PROJECT FINAL REPORT

Grant Agreement number: 223451

Project acronym: PHAGOSYS

Project title: Systems biology of phagosome formation and maturation, modulation by intracellular pathogens

Funding Scheme: FP7-CP-FP HEALTH-2007-2.1.2-5

Period covered: **from** 01-Nov-2008 **to** 30-Apr-2012

Name of the scientific representative of the project's co-ordinator, Title and Organisation:

Dr. Brian D. Robertson

Imperial College of Science Technology and Medicine,

Tel: +44-20-7594-3198

Fax: +44-20-7594-3095

E-mail: b.robertson@imperial.ac.uk

Project website address: www3.imperial.ac.uk/cisbic/research/phagosys

Part 1: Executive summary

One of the first lines of defence against bacterial pathogens that enter our bodies are phagocytes such as macrophages and monocytes, that literally eat the bacteria and destroy them. However some pathogenic bacteria, including *Salmonella* and *Mycobacterium tuberculosis*, have evolved mechanisms whereby they can survive inside these very cells; one way in which they do this is to inhibit the endosomal fusion processes inside the macrophage that expose the bacteria to lethal factors. The aim of this project was to identify the mechanisms bacteria use to block killing by macrophages, and by doing so find ways to overcome these blocks and promote intracellular killing. Helping the macrophage to do its job in this way could be the basis for therapeutic treatments that would be added to conventional anti-microbial drug treatment. This could help shorten treatment times for diseases like tuberculosis, which currently requires 6-9 months of therapy. Since this type of therapy would enhance macrophage function as opposed to killing bacteria, it would be unlikely to lead to the emergence of further microbial drug resistance.

This project targeted two areas of intracellular pathogen biology: the initial uptake into phagocytic cells, which may impact on the subsequent trafficking, and the later phagosome-lysosome fusion events, which are manipulated by *Mycobacterium tuberculosis* and *Salmonella* to promote their survival. We have used state-of-the-art cell-based screening techniques with libraries of small-inhibitory RNA (siRNA) molecules and pharmaceutically active compounds to identify potential macrophage targets, and combined this with the power of systems biology to analyse and integrate the different data sets. We have identifying genetic targets that can be manipulated by siRNA and enzymes that can be modulated with small molecules, which will be taken forward in studies to investigate their suitability as targets and drugs for antimicrobial therapy. A spin-off has been the identification of compounds that interfere with eukaryotic cell proliferation, and may have a role in cancer therapy. We identified bacterial mutants with altered trafficking in human macrophages, and these will form the basis of future work to examine the host macrophage transcriptional response to mutant and wild type bacteria.

We further developed methods for the automated modelling of dynamic systems (in the form of ordinary differential equation) using experimental data and expert knowledge, a typical task in systems biology modelling. We used these automated methods to model endosome maturation and LDL trafficking, using time-course data on protein concentrations. We also applied machine learning methods for structured output prediction and predictive clustering to a large number of datasets generated from high-throughput screens carried out during the project.

Molecules including small GTPases and phosphoinositides regulate phagosome maturation. We identified a surprising interaction between apparently antagonistic enzymes, and used mathematical modelling to investigate this. We found that complex formation can produce novel forms of switch-like and bell-shaped responses, and postulate this could have a functional role in the temporal regulation of phosphoinositides during phagosome maturation.

Part 2: Summary description of project context and objectives

Project context

Most bacteria that enter a host organism, by ingestion or inhalation, are engulfed and killed by phagocytic cells such as macrophages. However some pathogenic bacteria, including *Salmonella* and *Mycobacterium tuberculosis*, have evolved mechanisms whereby they can survive inside the very cells that our bodies have developed to kill the pathogens. The overall aim of this project was to try and identify how these bacteria prevent killing by macrophages, and by so doing find ways to overcome this block and promote intracellular killing. Helping the macrophage to do its job in this way could be the basis for therapeutic treatments that would be added to conventional anti-microbial drug treatment. This could help shorten treatment times for diseases like tuberculosis, which currently requires 6-9 months of therapy, and since this would enhance macrophage function as opposed to killing bacteria, would be unlikely to lead to the emergence of further drug resistance.

This project was designed to target two areas of intracellular pathogen biology; the initial uptake into phagocytic cells, which may impact on the subsequent trafficking, and the later phagosome-lysosome fusion events which are manipulated by *M. tuberculosis* and *Salmonella* to promote their survival. We have used state-of-the-art cell-based screening techniques with libraries of small-inhibitory RNA (siRNA) molecules and pharmaceutically active compounds to identify potential macrophage targets, and used the power of systems biology to analyse and integrate the different data sets. We report here on our success in identifying genetic targets that can be manipulated by siRNA and enzymes that can be modulated with small molecules, coupled with improved systems biology techniques that allow the interrogation of complex data sets.

Project objectives

The first objective of the project was to develop models of phagocytosis, based on existing expert knowledge and on data collected within the project or already available in the literature. This required the development of methods for automated modelling able to make use of both data and expert knowledge. Major progress was made towards achieving these goals, but some changes were required due to unforeseeable personnel issues, which meant that the scope of the models constructed was more limited than originally planned. We used these automated methods to model endosome maturation and LDL trafficking, using time-course data on protein concentrations. We also applied machine learning methods for structured output prediction and predictive clustering to a large number of datasets generated from high-throughput screens carried out by the project partners. Molecules including small GTPases and phosphoinositides regulate phagosome maturation. We identified a surprising interaction between apparently antagonistic enzymes, and used mathematical modelling to investigate this. We found that complex formation can produce novel forms of switch-like and bell-shaped responses, and postulate this could have a functional role in the temporal regulation of phosphoinositides during phagosome maturation.

The experimental objectives of the project ran in parallel to the modelling and we set out to characterise the bacteria-containing compartments, and determine the kinetics of intracellular phagosomal and endosomal trafficking. This was done using a combination of high throughput fluorescence microscopy cell-based assays and a novel quantitative multiparametric analysis developed at MPI-CBG. We successfully developed the required high-throughput, high-content screening assays using both murine and the more physiologically relevant primary human macrophages. The mycobacterial infection assay was optimized, combining recombinant *Mycobacterium bovis* bacillus Calmette-Guerin (BCG) expressing Green Fluorescent Protein, with LysoTracker® Red DND-99 staining of host-cell lysosomes. This provided a direct readout of the release of the phagosomal maturation block and degradative delivery of the bacteria to lysosomes when the system was perturbed with either siRNAs or small molecules.

The optimised assays and protocols were then used to identify, using RNAi or small molecules, the key regulators of the host cellular machinery, e.g. kinases, small GTP-binding proteins and their effectors, which are required to control infection of host cells, intracellular trafficking and growth or

clearance of non-pathogenic and pathogenic *Salmonella* and mycobacteria. It proved impracticable to cover the entire druggable genome of ~7000 targets, but using more targeted chemical and genetic approaches we have identified a number of promising hits that act on the host-cell and not the bacterium, and which are now being taken forward in studies that will continue beyond this project.

We also set out to identify the bacterial components that mediate the interactions with the macrophage, preventing phagosome-lysosome fusion, and therefore promoting bacterial survival. The approach was to make use of the mycobacterial mutants that have been selected on the basis of their inability to block phagosome-lysosome fusion during macrophage infection *in vitro,* and reported in the literature. A panel of mutants was collected and tested in human type 1 and type 2 macrophages, but only four mutants showed the same phenotype as originally reported. This is probably due to differences in the type of macrophage used in the original screens; murine or transformed cell lines are commonly used, but appear to differ fundamentally in the way they traffic intracellular mycobacteria. The next step will be to compare the transcriptional response of the macrophage to infection with wild type and mutant bacteria.

Overall we made significant progress towards achieving the main objective of the program, which was to use high-throughput genetic and chemical screens to identify host-cell targets that can be manipulated to overcome the phagosome-lysosome fusion block put in place by intracellular pathogens. To do this we optimised screening protocols, data analysis methods and systems biology with which to integrate and interrogate the large volumes of data generated. Target and pathway identification is underway and will continue beyond the life of the current project.

Part 3: Main S&T results/foregrounds

WP1 – Modelling Phagocytosis

The overall goal of WP1 was to develop models of phagocytosis, based on existing expert knowledge (ICTSM) and on data (JSI); both existing data and data collected within the project were to be used for this purpose. In addition, methods for automated modelling that can make use of both data and expert knowledge were to be further developed to meet the typical needs of systems biology modelling tasks (JSI).

Major progress was made towards achieving these goals, but some changes were inevitable due to unforeseeable circumstances. At ICTSM, WP1 experienced problems that affected the progress of the project including long-term illness and subsequent tragic death of the mathematician Professor Jaroslav Stark, and the late start and early departure of the research fellow at ICSTM, Dr Barbara Szomolay (April 2009-October 2011). A biomathematician, Dr. Vahid Shaherezei, was recruited to the project and managed the ICSTM side of WP1. Due to this, the scope of the models constructed is more limited than originally foreseen.

At JSI, the scope of WP1 was extended to include the analysis of the majority of data generated within the project. It was originally planned to use time-course data on protein concentrations to learn models of the dynamics of phagocytosis, however the limited availability of this kind of data meant we were not able to do as much of this as originally foreseen. Since the majority of the data generated in the project included different types of screens (genome or compound screens, flow-cytometry or image-based), a large amount of effort was devoted to integrating and analysing these, which is reflected in the outcomes of the project.

At ICSTM, we first compiled a signalling network for phagosome maturation based on an extensive literature survey. Based on the complexity of the initial network and the limited available data, we concluded that construction of a comprehensive mathematical model of this system is not yet feasible. We then decided to focus on particular aspects of this system and model them in detail.

Phagosome maturation is regulated by the key tagging molecules Rab proteins and phosphoinositides (PIPs). Inspecting the signalling network that regulates PIPs revealed a surprising interaction between antagonistic enzymes, including interactions between Vps34 kinase-myotubularin phosphatases (see Figure WP1-1).

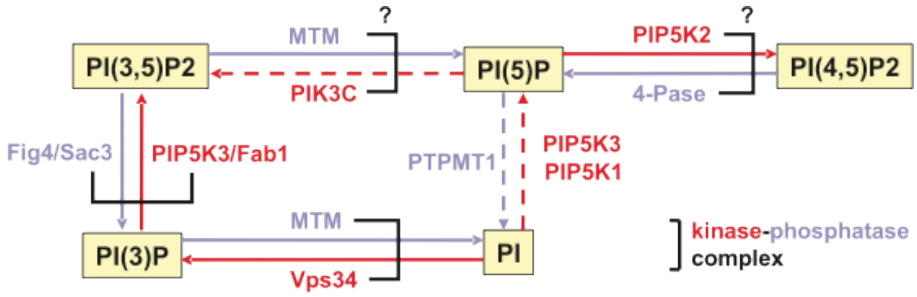


Figure WP1-1: Identified or hypothesized kinase-phosphatase complexes (PIK3C – class I PI 3-kinase, PTMT1 – PTEN-like phosphatase). Full lines represent *in vivo*, dashed line *in vitro* data.

Using mathematical modelling we investigated the role of these interactions, and found that the complex formation can produce novel forms of switch-like and bell-shaped response (Figure WP1-2). We have postulated that these could have a functional role in the temporal regulation of PIPs during phagosome maturation (Szomolay and Shahrezaei, 2012).

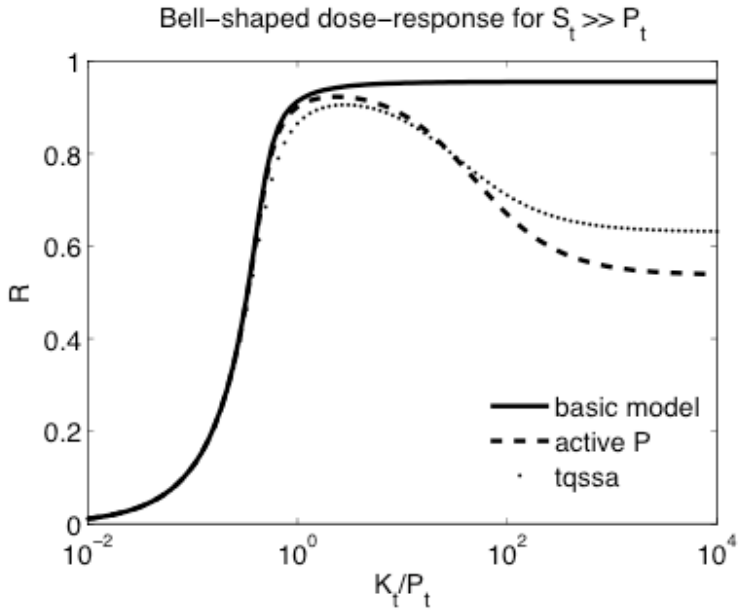


Figure WP1-2: Bell-shaped dose response for the active phosphatase case in the zero-order regime. Kt, Pt, St are the total kinase, phosphatase, and substrate concentrations, respectively. R denotes the fraction of total phosphorylated substrate. The dashed line represents the response curve for the extended model in the active phosphatase case and the solid line represents the response curve for the corresponding basic model of a phosphorylation-dephosphorylation cycle.

We have been also involved in statistical analysis of multiparametric imaging data (Collinet et al., 2010) from our partners at MPI-CBG, describing the intensity of EEA1 and APPL1 endosomes (Rab5 effector proteins necessary for endosomal trafficking) at the single endosome level for wild type and a range of RNAi knockdowns. In addition to MATLAB, image analysis software (Motion Tracking, developed by Y. Kaladzidis in the Zerial lab, Dresden) was used to analyse the data. A recent study has shown that APPL1-positive and EEA1-positive endosomes have a transient overlap (APPL1 endosomes represent an earlier stage than EEA1 endosomes) and that the switch from the APPL1 to the EEA1 stage is controlled by an important phospholipid, PI(3)P (Zoncu et al., 2009). One goal is to show that the single endosome data is consistent with this finding, but so far we have not found strong evidence for this. In addition we are looking at the heterogeneity of the markers and how this is regulated. We hope to add insights into the mechanisms of robustness in endosome trafficking. This is still an ongoing work in collaboration with MPI-CBG and we plan to publish these results as soon as all the analyses are finished.

At JSI, we pursued two major lines of research. First, we continued the development of methods for the automated modelling of dynamic systems (in the form of ordinary differential equations; ODEs) that can make use of both experimental data and expert knowledge, and so meet the needs of typical systems biology modelling tasks. We also investigated the use of these automated methods for modelling endosome maturation and LDL trafficking, where time-course data on protein concentrations was available. Second, we applied machine learning methods for structured output prediction, and in particular predictive clustering, to a large number of datasets generated from high-throughput screens related to phagocytosis of intracellular pathogens carried out by the project partners.

We have made significant advances in machine learning methods for automated modelling of dynamic systems (Cerepnalkoski et al., 2012), not only relevant to the automated modelling of phagocytosis but also addressing other modelling tasks in systems biology. First, an extended formalism for encoding domain-specific knowledge for automated modelling was developed. It is based on entities and processes, and modelling templates thereof: it facilitates the grouping of entities and processes into compartments, an important concept in systems biology. Second, a new software platform was implemented that supports the learning of models in this formalism. This software platform allows for the use of different local and global optimization methods for parameter estimation in ODE models. It also allows the use of different objective functions (in addition to the sum of squared errors) within these optimization methods.

