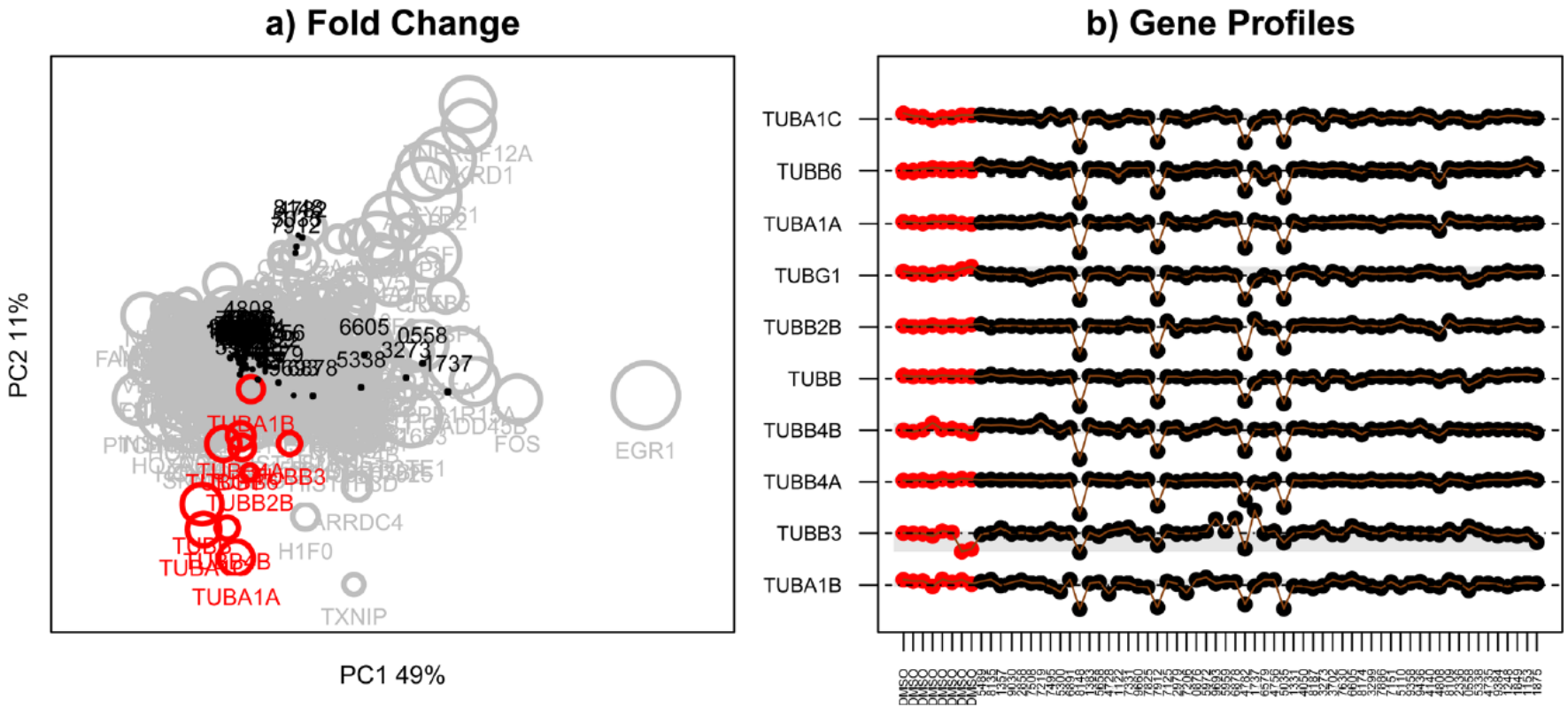
- Adverse effects potential concern.

Transcriptomics guiding lead optimization

Transcriptional profiling on:

- HEK 293 cells transfected with mouse homologue of PDE10A
- Equimolar concentrations (10µM) for 8 hours.
- Microarray hybridizations
- Filtering on informative genes (I/NI calls)
- Unsupervised analysis, spectral map analysis (pca like with double centering to remove size component).

