MODEL-IN Report Summary

Project ID: 222008
Funded under: FP7-HEALTH
Country: United Kingdom

Final Report Summary - MODEL-IN (Genomic determinants of inflammation: from physical measurements to system perturbation and mathematical modeling)

The MODEL-IN consortium was set up to foster the integration of elite European research teams from experimental and computational research communities to create a unique world-leading programme in the systems biology of inflammation.

Inflammation is an integrated biochemical and cellular response essential for the viability of complex organisms because of its requirement for both anti-microbial defence and protection from the consequences of tissue injury. However, uncontrolled or sustained inflammation leads to several diseases with a huge social impact, ranging from autoimmune diseases to septic shock and cancer. Inflammation requires that a complex gene expression program be deployed immediately after the encounter with inflammatory stimuli. Essential for the activation of this program are inflammatory transcription factors (TF), namely DNA-binding proteins that are activated by inflammatory stimuli, bind DNA and change the level of expression of genes whose products are involved at different stages of the inflammatory process. While the identity of many of such TF is known, quantitative models describing how their interactions with genomic sequences leads to normal or pathological outcomes are not available. Unravelling how the interplay between such TFs and genomic organisation underlies the orchestrated expression of hundreds of inflammatory genes is therefore crucial to understand how to pharmacologically tune pathological inflammatory responses.

The ??DEL-IN consortium made use of innovative technologies and computational approaches to generate quantitative genomic data and to integrate them into models describing the spatiotemporal coordination of transcription of hundreds of inflammatory genes by multiple TFs. Scientific successes achieved by ??DEL-IN include the identification of TFs involved in the control of genomic organisation of inflammatory cells (such as macrophages), the production of many genomic data sets of immediate relevance to the scientific community, the characterisation of TFs controlling the state of macrophage polarisation towards an anti-inflammatory versus a pro-inflammatory state, the generation of novel models describing the activity of inflammatory TFs in a quantitative manner and in the context of the laws of physics, the production and characterisation of mouse models for the in vivo investigation of quantitative aspects of inflammatory gene regulation, and finally the initial generation of novel technical approaches for the detection of protein-DNA interactions.

Significantly, ??DEL-IN provided training for a new generation of European scientists in systems biology as well as fostered interactions, exchange and synergy among the principal laboratories in the field of mammalian transcription. It engaged widely with European systems biology research community by disseminating the results obtained within the project via web-resources and annual workshops in different parts of Europe.

Project context and objectives:

The key objective of the ??DEL-IN consortium was to provide mechanism-based quantitative (physical and mathematical) models linking genomic determinants to transcriptional control of a basic biological process of great relevance to public health, namely the inflammatory response. ??DEL-IN brought together elite European research teams to create a world-
leading programme and to set up new standards for quantitative studies of key immunological processes by employing state-of-the-art multidisciplinary fundamental genomics, computational and molecular biology approaches and by developing novel dedicated computational and wet technologies for the analysis of inducible gene expression.

Inflammatory responses are first initiated and then maintained by transcription of genes whose products recruit and activate leukocytes, increase vascular permeability, amplify the response, protect inflammatory and tissue cells from apoptosis and eventually orchestrate tissue repair. Regulation of transcription and gene expression for many inflammatory cytokines involves combinations of TFs including those of the NF-kB, IRF and AP-1 families. The activity of a given transcription factor is determined first of all by its binding to a specific DNA sequence located either in the vicinity of target genes or at variable distances. DNA binding determines an array of distinct biological outcomes, ranging from the reorganisation of nucleosome positioning to the recruitment of the transcriptional machinery to activate gene expression. Therefore, basic events at the heart of every inflammatory response, either self-resolving or excessive and sustained, occur in cell nuclei and are underpinned by the dynamic relationships between genomic cis-regulatory determinants (namely DNA sequences) and TFs interacting with them.