The utility of the improved methods was evaluated within the simple context of estimating the parameters of a single model structure and a more challenging setting of selecting an appropriate model structure. For the single model structure, we used the model of endosome maturation proposed by del Conte-Zerial *et al* (Del Conte-Zerial et al., 2008). We compared the performance

of local derivative-based and global meta-heuristic optimization methods for parameter estimation under different observation scenarios and varying level of noise in the data. The scenarios cover a wide range of situations, from the simplest one of complete observability, where the concentrations of all state variables are assumed to be directly measurable, to the most complex (and realistic) scenario, where the observations are linear combinations of state variables (total concentration of protein active and passive states). Global meta-heuristic methods clearly and significantly outperform the local derivative-based method. These results hold for both real and artificial data, for all observation scenarios considered, and for all amounts of noise in the data (Tashkova et al., 2011).

We also considered the use of different optimization methods in the context of estimating the parameters in a large number of model structures, which is typical for automated modelling approaches (Cerepnalkoski et al., 2012). We compare a global and a local optimization method, as used in the context of the automated modelling tool, on four modelling tasks, each involving a large number of model structures. We considered the difference in performance of the best models found, the overall quality of all models considered, and each individual model considered; across all of these, the global optimization method used in the context of automated modelling performs much better than the local optimization method. For illustration, consider the error profiles of the models with parameter values fitted by each of the two methods, shown in increasing error order (Figure WP1-3): These clearly show the superiority of global optimization. The use of global optimization for parameter estimation eliminates a major weakness of automated modelling methods that we identified during the first phase of the project.

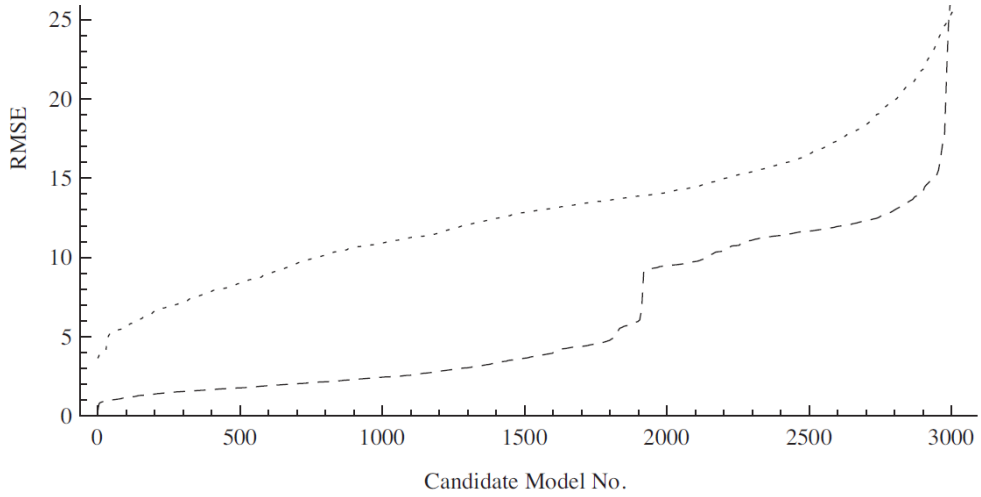


Figure WP1-3: Error profile curves for ODE model structures whose parameters have been fitted by a local (top curve) and a global (bottom curve) optimization method. Much better models are found when using global optimization methods for parameter fitting.

With the new software platform ready, in collaboration with MPI-CBG Dresden we formulated a library of entity and process templates for modelling Low-Density-Lipoprotein (LDL) trafficking, based on the domain knowledge shown in Figure WP1-4. Time-course profile data of LDL concentrations in different compartments (vesicular/non-vesicular) and stages (before/on/after) Rab5 were also provided, measured under stimulation (10 min. pulses) with different concentrations of EGF (from 30 to 3000 ng/ml). Good fit of the model with the estimated parameters to the measured data was achieved, both for the case of a single model structure and several alternative model structures.

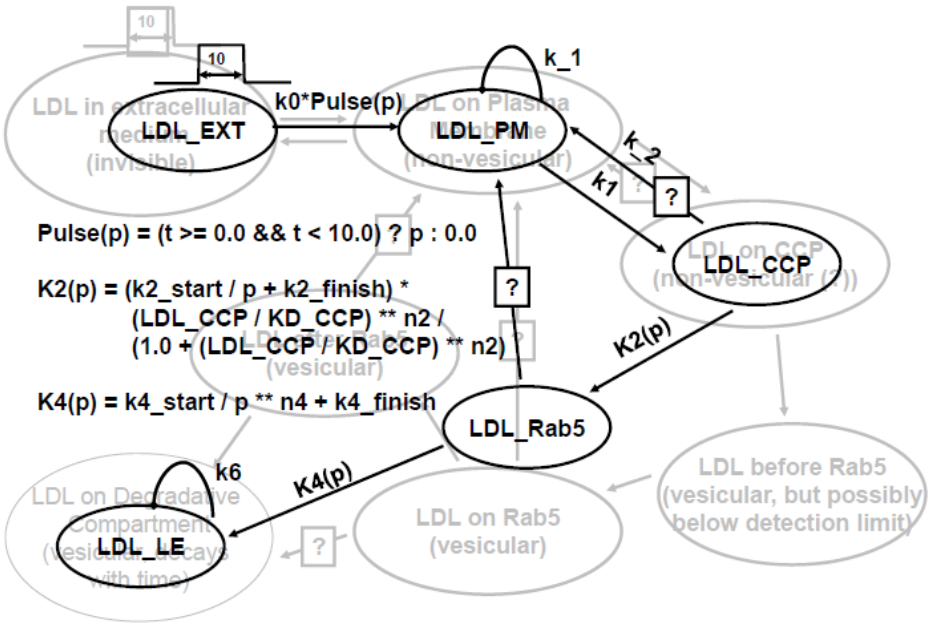


Figure WP1-4: Domain knowledge about modelling LDL traffic (grey), and the portion of knowledge encoded within the process-based modelling library (black).

In the first half of the project we started using different machine learning methods to analyse data provided by the PHAGOSYS partners, with the aim of gaining a better understanding of the process of phagocytosis. In particular, we used methods for structured output prediction, such as predictive clustering trees, to analyse a variety of data related to phagocytosis of mycobacteria. These included microarray data with time course profiles of gene expression levels in response to infection: Type 1 and Type 2 human macrophages were infected with *M. tuberculosis,* and Schwann cells were infected with *M. leprae* (data from LUMC). Interesting groups of genes with common functional annotations and clearly defined response patterns were identified for Schwann cells; the gene network showing some of these groups and the corresponding functions is shown in Figure WP1-5. This enabled the formulation of a hypothesis about a new candidate gene, previously unknown in the regulation of intercellular growth of mycobacteria. The hypothesis was subsequently investigated in experimentally and silencing of the gene was found to reduce bacterial load significantly, as predicted.

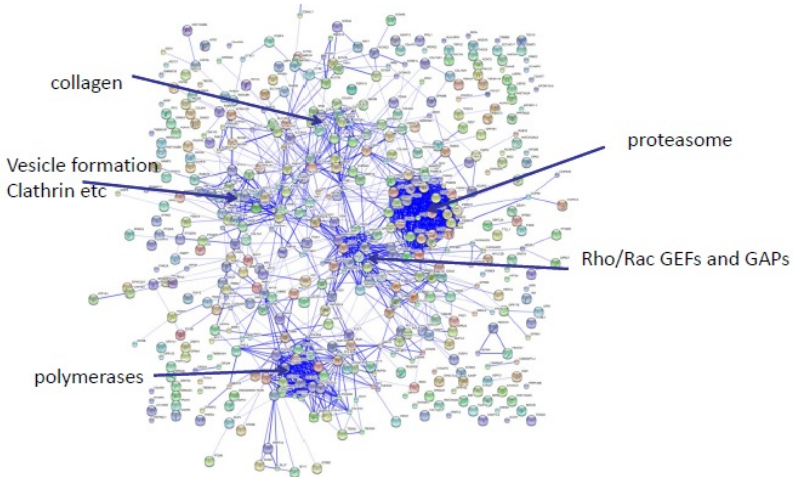


Figure WP1-5: The gene interaction network (from STRING) showing the groups of genes with common functional annotations and clearly defined response patterns for Schwann cells infected with *M. leprae*.

We have then used predictive clustering trees, as well as feature ranking methods, to analyse data from several screens collected by different PHAGOSYS partners. These included an siRNA image-based screen of endocytosis in HeLa cells (data from MPI-CBG), a flow cytometry screen and a smaller image-based siRNA screen aimed at studying MHC Class II antigen presentation (data from NKI), which we tried to inter-relate by our analyses. Finally, we analysed data from a large compound screen (from LUMC) of MelJuSo cells infected with *M. tuberculosis* and *S. typhimurium*.

We analysed data from a genome-wide siRNA image screen performed at MPI-CBG to study the process of endocytosis in the HeLa cell line. The phenotype images obtained after silencing each

gene were analysed using MotionTracking (MPI-CBG in-house software) to produce feature-based descriptions of the images. We described each gene with its functional annotation (using GO terms). A total of 20346 genes were analysed, 3678 of unknown function. The goal of the study was to provide knowledge about endocytosis and to annotate the hypothetical genes with possible functions using the methodology developed by Kocev (Kocev, 2011).

The NKI provided data from two screens concerning MHC Class II antigen presentation. The first screen includes flow cytometry data for 16675 genes, silenced in the MelJuSo cell line, where cell were stained with two fluorescent antibodies, L243 and CerCLIP. Then, 276 candidate genes from this screen (selected by a z-score cut-off) were used for the second screen, where images of the cells were taken and phenotypes analysed with the CellProfiler software to extract features. Using (semi)manual clustering on these features, 19 bins were identified (Paul et al., 2011), but a number of genes/images were left unassigned. We described each gene with its GO annotations and 13 CellProfiler features.

We performed several analyses in order to merge the data from these three studies (MPI-CBG image screen and the NKI's flow cytometry and image screens) and identify some false negatives, possibly resulting from the z-score cutoff in the flow cytometry study. More precisely, we aimed to find genes (from the MPI-CBG genome wide study) that were not selected for in the NKI imaging study, due to the stringent definition of a hit (a z-score larger than 2 or smaller than -2). Figure WP1-6 depicts the intersection of the genes analysed in the three studies. We conducted two types of analyses: gene set enrichment using SEGS-BioMine (Podpečan et al., 2011), and then predicted the possible assignment of the genes from MPI-CBG data to the bins from the NKI study.

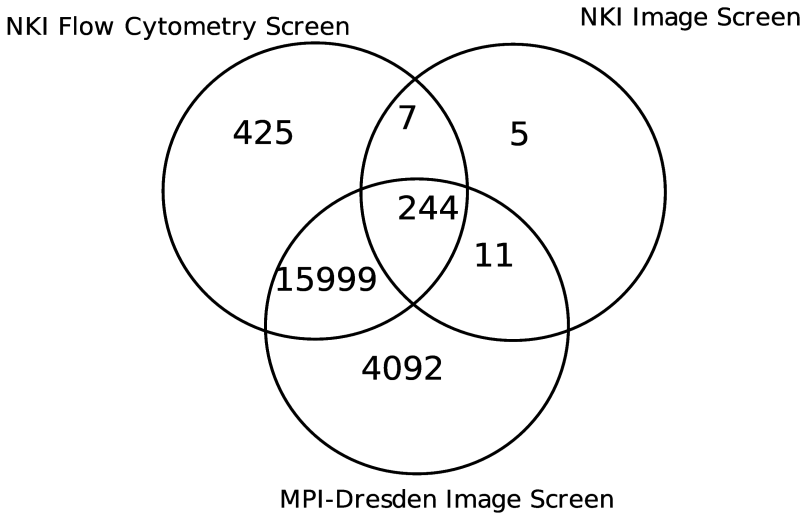


Figure WP1-6: The overlap of the genes from the three studies.

Using gene set enrichment, we discovered gene functions for the highly ranked genes from the flow cytometry study, dependent on the antibody used. For the L234 antibody, most prominent are genes involved in mRNA metabolic processes (GO:0016071), and genes that interact with genes active in processes in the endocytic vesicle membrane (GO:0030666) and the PPAR signalling pathway (KEGG:03320). For the CerCLIP antibody, most prominent are the genes involved in metabolic processes for cellular macromolecules (GO:0044260) and nucleic acids (GO:0090304), intracellular membrane-bounded organelles (GO:0043231), the spliceosomes (KEGG:03040), and genes that interact with genes involved in nuclear mRNA splicing, via the spliceosome (GO:0000398), and processes in the nuclear body (GO:0016604).

The second analysis investigated the possibility of merging the three studies, hoping that the analysis performed in the genome-wide study at MPI-CBG could help elucidate some genes that were found to be negatives in the NKI flow cytometry study. We first predicted the phenotype bins from the CellProfiler features using the NKI screen, the flow cytometry scores and the GO annotations. The results show that we can predict the values for the bins with limited success. We then predicted the phenotype bins using the MotionTracker features from MPI-CBG screen, the flow cytometry scores and the GO annotations. However the results showed that the predictive performance is quite low. Information about the phenotypes from the MPI-CBG study cannot be

(directly and in a straight-forward manner) connected to the phenotype bins defined by the NKI phenotype study. There are several factors that contribute to this outcome: the use of different cell-lines (HeLa *versus* MelJuSo), the different cellular mechanisms studies (endocytosis *versus* antigen processing), and different imaging and image description techniques (MotionTracking *versus* CellProfiler). The integration of the screen data would be more successful if a smaller subset of genes from the NKI study was selected, and the closest matches from the MPI-CBG study calculated. This is the topic of an ongoing collaboration between JSI and NKI.

In collaboration with LUMC, we analysed data from a compound screen they had conducted. The screen used the LOPAC library and measured the (reduction in) bacterial load with flow cytometry (FACS) in two related studies. In the first study, MelJuSo cells were infected with *M. tuberculosis* at a multiplicity of infection (MOI) of 10 bacteria per cell, while in the second study, HeLa cells were infected with *S. typhimurium* also at an MOI=10. The main goal of the analysis is to elucidate protein targets that contribute to the reduction/increase of the bacterial load, i.e., proteins that are involved in processes that promote bacterial killing or growth in the host cells.

We first performed data pre-processing, searching for descriptors of the 1260 compounds from the LOPAC library. We linked each of the compounds to the PubChem database at NCBI (pubchem.ncbi.nlm.nih.gov), and then extracted the proteins that were targeted by each compound in a human study, considering only studies where the compound was found to be active. In total, 711 protein targets were identified. We constructed three different descriptions of the dataset. First, we described 964 compounds through their respective protein targets (there are no protein targets for 296 compounds). Second, we described the compounds using the same GO terms as their respective protein targets. Third, we described the compounds with the GO terms of the target proteins and included gene-to-gene interaction information via the functions (terms) annotated with the interacting proteins. We then performed three different types of analyses on the three representations of the data.