Transcriptomics guiding lead optimization

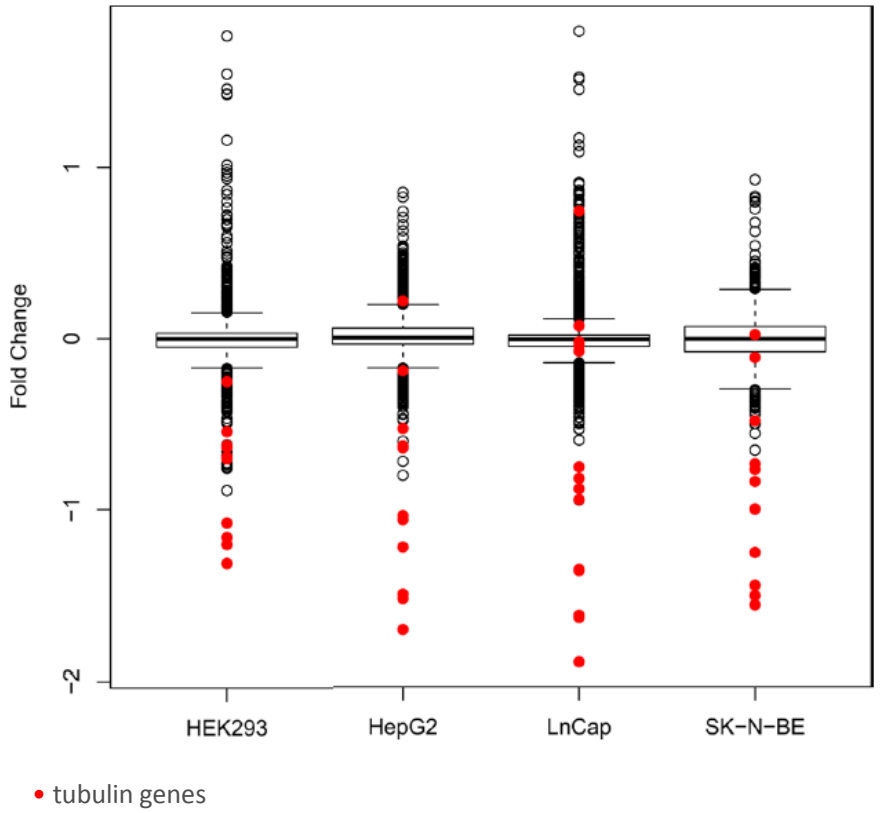


Four compounds (8148, 4782, 5035, 7912) show strong downregulation of tubulin genes.

Transcriptomics guiding lead optimization

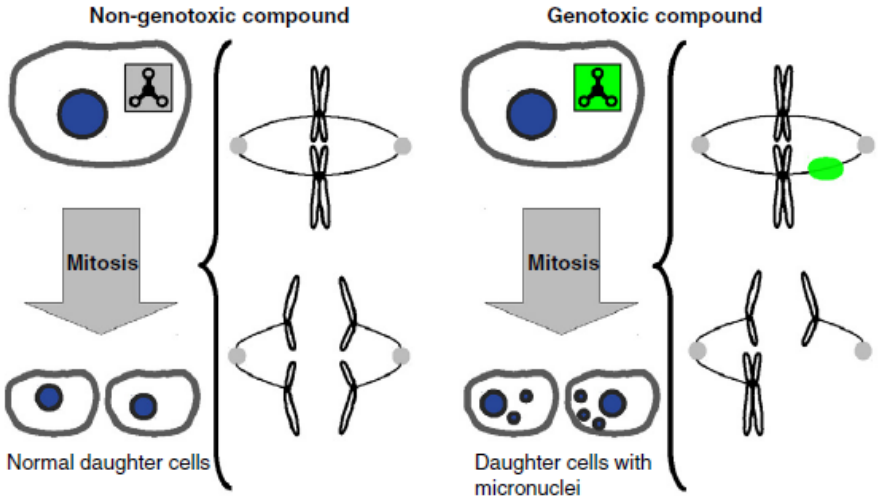
Compound 8148 profiled in different cell lines:

Tubulin genes among most down-regulated genes in different cellular context.



Transcriptomics guiding lead optimization

Tubulins know in context of chromosome segregation, hence downregulation of tubulin genes might suggest a genotoxic effect.