Therefore, one of the basic assumptions that justified the organisation of this project was that the first step towards a mechanistic understanding of inflammatory gene control was the accurate and quantitative annotation of the specific interactions between inflammatory TFs and their cognate binding sites (TFBS). A given TF typically recognises a specific but often relatively degenerate sequence pattern called a binding motif, which is usually 6-20 base pairs in length. In the motif, some nucleotides tend to occur more often than others at specific positions. Different TFs prefer different binding motifs, and multiple TFs can bind cooperatively to a cis-element that contains several different TFBSs clustered together. Moreover, TFs of the same family tend to bind similar sites and slight nucleotide preferences are often overlooked in standard analyses. At any time, the particular composition of transcription factors active in the cell nucleus determines which subset of cis-elements is bound and which genes are activated. The cis-regulatory elements are thus the hardwired 'control logic' in the genome. Although most protein-coding genes in the human and mouse genomes have been identified and annotated, the location, properties and physiological behaviour of cis-elements that control their expression are still largely unknown in spite of enormous advances in genomic sciences. The identification and characterisation of such elements and the ability to model their function is a key challenge of genomic biology. As detailed below, among the successes of the ??DEL-IN consortium was the quantitative characterisation, using a combination of distinct and novel high-throughput approaches, of the panel of binding sites specific for different inflammatory TFs, particularly (but not only) those of the NF-kB family. Specifically, we identified the slight differences in binding specificity that contribute to explain the unique biological activities of each TF of this family. This result is key to understand the in vivo specificity of these TFs and generate comprehensive models describing their activity.

A second major success was the identification of the regulatory logic of the 'context' in which these TFs acts, namely the specialised genome of inflammatory cells. Indeed, we found that Pu.1 a TF controlling the differentiation of macrophages (one of the essential inflammatory cell types), determines the identity of the specific genomic regulatory elements accessible in this cell type to inflammatory TFs, thus creating a macrophage-specific context or landscape in which these TFs act. This result was key to understand the molecular bases of the uniqueness of the macrophage inflammatory gene expression program. Moreover, from a conceptual point of view, this result supported the important notion that lineage-determining and highly cell-type specific TFs (commonly defined as master regulators) functionally interact with TFs activated by inflammatory stimuli (which are ubiquitous and therefore not cell type-specific) and in fact control the final transcriptional outcome of their activation. The molecular determinants of the selective ability of distinct subtypes of macrophages to activate pro-inflammatory rather anti-inflammatory gene expression programs have also been defined within the ??DEL-IN consortium, with the identification of the TF Irf5 as the main determinant of the pro-inflammatory macrophage subtype (usually indicated as M1).

Another important objective of this consortium was to create integrative models providing quantitative descriptions of the
functional activity of TFs induced or activated in response to stimulation. The relevance of such models would be to provide a framework for the correct quantitative (rather than qualitative) prediction of transcriptional responses to inflammatory agonists, namely predictions of the intensity of transcriptional outputs relative to the strength or dosage of the agonist. One such model we generated and reported was based on the translation to this biological response of a well-known physical model (the one-dimensional Ising model of ferro-magnets). This model relies on equilibrium statistical mechanics, motivated by the experimental observation that TFs operates at equilibrium. When active TF concentration varies, processes at gene regulatory elements (such as promoters) equilibrate kinetically, with changes in concentration being rapidly translated into changes in downstream gene transcription. The Ising model was originally developed as a simplified model of ferromagnetic materials, which contain ordered arrangements (lattices) of atoms possessing a magnetic moment (spin), and was later adapted to diverse physical systems. We demonstrated that in its one-dimensional (1D) formulation, it is formally equivalent to the problem of the equilibrium binding of a TF to a cluster of specific binding sites in a promoter sequence, the cluster of binding sites in fact corresponding to an ordered arrangement (i.e. a lattice) of atoms. Thus, we adapted the known solutions of the 1D Ising model to calculate the average value of the occupancy state of the RNA polymerase and TF binding sites. As compared to other existing approaches, this model has several advantages. First, it returns explicit expressions for the observables of interest, with no need for computer simulations, even in the presence of binding cooperativity. Second, once the number of sites and the TF binding constant are fixed, it contains a limited number of parameters all having a clear biological significance. Third, our approach allows calculating in a straightforward manner several important properties of the system, most notably the fluctuations in Pol II occupancy (which are related to the intrinsic noise in transcription).

Important objective of ??DEL-IN was also implementation of training activities that engaged a new generation of the systems biology scientists, such as periodic meetings and open workshops. During its run, ??DEL-IN trained 22 postdoctoral scientists and 11 postgraduate students in the state-of-the art functional genomic and computational technologies, as well as biochemistry, imaging, cell biology.