First, we conducted a discriminative analysis of the targets based on the hit compounds. Second, we performed analysis using predictive clustering trees trying to predict the z-score for reduction of the bacterial load using the three data representations. Third, we performed feature ranking using the three data representations.

For the discriminative analysis of the targets based on the hit compounds, we first reduced the number of targets under consideration, because the number of 711 target proteins is large and includes proteins that are targeted by a large variety of compounds. These proteins are not necessarily specific or directly relevant for the response to *M. tuberculosis* and *S. typhimurium* infection. More interesting are the protein targets that were found active in studies involving hit compounds (compounds with z-score larger than 2 or smaller than -2) and not active in the remainder of the compounds. To this end, we pruned the set of the proteins using the following rule: a protein target is removed from the set of protein targets if > 90% of the compounds that it is targeted by are not identified as hits.

Using this reduced set of proteins, we formulated three separate lists of compounds and their respective protein targets. The first list gives 60 protein targets that are obtained when the hit compounds were defined by using both the *M. tuberculosis* and *S. typhimurium* z-score values. The second list gives 58 protein targets that are relevant for the hit compounds for the *M. tuberculosis* study. The third list gives 15 protein targets relevant for the hit compounds for the *S. typhimurium* study. For these gene lists, we obtained three gene networks using the STRING database of gene/protein interactions.

Figure WP1-7 depicts the gene network for the first gene list. In this network, there are three parts. The largest part is positioned in the centre of the network and includes different kinases. The second part is at the right-hand side of the network and includes a group of somatostatin receptors, while the third part is at the left-hand side of the network and includes cell division cycle homologues, which are connected to the largest part of the network through topoisomerases.

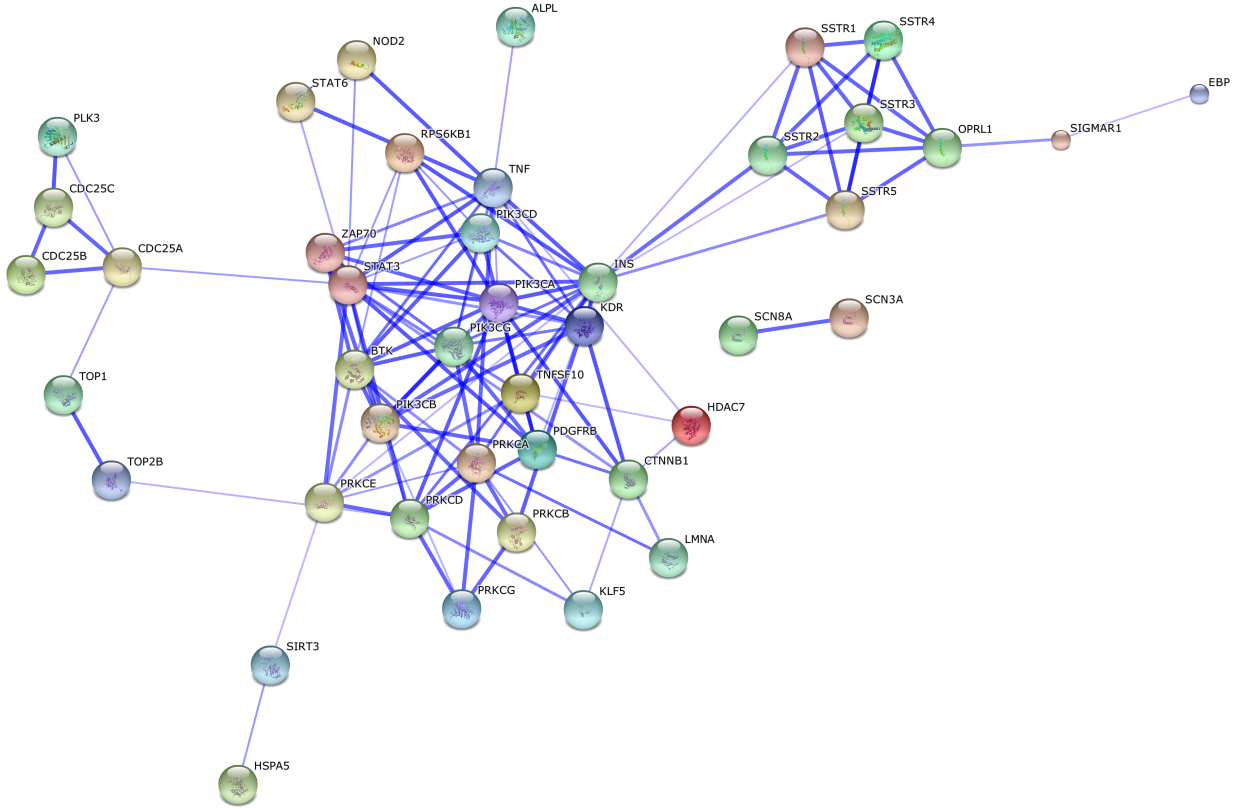


Figure WP1-7: The gene network (STRING-based) for the gene list with hits for both *M. tuberculosis* and *Salmonella typhimurium*.

The second analysis involved the construction of predictive clustering trees on the three data representations for both bacteria. This analysis identified protein targets, their functions (through the respective GO terms), and their gene interactions which are most responsible for the decrease or increase in bacterial load. In Figure WP1-8, we show the predictive clustering tree for predicting the increase/decrease of bacterial load for *M. tuberculosis*, in which each tree leaf lists the compounds that belong to it. This tree selects the following GO terms/functions as most important: negative regulation of type 2 immune response (GO:0002829), defence response to virus (GO:0051607), early endosome (GO:0005769), myeloid cell activation involved in immune response (GO:0002275), phagocytic cup (GO:0001891), ER-associated protein catabolic process (GO:0030433), circulatory system process (GO:0003013) and antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent (GO:0002479). All these functions are closely related to the process of *M. tuberculosis* infection and the defence response to it.

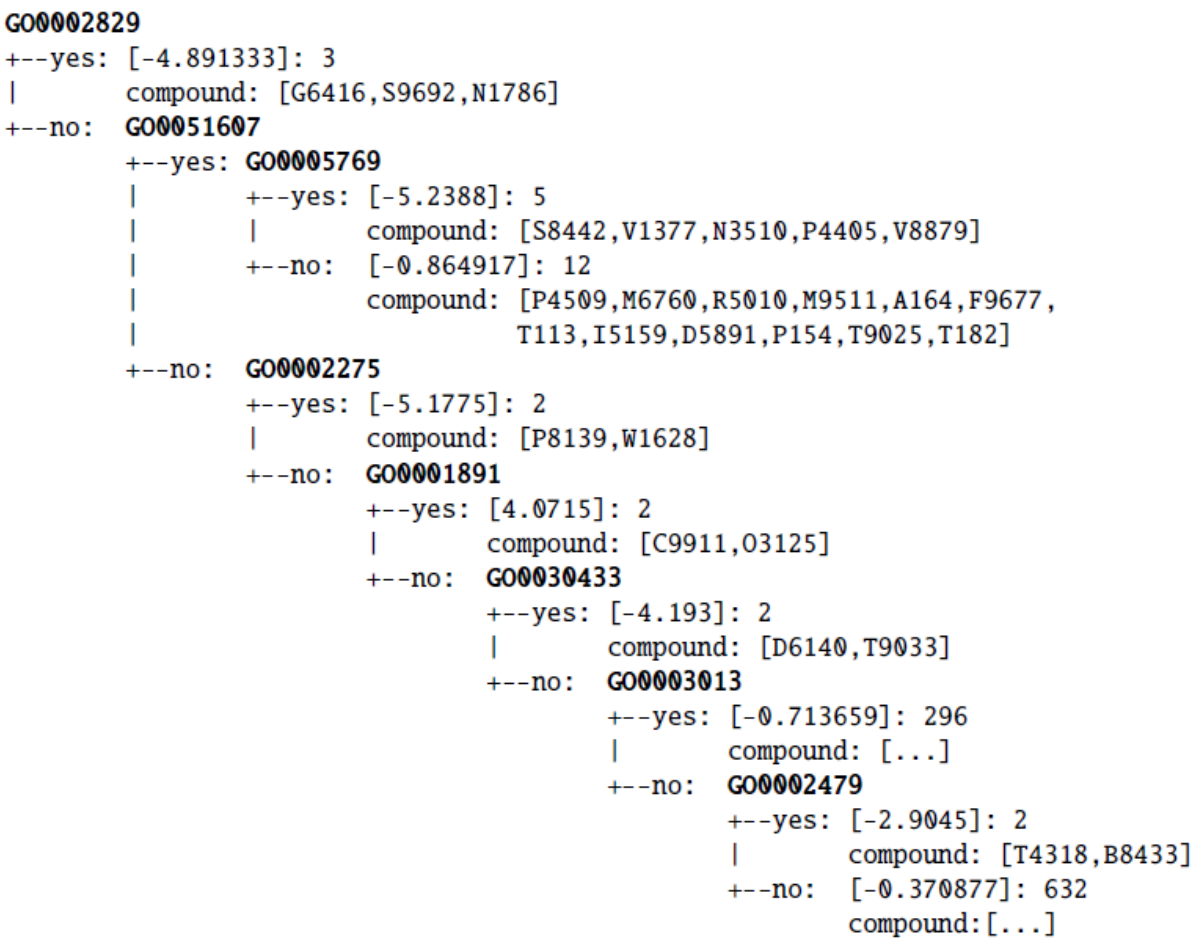


Figure WP1-8: The predictive clustering tree for predicting the increase/decrease of bacterial load for *M. tuberulosis*. Each tree leaf lists the compounds that belong to it. This tree is obtained from data where the compounds are described with the GO terms that the respective protein targets were annotated with.

For the third analysis, we performed feature ranking on the three data representations. We used RReliefF as a feature ranking algorithm (Robnik-Šikonja and Kononenko). Namely, we would like to obtain an ordered list of genes or gene functions that are most relevant for the decrease/increase of bacterial load. After that, using the top 40 ranked genes we constructed a gene network using the STRING database. We show the two networks in Figures WP1-9 and 10.

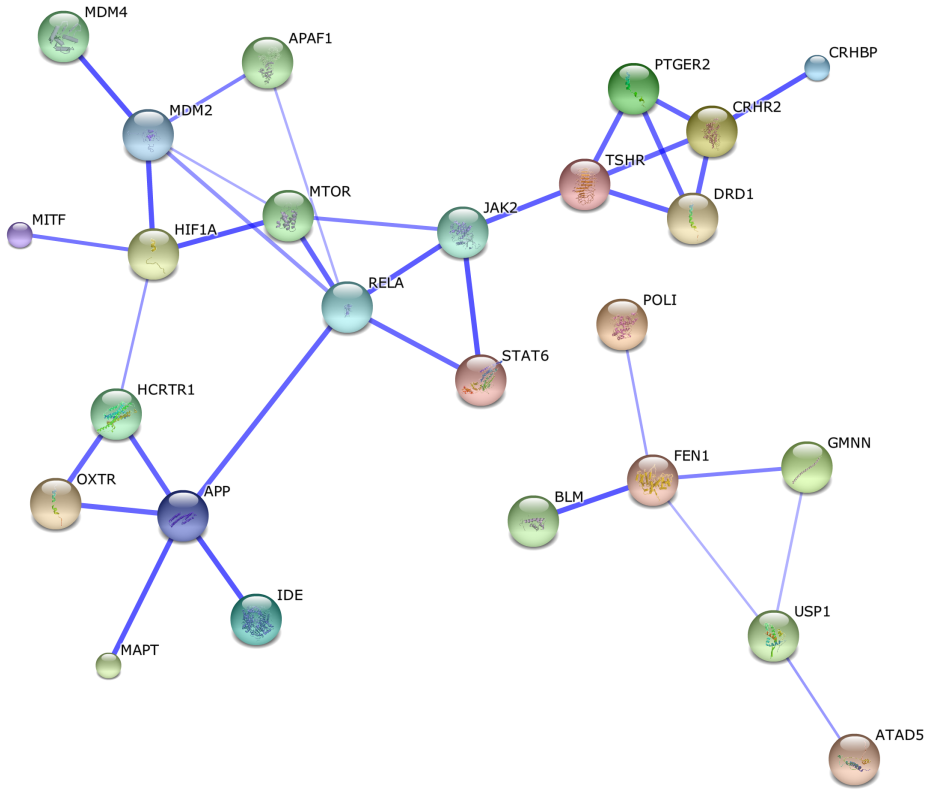


Figure WP1-9. Gene network for the gene list obtained by feature ranking on the *M. tuberculosis* data.

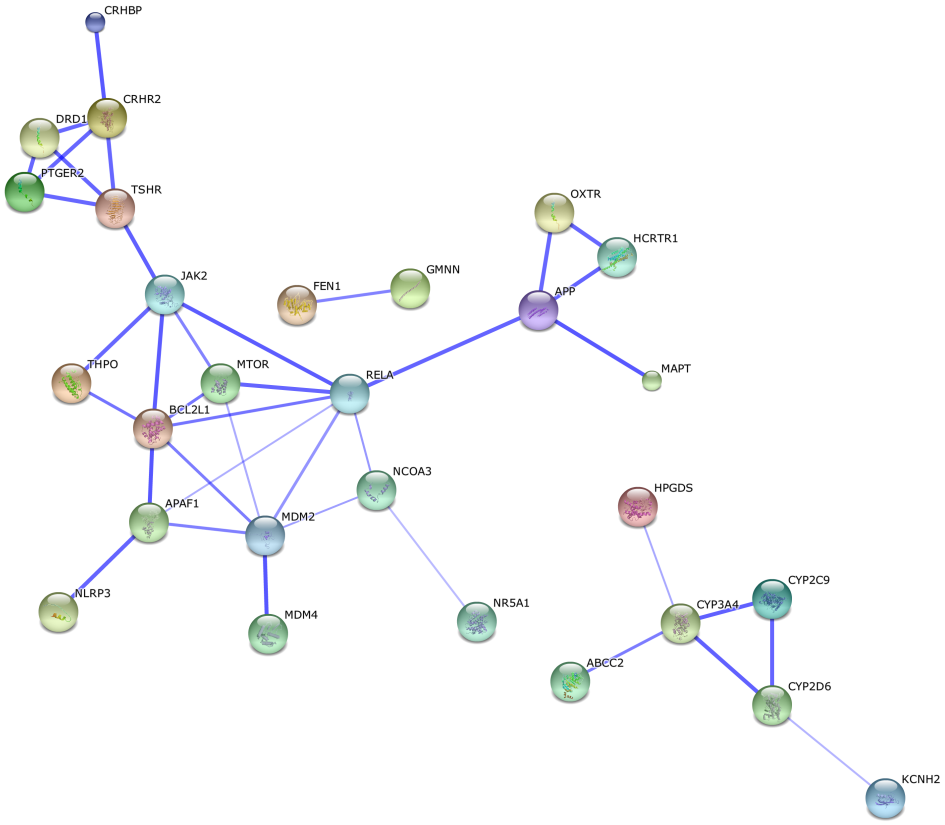
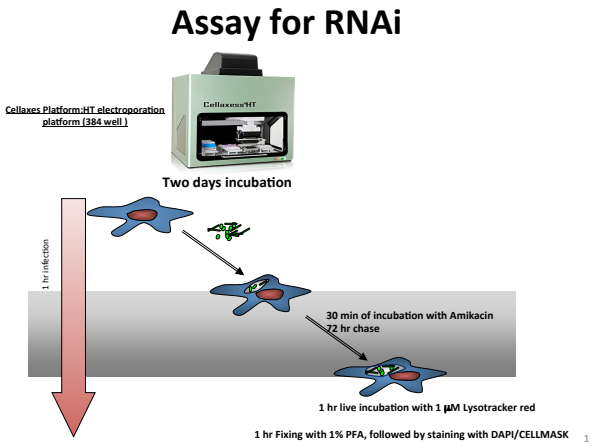


Figure WP1-10. Gene network for the gene list obtained by feature ranking on the *Salmonella typhimurium* data

WP2 - Assay Development and Phagosome characterisation

During the project we successfully developed high-throughput, high-content screening assays (see figure) using both murine and the more physiologically relevant primary human macrophages, derived from peripheral blood monocytes. We optimized the mycobacterial infection assay, combining the model organism *Mycobacterium bovis* bacillus Calmette-Guerin (BCG), genetically engineered to express GFP, with LysoTracker® Red DND-99 staining of lysosomes in the host cells. This provided a direct readout of the release of the phagosomal maturation block and degradative delivery of the bacteria to lysosomes when the system was perturbed.