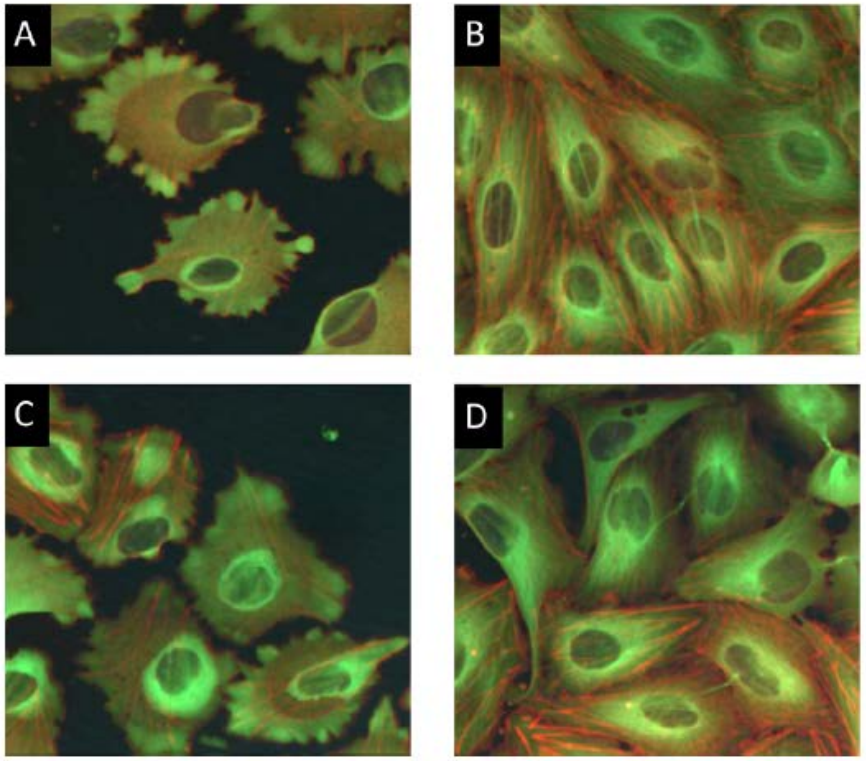


- Traditional micronucleus test (MNT) detects the formation of micronuclei (dose response).
- Positive MNT test > 2 fold count of micronuclei versus vehicle and concentration related increase

High-content imaging guiding lead optimization

High-content imaging experiment on osteosarcoma cells with fluorescent labelled tubulin:

- 11 compounds + genotoxic compounds at different concentrations.
- Genotoxic reference comp. shows microtubule aggregates at the edge.
- 661 features describe the images.



A: Nocodazole
B: DMSO
C: 8148
D: 0558

High-content imaging guiding lead optimization

- Three features extracted, which discriminates best image with microtubule aggregate versus other phenotypes.
- All three features derived from tubulin-GFP channel.
- Linear discriminant analysis to summarize the three featues as a linear combination = microtubule aggregate score
- Positive score is presence of microtubule aggregate

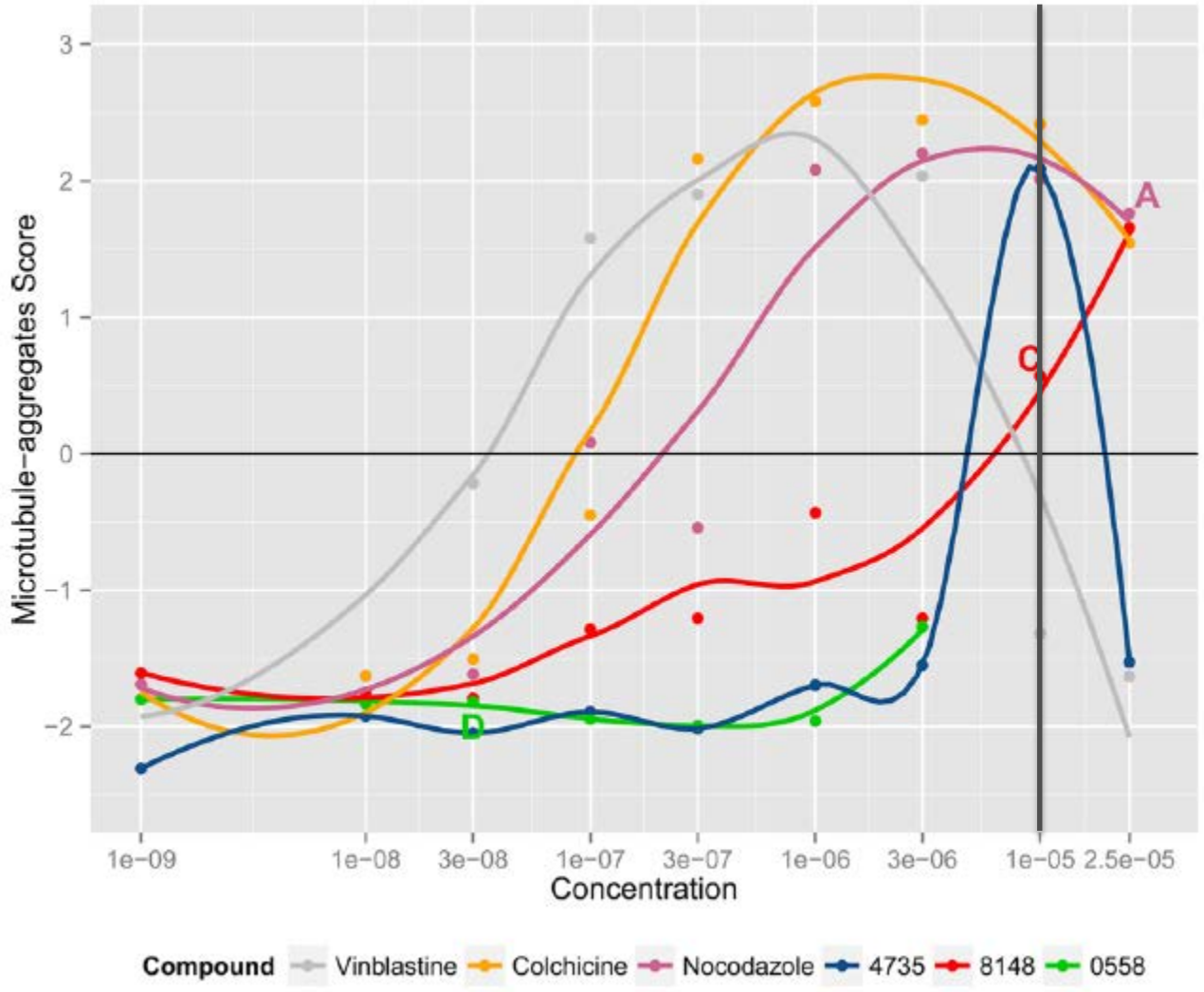
NOTE:

Gene-expression = 10uM

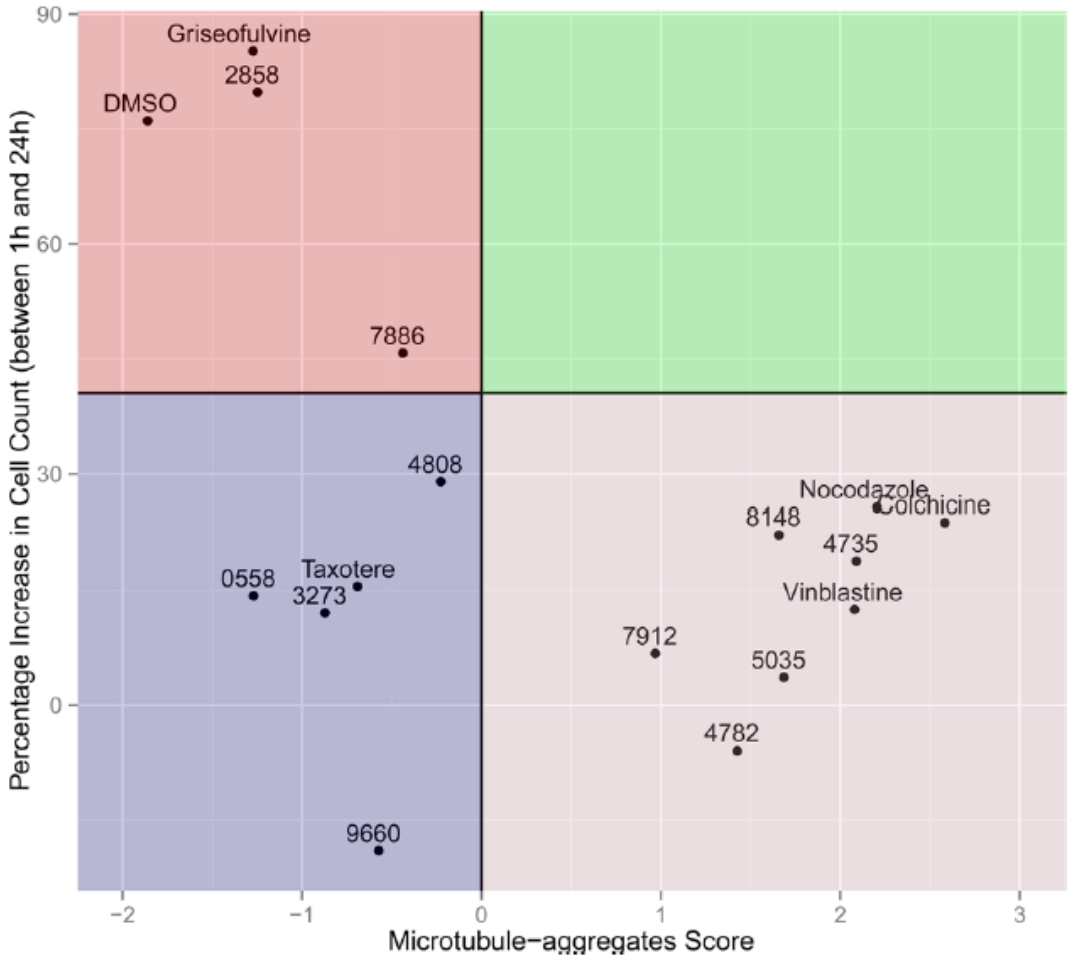
High-content imaging = concentration range (max score is plotted)

Aggregate score in function of concentration:

- Nocodazole and 8148 in agreement
- 4735 only in imaging, but single event (reliability?)
- Cholchicine no transcriptional effect, but retesting showed tubulin down-regulation.
- Vinblastine, bell shaped no aggregates at 10uM but paracrystals



Did we dose high enough to see tox effects ?