Project results:

Defining binding specificity of TFs controlling the inflammatory response

A major objective of the ??DEL-IN consortium was to define the precise binding specificities of TFs activated by inflammatory stimuli and directly controlling inflammatory gene transcription. This endeavour required the set up of novel technologies specifically dedicated to the in vitro analysis of DNA binding by TFs expressed in either bacterial or eukaryotic systems. An additional element of complexity was the fact that most relevant species of the pivotal inflammatory transcription factors of the NF-kB family are in fact heterodimers composed of the products of two distinct NF-kB genes. Therefore we had first of all to develop specific procedures for the in vitro production of NF-kB and IRF protein dimers. These proteins were used for in vitro binding assays using distinct technological approaches. The first one consisted in microarrays with approximately 15 000 probes (DNA binding sites) each. Each array contained several internal controls from the manufacturer (Agilent), as well as positive and negative controls for NF-kB and no motif probes. DNA in arrays was double stranded prior to hybridisation.

Data allowed us to establish the correlation between the relative binding affinities of the NF-kB family to the different sequences used. The proteins were compared with each other by calculating pair-wise correlation coefficients. The correlation coefficients between the pairs vary considerably according to the proteins. The pairs p50:RelAp50, RelAp52:RelBp52 and cRelp52:RelAp52 all showed correlations above 0.94; on the other side of the spectrum, RelA:RelBp50 showed a low correlation coefficient (0.52).

Within a collaboration between ??DEL-IN and Marta Bulyk's group (Harvard, United States (US)) we obtained NFkB protein binding data using an alternative protein binding microarray design. At present the work has produced the following binding site logos for several NF-kB protein combinations.
In order to obtain a more comprehensive analysis of the binding specificities of inflammatory TFs, ??DEL-IN developed a novel technological tool based on high-throughput sequencing, EMSA-Seq.

Conceptually, this technique capitalises on the well established EMSA assay for the isolation of nucleic acid sequences that have been bound by a transcription factor (TF). This material is then subjected to deep sequencing using Illumina’s SOLEXA platform. Computational analysis firstly reveals the complexity and extent of enrichment for individual DNA sequences bound by the TF. It can also infer the binding affinities a TF has for specific DNA motifs found within these sequences. The results from EMSA-Seq were highly complementary to information obtained from protein-DNA microarrays and ChIP-Seq platforms.

Our deep sequencing approach produced enough data to allow an exhaustive representation of every possible sequence up to a length of 11-mers. Approximately 10 - 13 % of all possible 11-mer combinations were bound by each of the three RelA-containing dimers. Binding models representing the 50 and 1000 highest affinity binders were created for each dimer. The profile of RelARelA was distinct from that of the heterodimers, RelAp50 and RelAp52. The proportions of sequences bound by RelARelA that are in common with RelAp50 and RelAp52 are 61 % and 63 % respectively. By comparison, RelAp50 and RelAp52 share 81 % of bound sequences. This is consistent with what we observed using microarrays in which binding profiles of the two RelA heterodimers are more similar to one another than when compared to that of the RelA homodimer. Moreover, binding specificities of selected IRF family transcription factors were generated, thus significantly expanding our knowledge of the sequence determinants of specific binding of inflammatory TFs.

To estimate the NF-kB-binding potential as measured by EMSA-Seq for the interpretation of in vivo NF-kB binding, we overlaid dimer-specific 11-mers from our datasets onto all binding region-summits (BRSs; methods) from both a study by Kasowski and co-workers (Kasowski et al., 2010) and datasets generated by the ??DEL-IN consortium. 11-mer binders identified by EMSA-Seq were mapped onto a 300bp-region, the BRS that is centred on the summit-point within a BR, indicating that our procedures effectively identified bona fide high affinity sites.

To prove that the newly identified sequences were bona fide TF binding sites, we focused on those most deviant from the already known sites and analyzed them by ultraviolet (UV)-laser cross-linking, confirming in all cases the EMSA-Seq data.

Relationship between nucleosomal occupancy and binding of inflammatory TFs

A second layer of objectives of ??DEL-IN referred to the interplay between nucleosomal occupancy and ability of TFs activated by inflammatory stimuli to bind DNA. This activity involved multiple in vitro and in vivo approaches.