Significant results

- Establishment of the high-throughput transfection protocol for efficient RNA knock-down in murine cell lines and primary human macrophages in 384 well format, using the high throughput electroporation platform CELLAXES
    - Upgrading of the QMPIA image analysis system for RNAi screens in both murine cell lines and primary human macrophages
- An RNAi library that targets ~300 genes (4 oligos per gene) was screened twice in *M. bovis* BCG-infected primary human macrophages
    - The screen was analysed based on identification of rare phenotypes by binning of multi-parametric profiles
- A library of 2000 chemical compounds was also screened in *M. bovis* BCG-infected primary human macrophages
    - Hits are being validated using laboratory and clinical isolates of *M. tuberculosis* in macrophage infection assays, measuring bacterial growth as the output
- The mode of action of hits identified in both chemical and RNAi screens is being inferred by comparing the multi-parametric phenotypic profiles obtained in both screens

Effective transfection protocols were established for infected human macrophages, but we encountered problems in identifying phenotypes associated with successful phagosome-lysosome fusion and mycobacterial degradation. Initial analysis using average values obtained from whole sample wells proved inadequate, so we implemented a binning analysis to improve the degree of separation between experimental and control samples. Finding the link between profiles obtained from the RNAi screen with those obtained from the chemical compounds screen proved more complex than initially expected: nevertheless by including additional functional genomics data we were able to identify host macrophage genetic nodes, such as autophagy and rates of endocytosis trafficking, that when perturbed by chemical treatment, led to clearance of the intracellular pathogen.

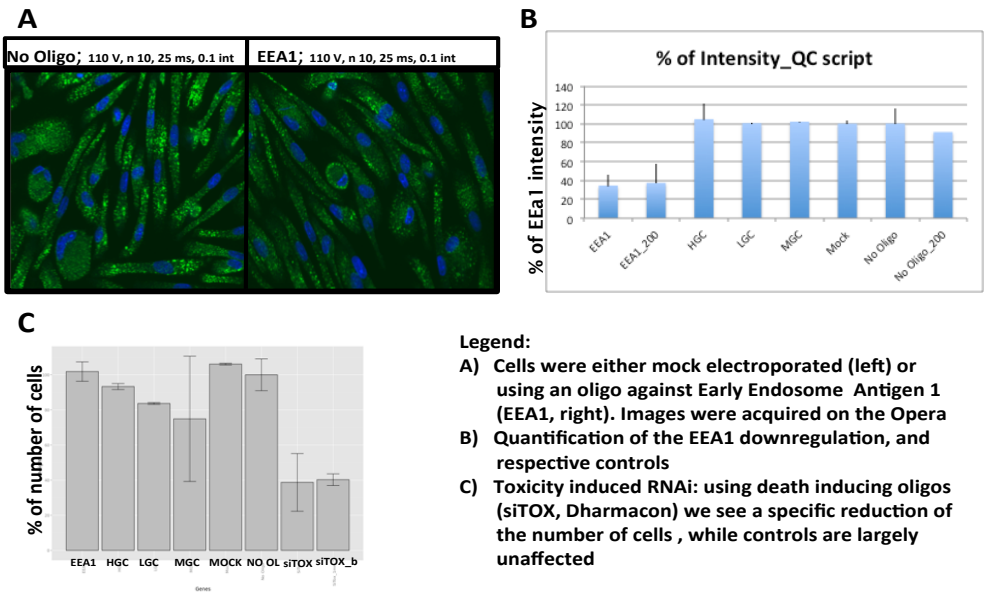
Deliverables 2.1 (High-throughput Imaging assay) and 2.2 (Predictive models of phagosome maturation)

We developed a phenotypic cell-based assay to use in an RNAi or chemical compound screen, aimed at systematically identifying host regulators of mycobacterial phagocytosis and, in the longer term, novel therapeutic targets. The optimised model uses primary human macrophages as host

cell, infected with *Mycobacterium bovis* bacillus Calmette-Guerin (BCG) genetically engineered to express GFP.

The processes established to implement the screen involve:

- transfection by electroporation to transfect primary human macrophages in optical 384-well plates with siRNAs, using the CELLAXES technology (see below for example).



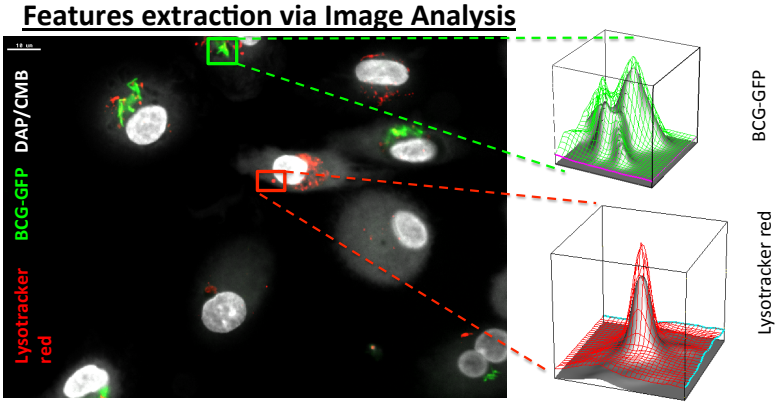
Legend:

A) Cells were either mock electroporated (left) or using an oligo against Early Endosome Antigen 1 (EEA1, right). Images were acquired on the Opera

B) Quantification of the EEA1 downregulation, and respective controls

C) Toxicity induced RNAi: using death inducing oligos (siTOX, Dharmacon) we see a specific reduction of the number of cells , while controls are largely unaffected

- automated protocols to wash away the transfection medium, administer the bacteria to the host cells and wash away the non-internalized bacteria after the pulse period
- imaging the fixed plates at high resolution (40x) on an automated confocal microscopy platform (Opera, Perkin Elmer) equipped with an automated plate stacker
- the adaptation of image analysis software (Motion Tracking, developed in house by Prof. Yannis Kalaidzidis) to extract quantitative descriptors of the images and generate multi-parametric fingerprints of the possible phenotypes (see example below).



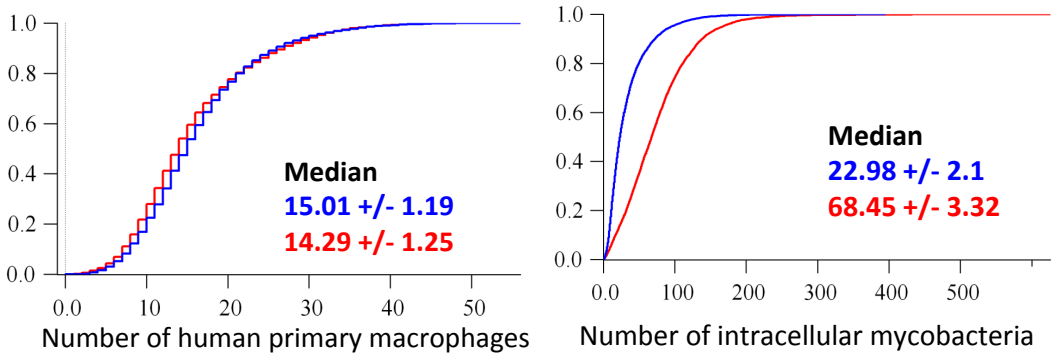
Extraction of quantitative object features such as :

Number, shape, size , intensity of cells (DAPI-CMB),
Bacteria (BCG-GFP) and lysosomes (lysotracker red)

Intra object relationships :

Bacteria within cells or not, bacteria within lysosomes or not

The plots below show examples of the cumulative distributions (expressed from 0 to 1 on the Y-axis) of the number of macrophages (left) and the number of bacteria (right) at two time points: 2 hour after infection (blue line) and 48 hour after infection (red line). While the number of macrophages stays constant, the number of intracellular bacteria is increasing, showing that this assay can be used to describe the fitness of the bacteria in the context of the cells (as well as their relation with intracellular organelles). The assay has been validated by measuring colony-forming units using the same conditions of the image based assay.

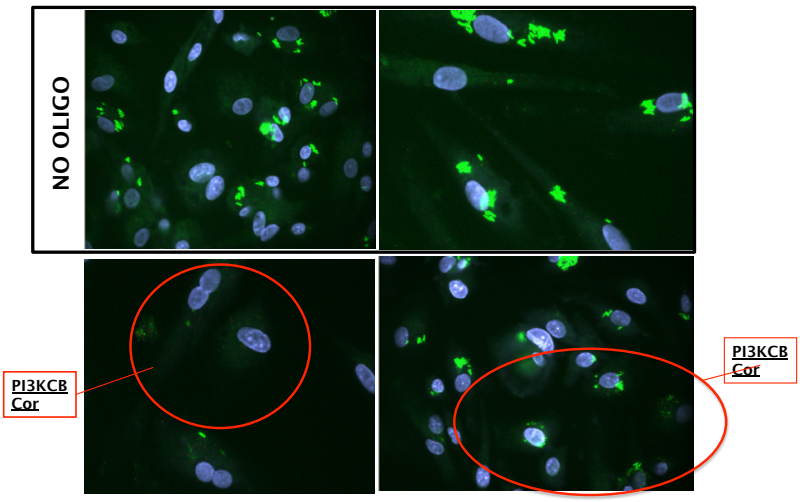


Deliverables 2.3 and 4.6 (data array/ phagosome maturation and endocytosis)

The human macrophage model was used to perform a screen of a targeted library of 300 genes. Due to the large amounts of oligonucleotides needed for the electroporation technology, we decided to use a smaller library than the human kinome and phosphatome libraries originally proposed for these deliverables.

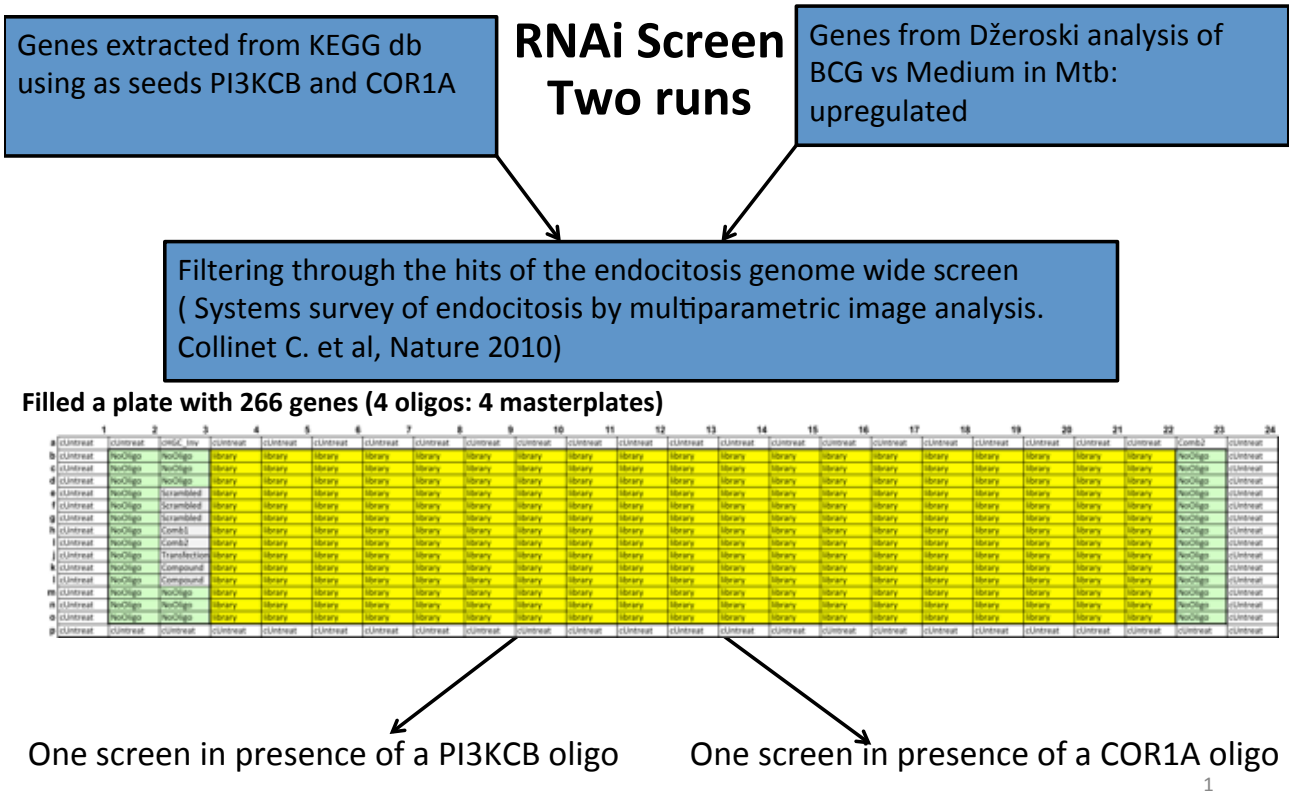
The screening campaign was preceded by an assay development phase in which we wanted to find positive controls: genes that, when down-regulated, would produce a phenotype resulting in degradation of intracellular bacteria (described by different parameters such as reduced intensity and number the bacteria, lysotracker signal associated with bacteria, all in the context of different host cells metrics).

We selected genes identified in the literature as having an effect on the host response to mycobacteria, but were unable to find a single gene that gave a measurable phenotype in our assay. A double knockdown approach however, provided appreciable phenotypes. In particular the combinations of oligos targeting Coronin1 (Cor) and Pi3KCB gave a reproducible bacterial degradation phenotype, illustrated below.