Griseofulvine, 2858, and 7886 not dosed high enough

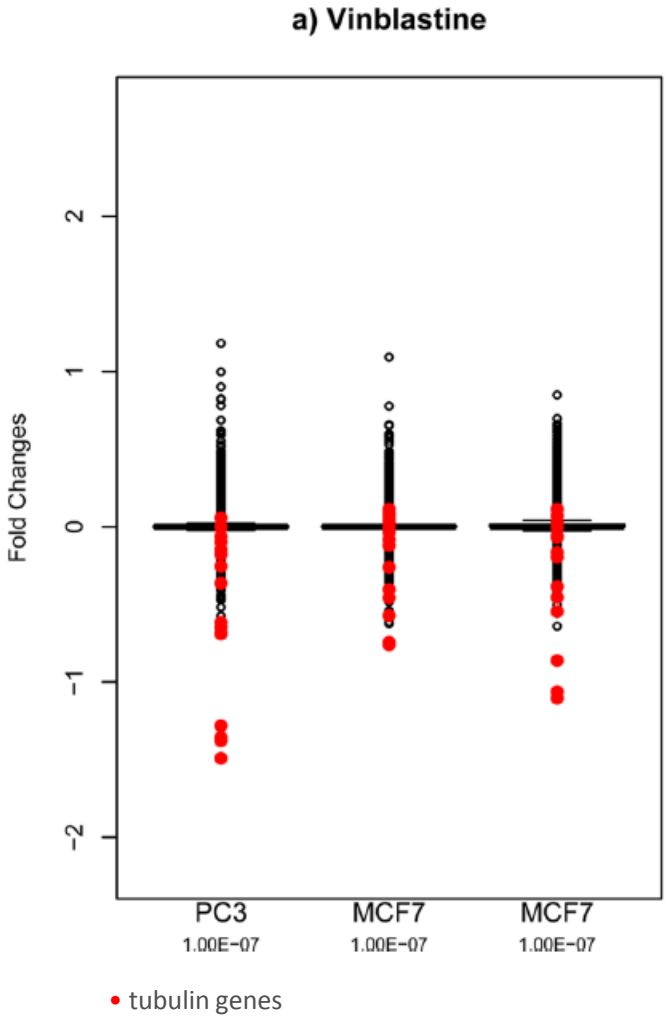
Similar increases (24h) in cell count compared to DMSO

Transcriptomics guiding lead optimization

Connectivity MAP (CMAP):

Database with transcriptional expression data of compound treatment.

- At 0.1µM vinblastine shows tubuline down-regulation



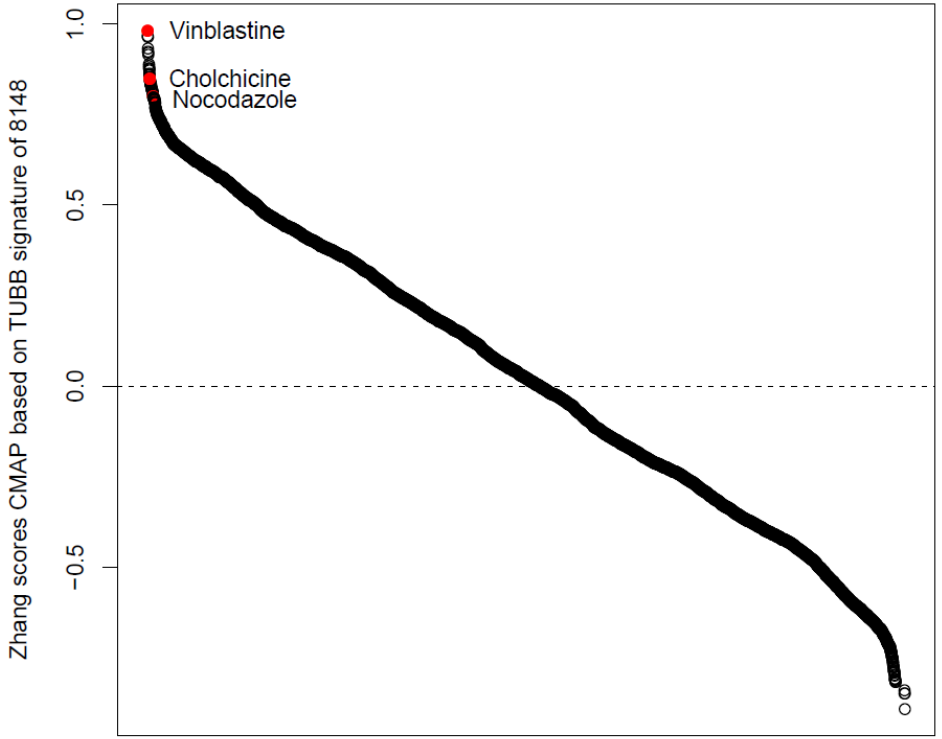
Transcriptomics guiding lead optimization

Connectivity MAP (CMAP):

Connectivity score* (rank based) for tubulin downregulation:

Top compds:

- Vinblastine (0.98)
- Fenbendazole (097, 0.96)
- Chelidonine (0.93)
- Mebendazole (0.93,0.92)



*Zhang and Gant (2008), BMC Bioinf 9, 258

Both transcriptional as well as high-content imaging data could flag potential genotox issues

But not seeing something doesn’t mean it is safe !