Examples of phenotypes: the upper row shows two examples of negative control images, where the bacteria shows a normal structure and GFP fluorescence; the lower panels show examples of fields with the double knockdown, where some cells with signs of degraded bacteria were clearly visible (highlighted by the red circles).

The decision was taken then to perform the screen using a double knockdown geometry, in which a set of 266 genes, chosen following the rationale schematized in the figure below, were silenced in the presence of an oligo against Cor1A OR against Pi3KCB.



Microarray expression data from BCG-infected human macrophages (analysed by WP1: Saso Džeroski) was combined with the hits from an endocytosis genome wide screen performed in the Zerial lab (Collinet et al., 2010): the rationale being to enrich for genes whose knock-down has a functional impact on the regulation of endocytosis (deliverable 4.6), a process which shares components with mycobacterial phagocytosis. We had originally planned to directly apply our assay to a panel of genes comprising the kinome and a set of regulators of endocytosis, however this proved impracticable due to difficulties associated with scaling up the technology.

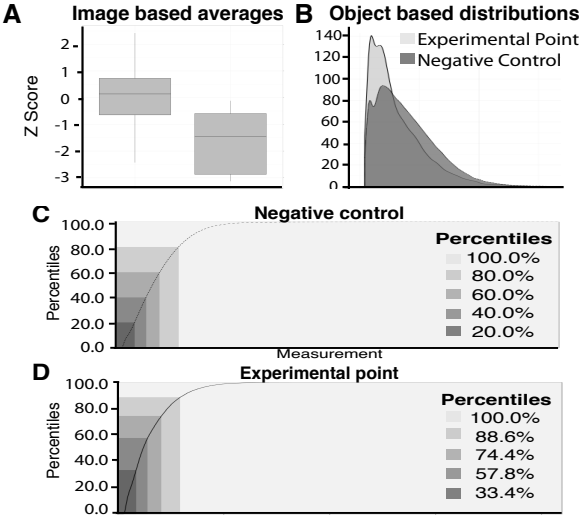
We have however increased the value of the screen by using a more relevant model system; primary human macrophages, instead of the originally proposed murine cell lines. In addition, we have also complemented the limited *Functional Genomic* activity, with a *Chemical Genomic* approach. The use of Chemical array data, and its integration with Functional Genomic data in the context of a different assay, has allowed us to produce interesting results regarding the mode of action of the compounds and the identification of host cellular pathways that could be exploited to re-establish the delivery of the pathogen to the phago-lysosome.

The analysis of a multi-parametric screen is a complex task, that was rendered even more complicated in this case by the fact that averaging the parameters measured in each well was insufficient to find phenotypes, due to the fact that only a subpopulation of cells in any given well shows a particular phenotypic response. In these circumstances we have found that dividing the negative control population into percentiles (binning) and comparing the population distribution in the corresponding parameter was a useful analysis method for scoring hits.

We have developed an R template implemented in KNIME for this subpopulation analysis. The script partitions the sorted population of a negative control into n equal sized bins (in this case n=5). Each bin covers a percentile of a parameter range. The script then determines the count of

objects in each parameter bin for each experimental well. Objects whose values lie either below or above the range of the negative control are attributed to either the lowest or highest bin respectively. The distribution of population counts is then used to compute z-score values relative to the negative control. This process is applied automatically for each parameter and each experimental condition in KNIME partitioning the population into a chosen number of bins. The final output of the template is a profile of n z-scores that we call 'z-score profiles'.

This analysis procedure is shown schematically below:



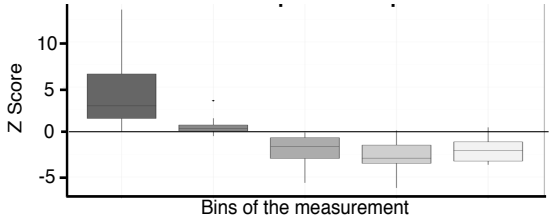
A) Image based averages of negative controls (left) and positive experimental points (right). The discrimination power of this experimental point is low.

B) Object based distributions.

C) Cumulative distribution of the object based measurement of a parameter: at this level the distribution is divided in 5.

D) The x axis of the distribution of the values of the same parameter is divided by the values found in **C**, and the percentiles of the counts of objects in each bin is the used to compute a Z value. The results is the "z-score profile" (see below) that has a higher discriminative power than the average based measurement.

Eliminating z-scores between +2 and -2 using a KNIME node named 'Range Filter' reveals the remaining parameters whose mean is at least two standard deviations away from the control population and have some discriminatory power.



In this manner is it possible to select parameters that score changes in population distributions. To reduce the noise we also operate a parameter selection, based on the evaluation of the correlation of the object based measurements, taking into account only parameters that have <0.4 of Pearson correlation between each other. The result of this series of steps is the identification of bins of a discrete series of parameters (15), which are then used to calculate distance metrics, like the Mahalinobis distance, to define hits based on strength of the phenotype and to cluster genes in phenotype specific classes.

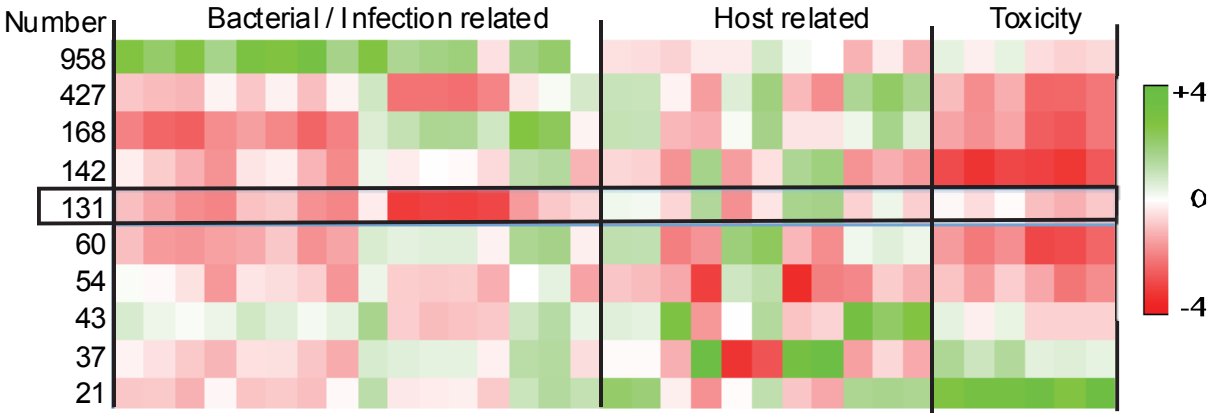
The result of the two different runs of the screen identified specific group of genes for each background, and we hypothesize that this is due to different epistatic relationships between the knockdown of the gene in the COR background and the PIK3CB background. A difficulty in providing a stable and definitive hit-list is the variability of the results from run to run. We will proceed with more runs to increase the statistical power of the analysis.

A complementary approach to find host regulators of bacterial survival was to use a Chemical Genomics screen in the same assay, to search for compounds able to stimulate the host cells to overcome the mycobacteria-induced phagosome-lysosome fusion block. This was achieved by screening a library of 2000 well-defined compounds, most of which are FDA approved drugs, in mycobacteria-infected primary human macrophages. In order to find compounds that were able to relieve the block without having a direct influence on bacterial uptake, we added the compounds to already infected cells.

The analysis in this case was more straightforward as the phenotypes identified were much more homogeneous in the cell population, compared to the functional genomics approach, meaning that averages derived from all the cells in a well had sufficient discriminative power. Phenotypic

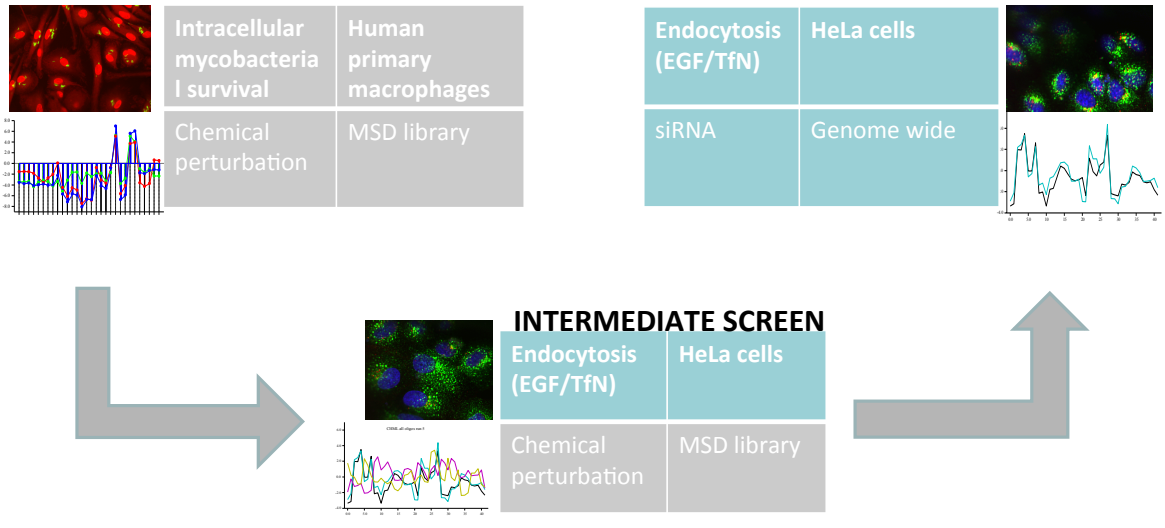
strength and clustering approach were used to discriminate compounds with no effect from various phenotypic classes.

In particular we focused our attention to a cluster of 131 compounds having a profound effect on the bacterial fitness and, at the same time, having little or no influence on the host cell parameters. These are highlighted in the following scheme where the clusters are colour coded, with each box representing a Z value of a single parameter. We were able to show that the majority of the compounds in this cluster were not directly bacteriostatic, but only exerted their function through the host cells.



Three of these compounds were validated using one lab strain (H37Rv) and two clinical isolates of *M. tuberculosis* to infect primary human macrophages that were then treated with the compounds. Lysing infected macrophages and counting colony-forming units assessed bacterial viability.

The analysis of the chemical genomic campaign was certainly more straightforward than for the functional genomics screen, with chemicals give better phenotypes in the assay, but as expected the identity of the mode of action of the compounds remains problematic. In the absence of a direct functional genomic/chemical genomic link, we decided to approach the problem by integrating data from different screening campaigns to try to describe important host features in the reaction to the mycobacterial infection. In particular we have access to a highly granular set of data, the endocytosis genomic screen, in which every gene is quantitatively annotated in an assay describing endocytosis (Collinet et al., 2010): this assay was performed on HeLa cells, so direct phenotypic comparison was difficult. Consequently we decided to perform a bridging screen, in which we used the HeLa system and the endocytosis assay, but in presence of the same library as the mycobacteria assay. This resulted in a description of endocytic profiles of the compounds which were active in the mycobacteria assay. The approach is schematized in the following cartoon



By comparing the multi-parametric profiles of these phenotypes with those of the genome wide assay of endocytosis, we were able to perform an enrichment analysis of the cellular processes that genes with similar profile in the endocytosis assay shared. Using this technique we were able

to predict the modes of action of the compounds in the active cluster of the mycobacteria-macrophage infection assay; this was then successfully validated experimentally for some of the hits.

The results of this activity are currently being summarized in a manuscript to be submitted for publication to "*Cell: Host and Microbe*".

WP3 – RNAi screens: uptake and model testing

Deliverable 3.1 (Focused data array / uptake)

The work in WP3 was delayed by the untimely and tragic death of Dr. E. Caron in July 2009. A cell biologist, Dr. Vania Braga, was recruited in September 2009 to manage WP3. Taken together with the high turnover of appointed post-doctoral staff (2 replacements in 18 months), this has had adverse effects on the project and led to delays in the progress of the experimental work.

Our initial screening work focused on setting up an appropriate experimental model to study *M. bovis* BCG uptake by the human macrophage-like cell line THP-1, suitable to use in an RNAi screen with a library targeting human mRNAs. We aim to (a) optimise a quantitative detection method that is suitable to test hypothesis-driven experiments, as well as medium-throughput screens; (b) validate the chosen model system and experimental conditions to identify and quantify signalling pathways relevant for bacterial uptake.

Identifying the receptor relevant for BCG internalization will facilitate mechanistic understanding of the signalling processes involved and how to interfere with them therapeutically. During the work to optimise the model system, we demonstrated that macrophages differentially uptake BCG found as single particles or as clumps. This is an interesting observation as mycobacteria characteristically grow as clumps, which show some similarities to biofilms formed by other bacteria, which provide environmental and survival advantages (Lasa, 2006). However, the functional consequences of clump uptake for phagosome maturation, intracellular mycobacteria survival or pathogenesis have not been investigated. In fact most research has focused on mycobacteria as a single particle because of the technical challenges of dealing with clumps.

An immunofluorescence-based assay was optimized to quantify levels of associated and internalised phagocytic targets, and differentiate between single or larger particles (Maryke Carsten, Imperial College). We demonstrated that in THP-1 macrophages, CR3 receptors are required for internalisation of single BCG, but not clumps. RhoA, involved in signalling pathways downstream of CR3, is also necessary for BCG uptake. Indeed, RNAi depletion of RhoA reduced the uptake of single BCG bacilli. Our data suggest that CR3 can act as a direct phagocytic receptor for BCG, and that BCG internalization utilises signalling from Rho small GTPases as shown for opsonized red blood cells.

The implications of our results are that, depending on the size of BCG particles presented to macrophages, different receptors are engaged and responsible for uptake. The corollary is that distinct intracellular signalling can be triggered and may influence the outcome of infection and immune response. To address the functional significance of our findings, we used two approaches. First, in collaboration with partners Dorhoi and Kaufmann (WP5), we determined that mycobacteria is found both as single cells and as clumps in the lungs of infected mice. Second we demonstrated that infection with single or clumped BCG leads to different patterns of cytokine production by human macrophages. The IL-1b and IL-6 expression is significantly increased by clumped BCG at 48 hrs compared to single BCG. However, IL-4 and TNF-b expression was not detected at all when THP-1 monocytes were infected with either BCG size. These data suggest that single BCG and clumped BCG induce cytokine expression differentially, consistent with their internalisation through different signalling receptors.

Deliverable 3.3 (Finalized data array / uptake)

We next proceeded to identify which signalling molecules could participate in the differential internalization of single or clumped BCG using RNAi to deplete Rho small GTPases GEFs, GAPs and effectors, including a number of kinases and phosphatases. The differential labelling detection method (associated or internalized particles) was successfully adapted for use with BCG uptake in THP-1 macrophages seeded in 96-well plates. Custom-made software was developed by Dr. Chris Tomlinson (Centre for Integrative Systems Biology and Bioinformatics, Imperial College) to allow automated quantification from immunofluorescence images to generate four parameters: attached or internalized particle detection, single or clumped BCG.

The software is able to (i) identify specific BCG shapes by setting a threshold to eliminate debris and differentiate single (20-80 pixels) and clumps (81-500 pixels) at the amplification of acquired images; (ii) determine the localization of mycobacterium on or inside the THP-1 cells. Single and clump BCG were coloured in different colour inside the cells for quality control and visual analysis. The final output gives the details of each well: the number of cells, the total BCG count, the BCG attached to the cells, BCG internalised, the number of single BCG and clump BCG attached and internalised. Furthermore, software was thoroughly validated with a subset of images that had been visually analysed or semi-automatically by using ImageJ software (Figures 1 and 2).

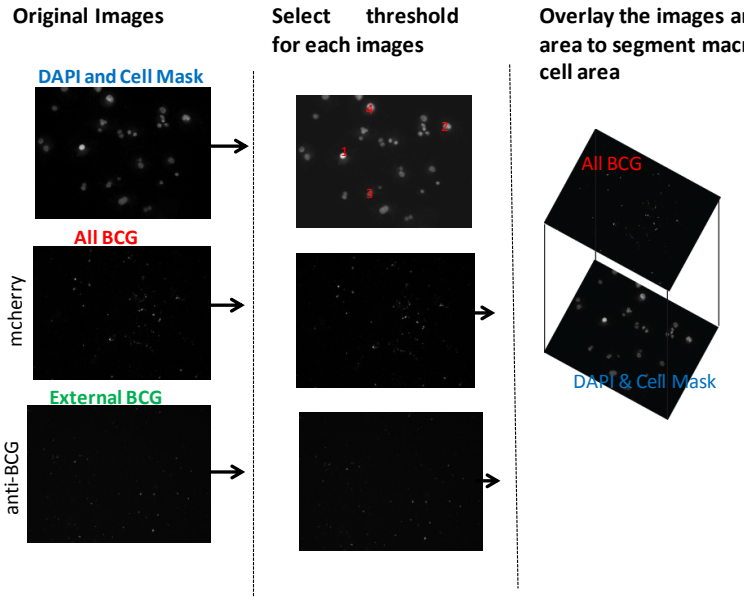


Figure 1: Workflow for software development using differentially stained images of THP-1 macrophages (DAPI and Cell Mask), mCherry BCG (All BCG) and attached BCG (secondary antibody, Dylight 649). The threshold was selected for each stained image, the number of macrophages was counted and the number of BCG attached or internalized by the cells.

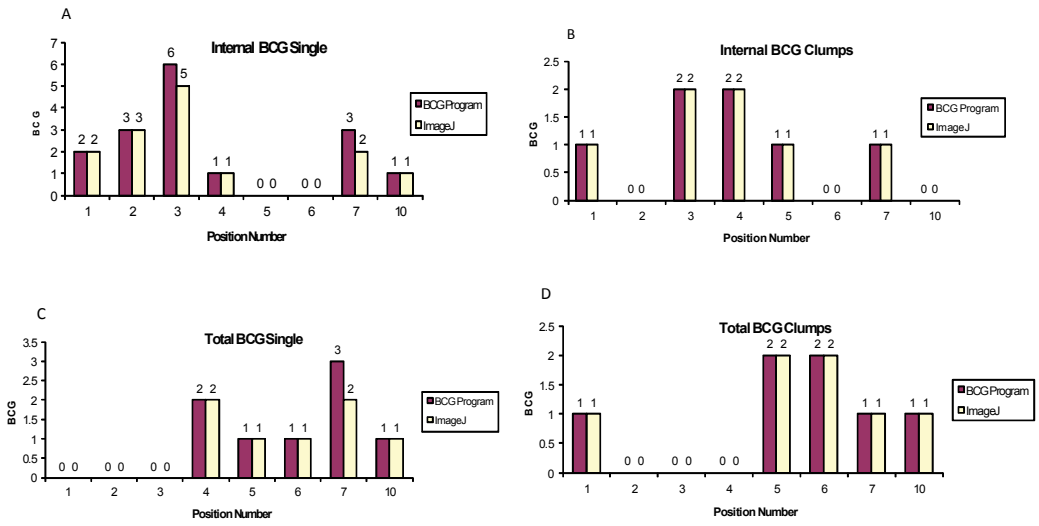
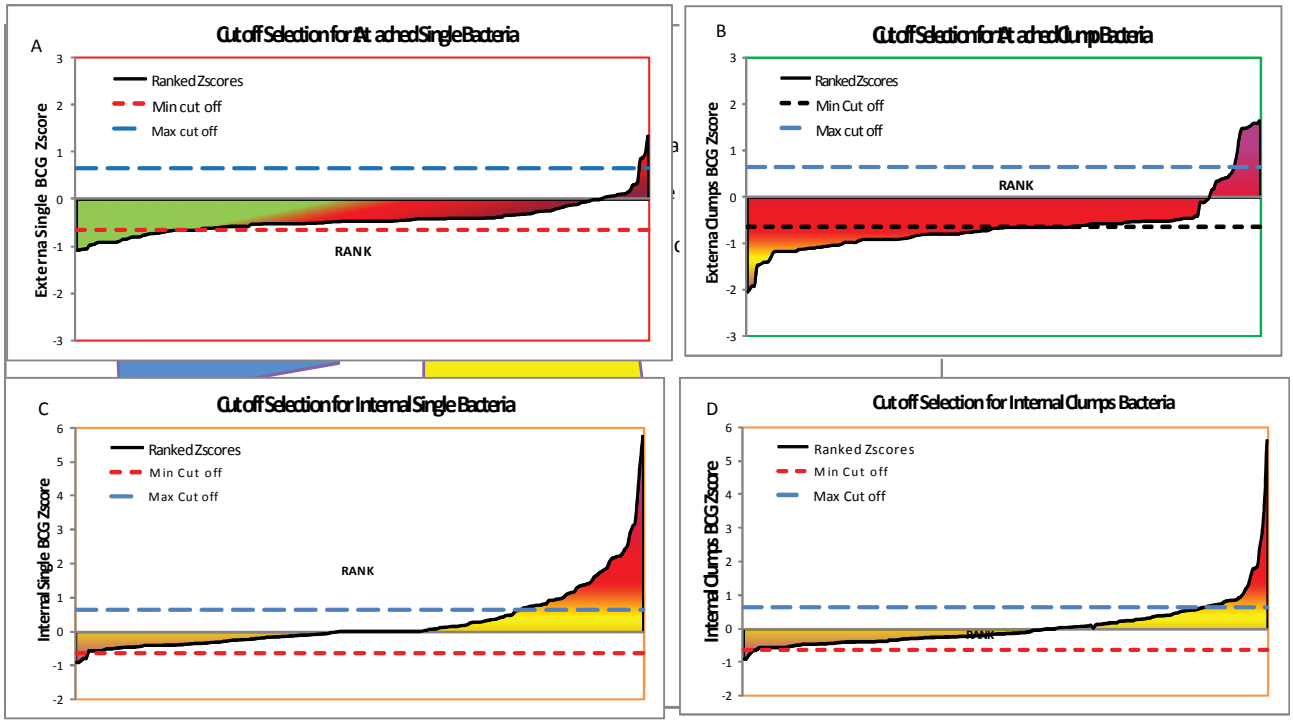


Figure 2: Validation comparing manual (ImageJ) and automated (BCG program) counting. The graphs represent the number of Single and Clumped BCG counted by both methods. (A) single internal BCG, (B) internal BCG clumps, (C) total single BCG, and (D) total BCG clumps. Actual numbers counted are shown above the bars.

RNAi screens were performed in triplicate. In collaboration with partners Barsacchi and Zerial (WP 2), the fluorescent images from the screen plates were obtained with an automated confocal microscope (Opera, Perkin Elmer) equipped with an automated plate-stacker (10 images per well were collected with four channel DAPI, Cy3, Cy5 and FITC). A large-scale dataset of images was generated and analysed by the custom-made software, including statistics to identify the target proteins required for BCG uptake (Figures 3 & 4). We selected 30 proteins for further validation with four individual oligos to confirm the phenotype.

Figure 3: Z-score for four parameters: (A) Attached single BCG, (B) Attached Clumped Bacteria, (C) Internal Single BCG, and (D) Internal Clumped BCG. A cut-off value of 0.65 was selected.



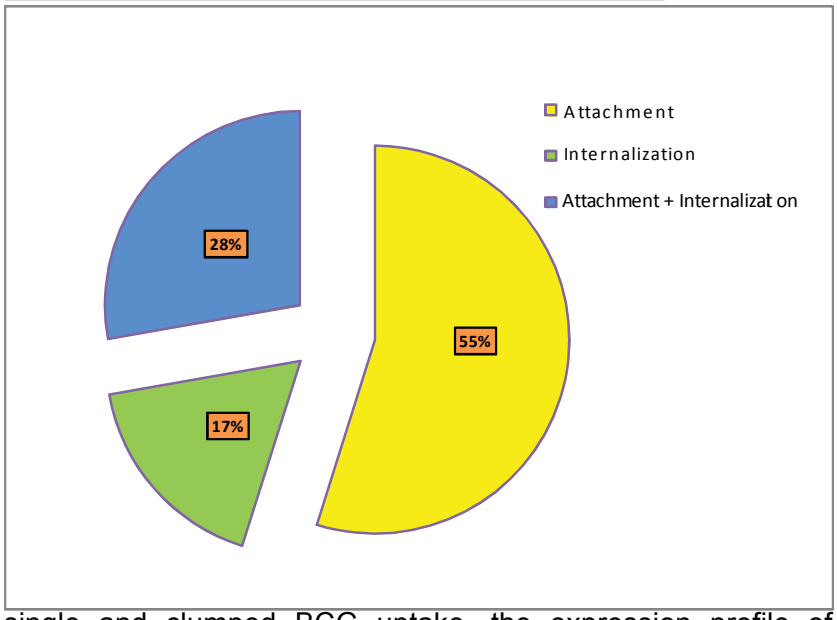


Figure 4: Percentage of targets grouped by phenotype, attachment (yellow), internalization (green) and both attachment and internalization (blue).

In summary, we have set up a robust assay to investigate the mechanisms of uptake of BCG by human macrophages. CR3 appears to be important for uptake of single bacteria but it is not required for attachment of bacteria to macrophages. Consistent with different receptors being required for single and clumped BCG uptake, the expression profile of cytokines was distinct if THP-1 macrophages were infected with single compared to clumped BCG. Finally, we have also completed the RNAi screen and developed software to identify the network of regulators mediating BCG uptake. In work continuing beyond the end of this project, we will (a) cross-reference our results with data provided by WP2, WP4 and WP5 and (b) test predictions and molecules identified in other work-packages (bacterial or host proteins) in the uptake of BCG. A manuscript on the characterisation of CR3 as a direct phagocytic receptor for BCG is being finalised for re-submission.

WP4 - RNAi screens – maturation & model testing

The two partners contributing to this work-package have established protocols and methods for screening *M. tuberculosis* (LUMC) and *Salmonella* (NKI/LUMC) infected phagocytic cells. The project required high throughput data acquisition so it was essential to establish a solid and robust system from which we can obtain reliable and reproducible data. LUMC developed and perfected the basic system, and optimised separate procedures for the *Salmonella* and *M. tuberculosis* screens.

Developing the assay took longer than originally planned, primarily due to fact that we had to establish multiple parameters, namely selection of suitable cell lines that were both easily transfectable with siRNA, as well as capable of phagocytosing mycobacteria. Significant results include:

- **Identification of cell lines** suitable for high throughput siRNA transfection and bacterial infection. HeLa are easily transfectable but non-phagocytic: they are suitable for optimising siRNA transfection conditions and *Salmonella* infection. MelJuSo phagocytose *M. tuberculosis,* and are easily transfectable. The results obtained were highly reproducible, making them a suitable system for the project screens. THP-1 cells, although monocytic and able to phagocytose bacteria, are harder to transfect and thus only suitable for the verification and validation on smaller sample sizes of the primary hits obtained in the screens.
- **Optimise siRNA transfections** in HeLa and MelJuSo cells to achieve up to 95% knockdown of AKT1, relative to scrambled siRNA control oligos; this was verified by immunoblotting. We also confirmed that siRNA knockdown of AKT1 in *Salmonella* Typhimurium infected HeLa cells decreases the bacterial load, and so mimics the effects of H-89, a chemical inhibitor of AKT1. The optimal time point for bacterial infection after siRNA knockdown in HeLa and MelJuSo cells was determined by following the decrease in AKT1 signal using immunoblotting. During the screening stages of the project RNAi protocols were further optimised to enable efficient gene knock-down in the human monocytic cell line THP-1 and in primary human macrophages generated from monocytes isolated from buffy coats of healthy donors. Successful knockdown was verified by qPCR and immunoblotting.
- **Infection protocols.** We optimised and streamlined protocols for infection of HeLa cells with *Salmonella* and MelJuSo cells with *M. tuberculosis* to increase our readout window. We verified that H-89 (chemical inhibitor of AKT1) treatment of HeLa and MelJuSo cells infected with *Salmonella* or *M. tuberculosis* reproduced our previous results in primary macrophages. Time-course experiments on MelJuSo cells infected with *M. tuberculosis* were performed to determine the optimal time-point post infection for readout.
- **Set up a 96 well screening assay.** Our HTS flow cytometry-based assay is a fast, robust and cost-effective semi-automated screen employing a flow cytometer coupled to a 96-well plate loader for measuring bacterial load using *Salmonella* or *M. tuberculosis* expressing DsRed, and estimating host cell survival/proliferation. Flow cytometry results are in agreement with colony forming unit (CFU) assays and fluorescence microscopy.
- **Developed novel tools for studying intracellular trafficking of *Salmonella* and mycobacteria.** These include bacteria (conditionally or constitutively) expressing stable or destabilized DsRed fluorescent protein and bacteria expressing multiple conditionally expressed reporter genes.
- **Data analysis.** We established semi-automated data analysis for the screening stages of the project, providing a standardised computer readable data format.

A paper outlining the RNAi and chemical compound screening assays is in preparation with intention to publish before 2013.

Deliverables

Utilizing the methods outlined above we performed siRNA and chemical compound screens on human cell lines infected with either *Salmonella* or *M. tuberculosis*. Promising preliminary screening results showing considerable bacterium-specific effects of host-targeted chemical compounds prompted us to perform all screens using both *Salmonella* and *M. tuberculosis* infected human cell lines. The overall focus was towards the study of regulatory proteins, relatively unbiased chemical compound libraries and integration of different screens to establish a comprehensive view of the regulatory host protein network for both *Salmonella* and *M. tuberculosis*. The focused RNAi array for Rab modifying proteins (**Deliverable 4.1**: Focused data array/pathogen survival) was abandoned in favour of an RNAi library for human deubiquitinases, which has a higher chance of identifying *novel* leads due to its unbiased nature (posttranslational protein modifications by deubiquitinases are not restricted to specific cellular mechanisms or protein families) and is as of yet left unexplored in the context of *M. tuberculosis* infection, increasing chances of generating high-impact results. RNAi screens for the human kinome were performed as part of **Deliverable 4.3** (Focused data array/phagosome maturation). Both RNAi screens identified hits that markedly decreased or increased bacterial load of *Salmonella* or *M. tuberculosis,* or both in our cell-based assay. Results of these screens were fed into **WP1** for gene ontology (GO) analysis and constructing predictive clustering trees (PCTs).