- Context dependency
- Time/concentration dependency
- Exposure levels

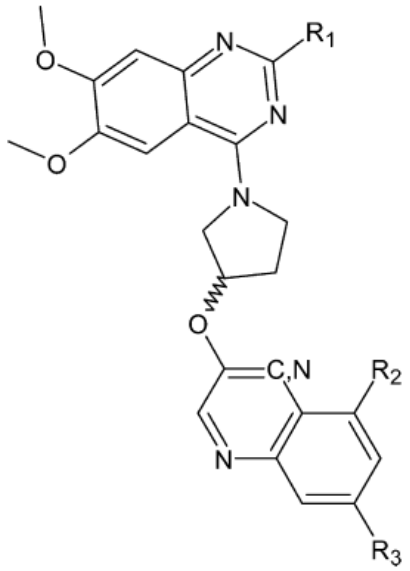
Next question: can we guide chemistry?

Transcriptomics guiding lead optimization

Tubulin downregulation only seen with specific substitution pattern on quinazoline scaffold:

R_1: small or electron-withdrawing

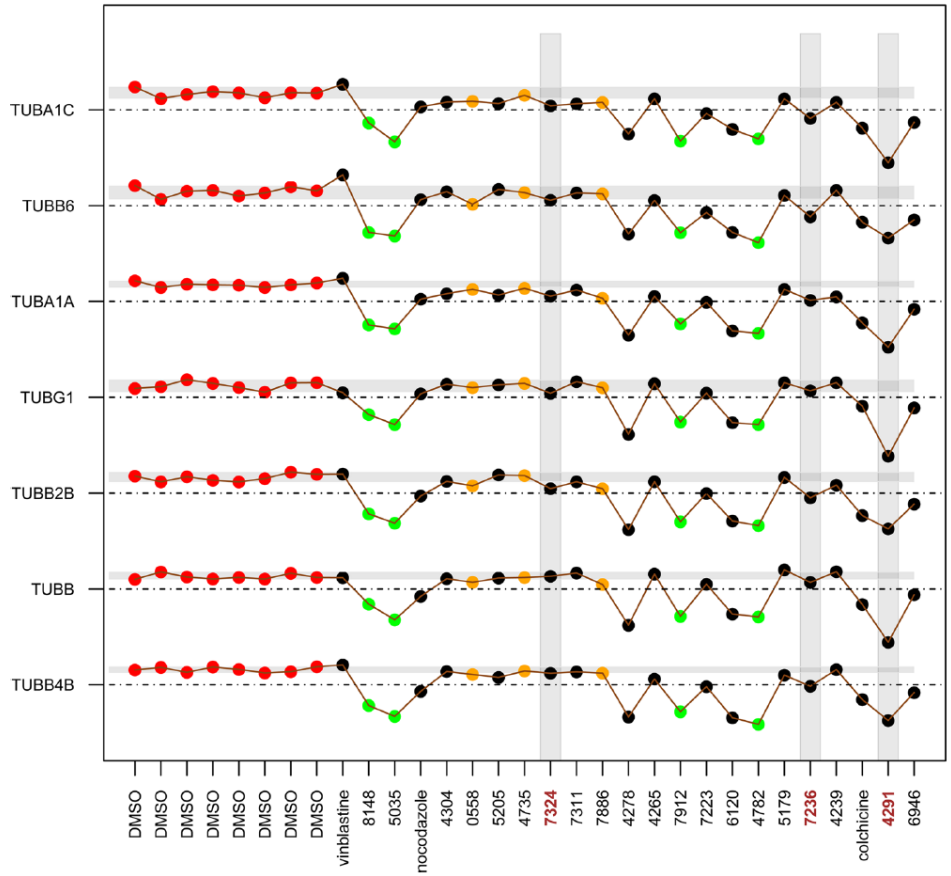
R_2/R_3: no substitution



→ 13 new compounds with high potential of showing tubulin down-regulation !

+ some genotoxic ref. compounds, and 4 pos / 3 neg from earlier experiments.