Due to some delays in final assay development for the siRNA screens we decided to include chemical compound libraries as part of the drugome array (**Deliverable 4.4**: Drugome array/*Salmonella*). To identify druggable targets involved in regulation of phagosome maturation in human cell lines infected with *Salmonella* or *M. tuberculosis*, we performed four separate chemical compound screens:

- A **Library of Pharmacologically Active Compounds** (LOPAC; commercial, Sigma-Aldrich) was used to probe host regulatory networks for druggable targets regulating intracellular survival of *Salmonella* or *M. tuberculosis*. Results of this screen were fed into **WP1** to develop an automated method for aggregation of known protein targets of chemical compounds in public repositories (PubChem, ChEMBL), outlined in the section on WP1 in this report. This provided a valuable protein target framework for identifying novel regulatory networks, due to its unbiased nature (chemical compounds may target any type of host protein). Targets were analyzed for enrichment, GO analysis was performed and PCTs were constructed to predict targets or combinations of targets for follow-up (**WP1**).
- A screen of the **third generation PKB/Akt1 inhibitors derived from H-89** (H. Overkleeft, Leiden Institute of Chemistry) on human cells infected with *Salmonella* or *M. tuberculosis* led to the identification of several molecules with strong inhibitory effects on intracellular *M. tuberculosis*, *Salmonella* or both. Kinase profiling of a selection of compounds with distinct effects on bacterial load and PKB/Akt1 inhibition, identified candidate kinases that may regulate intracellular survival of *Salmonella* or *M. tuberculosis*. The role of these kinases in regulation of phagosome maturation is currently under investigation using RNAi. Of note, several kinases identified by this approach were also identified in the kinome RNAi screen. A paper on the PKB/Akt1 inhibitor screens is in preparation with intention to publish before 2013.
- A **library of potential phosphatase inhibitors** (H. Ovaa, Netherlands Cancer Institute) was screened on human cells infected with *Salmonella* or *M. tuberculosis* and resulted in the identification of a novel compound that reduced *M. tuberculosis* bacterial load.
- A **library of autophagy inhibitors/activators** (commercial, ENZO Life Science) was screened on human cells infected with *Salmonella* or *M. tuberculosis*, identifying multiple chemical compounds affecting bacterial load of either *Salmonella* or *M. tuberculosis* or both. Of note, the top hit in this screen was structurally similar to two compounds from the LOPAC library screen, all three resulting in a decrease in *M. tuberculosis* bacterial load. These compounds all shared a deubiquitinase as a common target, which was also identified in the siDub screen.

Generally, all screens identified compounds or siRNA oligos that affected bacterial load of either *Salmonella* or *M. tuberculosis* or both. Hits from all the screens were integrated in a single network analysis to identify their possible interactions and/or participation in common processes. Here, targets identified in the LOPAC library screen were used to construct a general network consisting of targets with a regulatory role as well as targets with effector function, whereas the RNAi screens

provided valuable data to identify crucial regulatory nodes within these networks. In addition to the unbiased, data-driven predictions performed in **WP1**, targets that were either identified in multiple independent screens or that complemented each other in the different regulatory networks identified using STRING or Ingenuity Pathway Analysis were selected for follow-up. Although the final analyses are still on going, the combined efforts have already led to the identification of several lead targets and testable hypotheses that are currently being tested experimentally. The combined efforts will be pulled together as **Deliverable 4.5**, and result in several papers with intention to publish in the course of 2013.

Interfere with the human kinome and validation of positive hits: In our previous screen at the NKI (Kuijl et al., 2007), we have shown that kinases involved in pathways related to cytoskeletal re-arrangements affect the maturation of *Salmonella*-containing phagosomes. During this work-package we focused our studies on the ABL1 tyrosine kinases that have been shown to regulate Rac-dependent cytoskeletal dynamics in mammalian(Van Etten et al., 1994; Wang et al., 2001) suggesting its role in bacterial uptake(Ly and Casanova, 2009). So far we have demonstrated, in agreement with previous studies(Ly and Casanova, 2009), how impairment of ABL1 tyrosine kinase activity before *Salmonella* infection (i.e two hours before infection) by using a specific ABL-inhibitor Gleevec (also called Imatinib;(Khorashad et al., 2009)) affects *Salmonella* uptake. The same result was obtained by using Gleevec-like inhibitors NTK-9 and NTK-12 synthesized by H. Owa's group (NKI). In an effort to investigate the pathways involved in phagosome maturation, we searched for proteins that can affect *Salmonella* internalization and/or replication together with the kinase ABL1.

By performing an siRNA screen for phosphatases, we identified 14 targets involved in the control of *Salmonella* uptake and intracellular growth (C. Kuijl, personal communication). The silencing of some tyrosine phosphatases that we previously identified as proteins impairing *Salmonella* infection, and the treatment of the same cells with Gleevec before their infection with the pathogen, suggested that MTMR7 phosphatase, among others, might have a role in *Salmonella* infection in MCF7 cells. More specifically, the silencing of tyrosine phosphatase MTMR7(Mochizuki and Majerus, 2003), together with inactivation of tyrosine kinase ABL1 by Gleevec treatment, resulted in a strong synergistic decrease in *Salmonella* internalization/replication, suggesting that these two proteins might act together in regulating *Salmonella* infection. However, we were not able to reproduce these data in other cell lines, such HeLa cells, nor in other batches of MCF7 cells, suggesting that MTMR7 phosphatases possibly represent a false-positive hit of the phosphatases siRNA screen mentioned above.

Besides MTMR7, the silencing of some DUSP phosphatases revealed a synergistic decrease in *Salmonella* infection in MCF7 and HeLa cells, in combination with ABL1 kinase inhibition by Gleevec. In particular, the silencing of DUSP 27 and DUSP 11 in Gleevec-treated cells produced an additional effect in decreasing the *Salmonella* infection compared to Gleevec alone. At the same time, we generated a phosphatase chemical library (in collaboration with H. Ovaa, NKI) and purified a number of the phosphatases identified in order to test their activity in the presence of the biologically active phosphatase inhibitors. We have now identified one phosphatase inhibitor that inhibits DUSP 27, DUSP 11 and DUSP 3, and produces an additional effect in decreasing *Salmonella* infection when combined with Gleevec-inhibition of ABL1. This suggests that DUSP phosphatases work together with ABL1 kinase in controlling *Salmonella* infection. We are currently improving the phosphatase inhibitors to generate selectivity for one single phosphatase. These compounds will be used in combination with silencing of kinases or over-expressing their constitutive active form to generate biologically active phosphatase-kinase pairs; this is an important step to generate the data required for understanding kinase/phosphatase networks. A paper outlining the role of kinases and phosphatases in controlling *Salmonella* infection is in preparation with intention to publish before 2013, as part of **Deliverable 4.5**.

Finally, we have shown how to modify our bacteriostatic compound, the PKA inhibitor H-89 into a PKB/Akt1 inhibitor with low activity on PKA, by using a secondary modified compound library. In addition, we have generated potential PKD1 inhibitors based on the H-89 core structure, which will also inhibit PKB/Akt1 and have low inhibitory effects on PKA and PKB/Akt1 using in vitro assays. We are building the third generation library of these structures to further improve selectivity and activity. A paper showing the conversion of H-89 into PKB/Akt1 and PKD1 inhibitors is in preparation.

WP5 - Manipulation of phagosome maturation by pathogens & model testing

Deliverable 5.1 (Data array/pathogen genes)

The original plan was to use siRNA to knock-down target genes in mycobacteria in order to establish their role in intracellular survival. However due to technical problems it proved impossible to use this approach. We therefore switched to the more conventional approach of using mycobacterial mutants attenuated in their intracellular survival as tools with which to probe the intracellular response of the host macrophage.

Many of the attenuated mutants reported in the literature were selected using murine host cells, so we first set out to confirm the reported phenotypes of a panel of mycobacterial-mutants using human cells. The intracellular fitness of mycobacterial mutants was tested in human type 1 and type 2 macrophages(Verreck et al., 2004; 2006), by determining bacterial counts with the aim to identify mutants with modified intracellular fitness in macrophages. The transcriptome of the macrophages would then be analysed and compared to wild-type infected and uninfected cells. This would enable us to pinpoint host genes that are differentially regulated upon infection with these mutants and may therefore play a key role in phagosome maturation (Ehrt et al., 2001). This would provide deeper insight into the process of phagosome maturation possibly revealing new host targets for therapeutic intervention.

We investigated a set of BCG mutants (Stewart et al., 2005) (see table below). Rv3707c, Rv442c (PPE), Rv2301 (cutinase) and Rv1093 (*glyA*) are unable to control early acidification of the phagosome in the murine macrophage cell line J774 and showed attenuated growth (Stewart et al., 2005). However, no growth attenuation was observed in human type 1 or type 2 macrophages. A urease deficient recombinant BCG vaccine strain genetically modified to express listeriolysin (hly+) (Grode et al., 2005) was also tested. In this strain absence of *ureC* results in increased acidification of the phagosome and thus inhibits maturation of the phago-lysosome, whereas listeriolysin is a protein that enables *Listeria monocytogenes* to escape from the phagosome of infected cells. The aim of this is to improve access of mycobacterial antigens to the MHC I pathway which would result in better CD8 T-cell stimulation and therefore better vaccine efficacy. Investigating the intracellular fitness of ΔureC hly+ rBCG in type 1 and type 2 macrophages we observed a transient reduction in intracellular fitness in both macrophage types at 72-120h post-infection. However, after 120h of incubation outgrowth of *ΔureC* hly+ rBCG was observed. This outgrowth at late time-points is due to increased apoptosis of macrophages, as a loss of cells was observed in the wells infected with *ΔureC* hly+ rBCG. We also tested BCG mutants with *in vitro* growth defects (Beste et al., 2009) (*glpK* KO, *hspR* KO) for their intracellular fitness in human macrophages. The BCG *glpK* KO mutant grew significantly slower than the parental strain in batch culture. The BCG *hspR* KO mutant showed comparable growth to wild type in batch culture, however, it has been reported to be unable to switch to slow growth in chemostat experiments (Beste et al., 2009). Infection of type 1 and type 2 macrophages with the *glpK* KO mutant resulted in higher intracellular survival compared to the parental strain, although the BCG *ΔhspR* showed reduced intracellular fitness in human macrophages. Intracellular killing of BCG *ΔhspR* was more pronounced in type 2 than in type 1 macrophages.

BCG mutants that were reported to show a reduced intracellular survival in the murine macrophage cell line J774 were not attenuated for growth in the human macrophages tested, which points to an important potential difference in the phagosome maturation between murine and human macrophages (Paul et al., 1996) or between cell lines and primary macrophages, and requires further investigation. The recombinant BCG vaccine strain *ΔureC* hly+ rBCG showed a transient decrease in intracellular fitness in human macrophages but induced apoptosis of host cells at later stages of infection. Therefore it will be difficult to isolate sufficient amounts of RNA from host macrophages infected with *ΔureC* hly+ rBCG to perform transcriptome analysis. Growth defects in BCG mutants appear to have a significant effect on their intracellular fitness depending on the nature of the growth defect. Slower growth in batch culture appears to correlate with higher intracellular fitness whereas the inability to switch to slow growth leads to a loss of intracellular fitness. We will therefore study the human macrophage host response of these “growth defect” mutants by analysing the macrophage transcriptome.

To further test our hypothesis that mutants with growth defects *in vitro* have a modified susceptibility to phagosomal killing by macrophages we investigated another two mutants (*M. tuberculosis ΔhspR* and *Δmce1*) with well-characterized *in vitro* growth defects (Beste et al., 2009). H37Rv *Δmce1* grows like the wild type in batch culture; however, in chemostat experiments H37Rv *Δmce1* exhibits a strong competitive advantage compared to the wild type at slow growth rate. Furthermore, H37Rv *Δmce1* has been reported to be hypervirulent in mice *in vivo* (Weber et al., 2000; 2000; Lima et al., 2007). In contrast to these findings we did not observe any increased outgrowth of *M. tuberculosis Δmce1* mutants in human macrophages *in vitro*. We infected human macrophages with *M. tuberculosis ΔhspR* and observed a reduced intracellular fitness similar to its BCG counterpart

We tested addition another two *M. tuberculosis* mutants generated at MPIIB: *ΔRv2017* and *ΔRv2528 mrr*. Rv2017 is possibly a transcriptional regulatory protein, and reduced intracellular fitness was observed in murine macrophages, however no reduced intracellular fitness was found in human type 1 and type 2 macrophages. Rv2528 *mrr* encodes a potential restriction protein, and previous analysis revealed the mutant was slightly hypervirulent relative to wild type in mice *in vivo* and murine macrophages *in vitro*. However, upon infection of human macrophages, *ΔRv2528 mrr* showed only a very low, statistically insignificant, increase in intracellular fitness compared to the wild type.

We further tested the BCG metabolic mutant *ΔnarG*, which lacks the ability to reduce nitrite under anaerobic conditions (Weber et al., 2000). It has been reported that BCG *ΔnarG* fails to persist in immune competent (BALB/c) mice, but this attenuation was not reproduced in in human type 1 or type 2 macrophages. We also included the recombinant BCG strain complemented with RD1. RD-1 is the chromosomal locus that encodes for the ESAT-6 secretion system 1, that was deleted during attenuation of *M. bovis* to produce BCG. Enhanced virulence of BCG::RD1has been reported in immunodeficient mice *in vivo* (Brodin et al., 2006), however our investigations did not reveal any hypervirulence in human macrophages. Recent studies suggest that the RD-1 locus is critical in promoting IL-1beta secretion by infected mouse and human macrophages (Mishra et al., 2010; Dorhoi et al., 2012). Together these observations suggest that bacteria within the phagosome secrete virulence products encoded within the RD1 locus to control macrophage inflammatory responses, without any intracellular fitness penalty.

In summary we found four mycobacterial mutants that modulated phagocytosis and replication in human macrophages: *M. bovis* BCG mutants Rv3696c, Rv0353, recombinant BCG *ΔureC* hly+ rBCG, and *M. tuberculosis* H37Rv Rv0353. Only minor differences were observed between type 1 and type 2 human macrophages. The recombinant BCG strain *ΔureC* hly+ showed a transient decrease in intracellular fitness in human macrophages, but induced apoptosis of host cells at later stages of infection, making it impossible to isolate sufficient RNA amounts to perform host macrophages transcriptome analysis. Interestingly, three of the four mutants that modulate phagosome maturation have confirmed growth deficits. *In vitro* growth defects in BCG and *Mtb* mutants appear to have a significant effect on their intracellular fitness. Slower growth in batch culture appears to correlate with higher intracellular fitness, whereas the inability to switch to slow growth leads to a loss of intracellular fitness. BCG *ΔureC* hly+ shows no growth deficit in batch culture, but remains untested in a chemostat.

The original goal of this workpackage was to analyse the transcriptome of infected macrophages in order to pinpoint host genes differentially regulated upon infection with mycobacterial mutant; this information would lead to a better understand of phagosome maturation and it manipulation by pathogens. *M. tuberculosis* and BCG mutants that influence phagosome maturation in human host cells were identified and characterised, but so far we have been unable to obtain usable data due to the low yield of RNA: one of the selected strains (*ΔureC* hly+ rBCG) induced death of the infected cells, which further complicated the use of the strain in this study. Two of the other strains investigated, BCG*ΔhspR* and H37Rv*ΔhspR*, are carrying the same mutation, which further reduced the number of actual candidate genes to two, even though significant growth differences were observed depending on their background. However, due to the considerable discrepancies observed between the human and mouse macrophages used to test all the mutants, the validation of genes in the mouse model (D5.3) became obsolete.

Table 1: List of mycobacterial candidate genes (BCG strains) investigated for modulation of phagocytosis in host cells

Gene ID of drug target-candidate pathogen genes	knocked out protein/ information on the mutant	Reference	Intracellular fitness in human type 1 and type 2 macrophages
Rv3707c	CHP; unknown protein • predicted extracellular	Stewart et al, 2005	Reduced intracellular fitness was observed for these mutants upon infection of the murine macrophage cell line J774. However, in contrast to these experiments no reduced intracellular fitness was observed in human type 1 and type 2 macrophages.
Rv442c	PPE10; PPE proteins derive their name from a proline–proline–glutamic acid (PPE) motif • the PPE protein family is unique to Mycobacteria and is highly expanded in the pathogenic members of this genus		
Rv2301	cfp25; secreted cutinase		
Rv1093	glyA1; serine hydroxymethyltransferase		
Rv3696c	glpK; glycerol kinase • glpK is essential for glycerol metabolism • mutant exhibits dysgenic growth on plate and in batch culture	Beste et al, 2009	Infection of type 1 and type 2 macrophages with the glpK KO mutant resulted in higher intracellular survival compared to the parental strain.
Rv0353	hspR; heat shock protein repressor gene • negatively regulates several heat shock proteins that are consequently overexpressed in the KO mutant. • Exhibits identical growth during batch phase but significantly reduced fitness during slow growth rate (chemostat)	Beste et al, 2009	The BCG mutant ΔhspR showed reduced intracellular fitness in human macrophages. Intracellular killing of ΔhspR was more pronounced in type 2 than in type 1 macrophages
Recombinant BCG strain	ΔureC hly+ rBCG • absence of ureC results in increased acidification of the phagosome • genetically modified to express listeriolysin (hly), a protein that enables Listeria monocytogenes to escape from the phagosome of infected cells • vaccine candidate in phase I clinical trials	Grode et al, 2005	• reduced intracellular fitness in in human type 1 and type 2 macrophages at 72-120h postinfection • after 120h postinfection outgrowth of ΔureC hly+ rBCG was observed due to increased macrophage apoptosis followed by liberation of bacteria into the medium
BCG::RD1	RD1 is the chromosomal locus that encodes for the ESAT-6 secretion system 1 and was deleted during the divergence of *M. bovis* from *M. tuberculosis.* BCG::RD1 is a recombinant BCG strain that has been complemented with the RD-1 region of *M. tuberculosis.*	Brodin et al. 2006	Enhanced virulence of BCG::RD1 was observed in immunodeficient mice *in vivo.* However, no hypervirulence of BCG::RD1 could be observed in human type 1 and type 2 macrophages.
Rv1161	ΔnarG; nitrate reductase • ΔnarG failed to persist in in immune competent (BALB/C) mice	Fritz et al. 2002	No reduced or increased intracellular fitness was observed in human type 1 and type 2 macrophages

Table 2: List of mycobacterial candidate genes (MTB strains) investigated for modulation of phagocytosis in host cells

Gene ID of drug target-candidate pathogen genes	knocked out protein/ information on the mutant	Reference	Intracellular fitness in human type 1 and type 2 macrophages
Rv2017	Possible transcriptional regulatory protein	generated at the MPIIB	Reduced intracellular fitness of Δ2017has been observed in murine macrophages. However, no reduced intracellular fitness of Δ2017could be observed in human type 1 and type 2 macrophages
Rv2528 mrr	Probable restriction protein mrr	generated at the MPIIB	Slightly hypervirulent relative to the wild type in mice *in vivo* and *in vitro*. However, no hypervirulence could be observed in human type 1 and type 2 macrophages
Rv0168	mce1; hypothetical - integral membrane protein • no difference in growth rate in batch culture • in chemostat experiments Δmce1 exhibits a strong competitive advantage over the wild type at slow growth rate	Beste et al, 2009	Increased intracellular fitness has been reported for Δmce1 upon infection of mice. However, in contrast to these experiments no hypervirulence of Δmce1 was observed in human macrophages.
Rv0353	hspR; heat shock protein repressor gene • negatively regulates several heat shock proteins that are consequently overexpressed in the KO mutant. • exhibits a growth defect compared to the wild type in batch culture but no growth defects in the chemostat	Beste et al, 2009	ΔhspR showed a slightly reduced intracellular fitness in human macrophages.

Part 4: Potential impact, main dissemination activities and the exploitation of results

As envisaged in the original project aims we have identified a number of hits that may have impact, subject to further development, as therapeutic agents. These may have impact in the treatment of a number of diseases, including cancer, autoimmune disease, and bacterial infection, and the results will be exploited by the respective partner institutions.

We identified inhibitors of MHC class II expression, specifically one that down-regulated gamma-interferon mediated increase in MHC class II expression. The exact molecular mechanism is as yet unknown, this approach may be of interest for controlling autoimmune responses that are mediated and maintained by MHC class II expression, such as multiple sclerosis and colitis. The current hits have the advantage of not affecting normal MHC class II expression, making side effects less likely.

Partners has previously identified the host kinase AKt1 as key in determining the fate of intracellular bacteria. In this project inhibitors of Akt1 with increased potency and specificity. These Akt1 inhibitors could have application in the control of bacterial infections but are also important in oncology where Akt1 is often activated by other factors. A spin-off of this was the identification of another kinase not only essential in the normal activation and proliferation of myeloid cells but also often upregulated in some leukemias. Compounds were identified that interfere with kinase activity and cell proliferation. Some of these activities will be take forward in Spin-out companies.

Significant advances were also made in data analysis tools that have application in other areas of systems biology, where disparate types of date need to be integrated and interrogated together to leverage the maximum output.

Further details of the exploitable foreground are provided in section B of 4.2.

The project has resulted in 14 publications so far, with several more in preparation. These cover all aspects of the project, and it is anticipated that some will have significant impact in their respective fields. The systems biology methods are applicable to a broad range of biomedical fields. We have used open access publishing where possible to maximize the availability of our findings to interested researchers.

We identified 31 dissemination activities involving over 2,000 people at which various aspects of the work were presented to national and international audiences of scientists and the general public.

Details of publications and dissemination events are provided in section A of 4.2.

References

Beste, D.J.V., Espasa, M., Bonde, B., Kierzek, A.M., Stewart, G.R., and McFadden, J. (2009). The genetic requirements for fast and slow growth in mycobacteria. PLoS ONE *4*, e5349.

Brodin, P., Majlessi, L., Marsollier, L., de Jonge, M.I., Bottai, D., Demangel, C., Hinds, J., Neyrolles, O., Butcher, P.D., Leclerc, C., et al. (2006). Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence. Infect Immun *74*, 88–98.

Cerepnalkoski, D., Tashkova, K., Todorovski, L., Atanasova, N., and Dzeroski, S. (2012). The influence of parameter fitting methods on model structure selection in automated modeling of aquatic ecosystems. Ecological Modeling, in Press.

Collinet, C., Stöter, M., Bradshaw, C.R., Samusik, N., Rink, J.C., Kenski, D., Habermann, B., Buchholz, F., Henschel, R., Mueller, M.S., et al. (2010). Systems survey of endocytosis by multiparametric image analysis. Nature *464*, 243–249.

Del Conte-Zerial, P., Brusch, L., Rink, J.C., Collinet, C., Kalaidzidis, Y., Zerial, M., and Deutsch, A. (2008). Membrane identity and GTPase cascades regulated by toggle and cut-out switches. Mol Syst Biol *4*, 206.

Dorhoi, A., Nouailles, G., Jorg, S., Hagens, K., Heinemann, E., Pradl, L., Oberbeck-Müller, D., Duque-Correa, M.A., Reece, S.T., Ruland, J., et al. (2012). Activation of the NLRP3 inflammasome by *Mycobacterium tuberculosis* is uncoupled from susceptibility to active tuberculosis. Eur J Immunol *42*, 374–384.

Ehrt, S., Schnappinger, D., Bekiranov, S., Drenkow, J., Shi, S., Gingeras, T.R., Gaasterland, T., Schoolnik, G., and Nathan, C. (2001). Reprogramming of the macrophage transcriptome in response to interferon-gamma and *Mycobacterium tuberculosis*: signaling roles of nitric oxide synthase-2 and phagocyte oxidase. J Exp Med *194*, 1123–1140.

Grode, L., Seiler, P., Baumann, S., Hess, J., Brinkmann, V., Nasser Eddine, A., Mann, P., Goosmann, C., Bandermann, S., Smith, D., et al. (2005). Increased vaccine efficacy against tuberculosis of recombinant *Mycobacterium bovis* bacille Calmette-Guérin mutants that secrete listeriolysin. J Clin Invest *115*, 2472–2479.

Khorashad, J.S., Wagner, S., Greener, L., Marin, D., Reid, A., Milojkovic, D., Patel, H., Willimott, S., Rezvani, K., Gerrard, G., et al. (2009). The level of BCR-ABL1 kinase activity before treatment does not identify chronic myeloid leukemia patients who fail to achieve a complete cytogenetic response on imatinib. Haematologica *94*, 861–864.

Kocev, D. Ensembles for predicting structured outputs. PhD thesis. 2011. International Postgraduate School Jožef Stefan, Ljubljana, Slovenia.

Kuijl, C., Savage, N.D.L., Marsman, M., Tuin, A.W., Janssen, L., Egan, D.A., Ketema, M., van den Nieuwendijk, R., van den Eeden, S.J.F., Geluk, A., et al. (2007). Intracellular bacterial growth is controlled by a kinase network around PKB/AKT1. Nature *450*, 725–730.

Lasa, I. (2006). Towards the identification of the common features of bacterial biofilm development. Int. Microbiol. *9*, 21–28.

Lima, P., Sidders, B., Morici, L., Reader, R., Senaratne, R., Casali, N., and Riley, L.W. (2007). Enhanced mortality despite control of lung infection in mice aerogenically infected with a *Mycobacterium tuberculosis mce1* operon mutant. Microbes Infect *9*, 1285–1290.

Ly, K.T., and Casanova, J.E. (2009). Abelson tyrosine kinase facilitates Salmonella enterica serovar Typhimurium entry into epithelial cells. Infect Immun *77*, 60–69.

Mishra, B.B., Moura-Alves, P., Sonawane, A., Hacohen, N., Griffiths, G., Moita, L.F., and Anes, E. (2010). *Mycobacterium tuberculosis* protein ESAT-6 is a potent activator of the NLRP3/ASC inflammasome. Cell Microbiol *12*, 1046–1063.

Mochizuki, Y., and Majerus, P.W. (2003). Characterization of myotubularin-related protein 7 and its binding partner, myotubularin-related protein 9. Proc Natl Acad Sci USA *100*, 9768–9773.

Paul, P., van den Hoorn, T., Jongsma, M.L.M., Bakker, M.J., Hengeveld, R., Janssen, L.,

Cresswell, P., Egan, D.A., van Ham, M., Brinke, Ten, A., et al. (2011). A Genome-wide Multidimensional RNAi Screen Reveals Pathways Controlling MHC Class II Antigen Presentation. Cell. *145*, 268-283.

Paul, S., Laochumroonvorapong, P., and Kaplan, G. (1996). Comparable growth of virulent and avirulent *Mycobacterium tuberculosis* in human macrophages *in vitro*. J Infect Dis *174*, 105–112.

Podpečan, V., Lavrač, N., Mozetič, I., Novak, P.K., Trajkovski, I., Langohr, L., Kulovesi, K., Toivonen, H., Petek, M., Motaln, H., et al. (2011). SegMine workflows for semantic microarray data analysis in Orange4WS. BMC Bioinformatics *12*, 416.

Robnik-Šikonja, M., and Kononenko, I. Theoretical and Empirical Analysis of ReliefF and RReliefF. Machine Learning *53*, 23–69.

Stewart, G.R., Patel, J., Robertson, B.D., Rae, A., and Young, D.B. (2005). Mycobacterial mutants with defective control of phagosomal acidification. PLoS Pathog *1*, 269–278.

Szomolay, B., and Shahrezaei, V. (2012). Bell-shaped and ultrasensitive dose-response in phosphorylation-dephosphorylation cycles: the role of kinase-phosphatase complex formation. BMC Syst Biol *6*, 26.

Tashkova, K., Korošec, P., Silc, J., Todorovski, L., and Dzeroski, S. (2011). Parameter estimation with bio-inspired meta-heuristic optimization: modeling the dynamics of endocytosis. BMC Syst Biol *5*, 159.

Van Etten, R., Jackson, P., and Baltimore, D. (1994). The COOH terminus of the c-Abl tyrosine kinase contains distinct F-and G-actin binding domains with bundling activity. J Cell Biol *124*, 325-340.

Verreck, F.A.W., de Boer, T., Langenberg, D.M.L., Hoeve, M.A., Kramer, M., Vaisberg, E., Kastelein, R., Kolk, A., de Waal-Malefyt, R., and Ottenhoff, T.H.M. (2004). Human IL-23-producing type 1 macrophages promote but IL-10-producing type 2 macrophages subvert immunity to (myco)bacteria. Proc Natl Acad Sci USA *101*, 4560–4565.

Verreck, F.A.W., de Boer, T., Langenberg, D.M.L., van der Zanden, L., and Ottenhoff, T.H.M. (2006). Phenotypic and functional profiling of human proinflammatory type-1 and anti-inflammatory type-2 macrophages in response to microbial antigens and IFN-gamma- and CD40L-mediated costimulation. J Leuk Biol *79*, 285–293.

Wang, Y., Miller, A.L., Mooseker, M.S., and Koleske, A.J. (2001). The Abl-related gene (Arg) nonreceptor tyrosine kinase uses two F-actin-binding domains to bundle F-actin. Proc Natl Acad Sci USA *98*, 14865–14870.

Weber, I., Fritz, C., Ruttkowski, S., Kreft, A., and Bange, F.C. (2000). Anaerobic nitrate reductase (narGHJI) activity of *Mycobacterium bovis* BCG in vitro and its contribution to virulence in immunodeficient mice. Mol Microbiol *35*, 1017–1025.

Zoncu, R., Perera, R.M., Balkin, D.M., Pirruccello, M., Toomre, D., and De Camilli, P. (2009). A phosphoinositide switch controls the maturation and signaling properties of APPL endosomes. Cell *136*, 1110–1121.