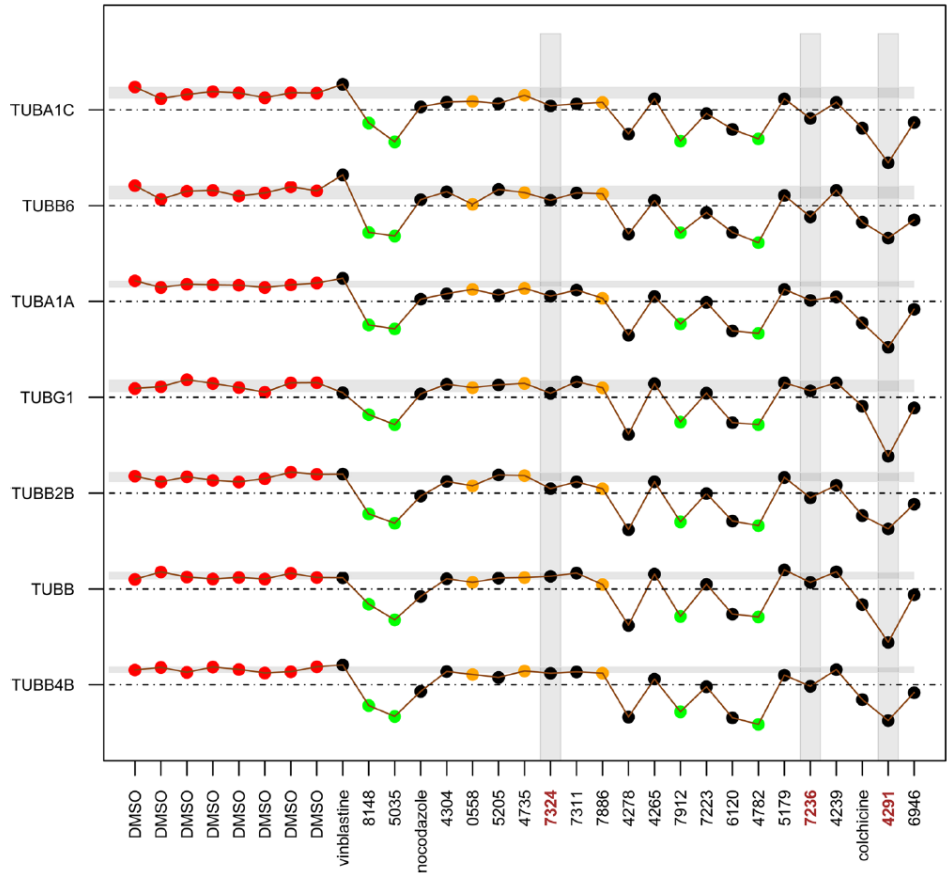
Transcriptomics guiding lead optimization



- Down-regulation in previous experiment
- No down-regulation in previous experiment

- Vinblastine no down-regulation.
- Nocodazole/Colchicine down regulation
- Fair amount of compounds (below grey bar) show downregulation

Transcriptomics guiding lead optimization



- Down-regulation in previous experiment
- No down-regulation in previous experiment

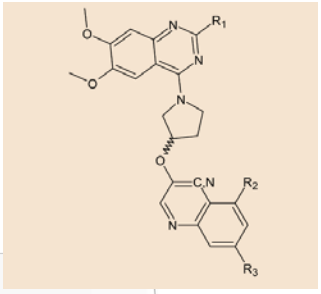
Three compounds (different gradation tubuline downregulation) tested in MNT:

7324 < 7236 < 4291

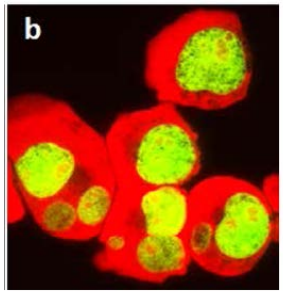
IC50 in MNT: 16.0µM 4.17µM 2.17µM

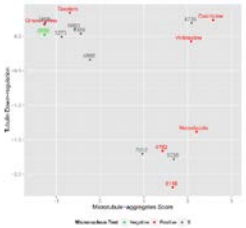
→ For 7324, probably exposure not high enough in transcriptomics

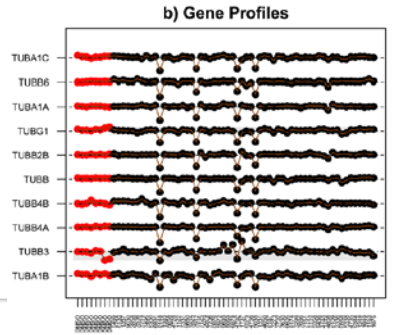
Conclusion:



Structure



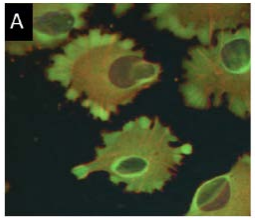




MNT

Transcriptomics

High-Content Imaging



Wrap-up

Assess the utility of high-troughput, high-dimensional data in decision making:

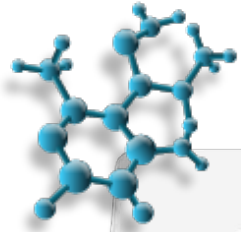
- Transcriptomics as a proxy of selectivity measure
- Early identification of toxicity
- Integrated in traditional (Q)SAR

However clear limitations:

- Context dependency (transcriptomics ≠ changes at protein/metabolite/… level)
- Time/dose/cell line dependency
- Exposure levels for tox effects
- Interpretation of discoverd effects not always straigtforward
  - Risk of introducing more questions than answers.
- Hypothesis generation, validation still required

Wrap-up

Integrated approaches are difficult:



- Highly correlated features
- Bioisosterism

Chemistry:
Quantification

Bio activity:
Sparseness

Gene expression:
Context-dependency

- Time, concentration and cell line
- Interpretation

Acknowledgements

Hinrich W. H. Göhlmann[3], Andreas Bender[5], Sepp Hochreiter[2], Olivier Thas[1], Ziv Shkedy[4], Adetayo Kasim[6], Patrick Angibaud[3], Dhammika Amaratunga[3], Aditya Bhagwat, Theophile Bigirumurame[4], Luc Bijnens[3], Ulrich Bodenhofer[2], Mark Booth[6], Hugo Ceulemans[3], Lieven Clement[1], Djork-Arné Clevert[2], Laure Cougnaud[7], An De Bondt[3], Harrie Gijsen[3], Martin Heusel[2], Tatsiana Khamiakova[4], Günter Klambauer[2], Federico Matiello[1], Andreas Mayr[2], Matthew McCall, Andreas Mitterecker[2], Marjolein Crabbe[3], Steven Osselaer[3], Martin Otava[4], Pieter Peeters[3], Nolen Perualila[4], Aakash C. Ravindranath[5], Marvin Steijaert[7], Willem Talloen[3], Pushpike Thilakarathne[4], Gary Tresadern[3], Thomas Unterthiner[2], Ilse van den Wyngaert[3], Freddy Van Goethem[3], Evelyne Vanneste[9], Herman W. T. van Vlijmen[3], Tobias Verbeke[7], Bie Verbist[3], Geert Verheyen[3], Liesbet Vervoort[3], Alain Visscher[1], Jörg K. Wegner[3], Berthold Wroblowski[3] and Dirk Wuyts[3]

[1]UGent, [2]Johannes Kepler University (Linz, Austria), [3]Janssen Pharmaceutica, [4]Uhasselt, [5]University of Cambridge (UK), [6]Durham University (UK), [7]OpenAnalytics, [8]University of Rochester (USA), [9]CMAST.

Acknowledgements

Hinrich W. H. Göhlmann[3], Andreas Bender[5], Sepp Hochreiter[2], Olivier Thas[1], Ziv Shkedy[4], Adetayo Kasim[6], Patrick Angibaud[3], Dhammika Amaratunga[3], Aditya Bhagwat, Theophile Bigirumurame[4], Luc Bijnens[3], Ulrich Bodenhofer[2], Mark Booth[6], Hugo Ceulemans[3], Lieven Clement[1], Djork-Arné Clevert[2], Laure Cougnaud[7], An De Bondt[3], Harrie Gijsen[3], Martin Heusel[2], Tatsiana Khamiakova[4], Günter Klambauer[2], Federico Matiello[1], Andreas Mayr[2], Matthew McCall, Andreas Mitterecker[2], **Marjolein Crabbe**[3], Steven Osselaer[3], Martin Otava[4], Pieter Peeters[3], Nolen Perualila[4], Aakash C. Ravindranath[5], Marvin Steijaert[7], Willem Talloen[3], Pushpike Thilakarathne[4], Gary Tresadern[3], Thomas Unterthiner[2], Ilse van den Wyngaert[3], **Freddy Van Goethem**[3], Evelyne Vanneste[9], Herman W. T. van Vlijmen[3], Tobias Verbeke[7], Bie Verbist[3], Geert Verheyen[3], Liesbet Vervoort[3], Alain Visscher[1], Jörg K. Wegner[3], Berthold Wroblowski[3] and Dirk Wuyts[3]

References

Verbist, B., Klambauer, G., Vervoort, L., Talloen, W., QSTAR Consortium, Shkedy, Z., Thas, O., Bender, A., Göhlmann, H., & Hochreiter, S. (2015). **Using Transcriptomics to Guide Lead Optimization in Drug Discovery Projects**. *Drug Discovery Today*, 20, 5, 505.

Verbist, B.M.P, Verheyen, G., Vervoort, L., Crabbe, M., Beerens, D., Boxmans, C;, Jaensch, S., Osselaer, S., Talloen, W., Van den Wyngaert, I., Van Hecke, G., Wuyts, D., QSTAR Constorium, Van Goethem, F., Göhlmann, H.W.H. (2015) **Integrating High Dimensional Transcriptomics and Image Analysis Tools into Early Safety Screening: Proof of Concept for a New Early Drug Development Strategy**, 28, 1914.

Göhlmann, H., Verbist, B., Crabbe, M., Van Goethem,F., **Addressing genotoxicity risk: A PDE10A inhibitor case study** chapter in Early Drug Development:Bringing a Preclinical Candidate to the Clinic (2018), Wiley (in press)