Our data evidence that NF-kB was unable to invade remodeled nucleosomes when the NF-kB recognition sequence was inserted close to the nucleosomal dyad. By contrasts NF-kB is able to bind to his cognate sequences at the edge of the nucleosome core particle. Sliding of the nucleosome by remodeling complexes (RSC) made the nucleosomal DNA accessible to the transcription factor, since the binding site is located at the edge of the slid nucleosome. Interestingly, upon nucleosome dilution driven removal of the H2A-H2B dimer, NF-kB was able to interact with the H3-H4 tetrameric particle (Figure, compare lanes 12-23 with lanes 24-29. This suggests that removal of H2A-H2B dimers from the nucleosome is required for proper NF-kB binding.

We have next studied if and how linker histone H1 affects the interaction of NF-kB with reconstituted nucleosome. To shed light on the effect of histone H1 on NF-kB interactions, we developed a novel method for reconstitution of histone H1-containing nucleosomes and mapped at a single base pair resolution the binding of H1 to nucleosomal DNA. This allowed us to analyse the role of histone H1 in the interactions of NF-kB proteins with H1-containing nucleosomes. Our data revealed that in H1 containing nucleosomes NF-kB displaces histone H1 and binds to the high-affinity cognate sites located either at the core nucleosome edge or in the linker DNA. Noteworthy, H1 was unable to bind to nucleosomes containing already bound NF-kB at the nucleosome core edge or in the linker DNA.
In parallel to in vitro analyses, the DEL-IN consortium generated the first nucleosomal occupancy maps in inflammatory cells, specifically macrophages, using an MNase-Seq approach. This approach was made possible by the increased throughput of current multiparallel sequencers (before their availability this type of experiments was excessively expensive). We compared two levels of MNase digestion (20 min and 100 min) to understand how the protocol affects the quality of the results and to map nucleosome positions in inflammation genes. The sequencing resulted in an average of 164 M reads/sample, of which 152 M were reliably mapped to the mouse genome.

From these data, it is still not trivial to generate reliable genome wide nucleosome maps since each nucleosome is covered by 2-10 reads on average, still for specific regions with relatively high coverage single nucleosome positions can be determined. As expected, the 100 min sample exhibits stronger boundaries between adjacent nucleosomes leading to clearer nucleosome calls.

We generated 'phasograms', histograms of distances between mapped reads' start positions of the two digestion levels. The phasogram exhibit a wave-like pattern where the period represents the genome-wide average spacing between adjacent nucleosomes. The decay at longer distances shows the effect of partially localised nucleosomes and variable spacing (e.g. nucleosome depleted regions). We observe a period of about 190 bp in both samples, matching the expected center-to-center distance between nucleosomes. Importantly, more mono-nucleosomal distances of < 200 bp were observed in the longer digestion sample, signifying the higher resolution of the 100 min MNase digestion sample.

To further explore nucleosome organisation in our data we examined the patterns around several cis regulatory elements, which are specific for bone marrow tissue. We then asked whether active enhancers are characterised by a distinct nucleosomes pattern, we used the recently published tissue specific enhancers data from the Ren's lab and plotted the average patterns of 17 enhancers groups including bone marrow specific enhancers.

We first computed the average pattern around predicted CTCF sites. As expected we find a periodicity of about 190 bp from both ends of CTCF insulator with a depletion over the binding cite. The overrepresentation of 'well defined' nucleosomes which are adjacent to some barrier (established either by a protein or by a sequence element) in the 20 min sample is apparent here as well. Another example of this observation is the average profile next to Polymerase II (PolII) binding sites. Downstream to PolII peaks there is (on average) a well-localised nucleosome, supporting a contact between this nucleosome and Pol2 initiation complex. However, 150 bp upstream to the binding site there is a depletion in the 100 min sample but strong signal in the 20 min sample, suggesting that we either digest this nucleosome in the 100 min sample or that we capture the binding of other members of the initiation complex and not a real nucleosome.

All 16 enhancers groups, which are not active in bone marrow cells, showed signal similar to randomly selected positions and the bone-marrow specific enhancers contain a strong well-positioned nucleosome. It is important to note that enhancers were detected using chromatin pattern, i.e. the presence of H3K4me1 but absence of H3K4me3, thus the peak we detect in the bone-marrow group is probably due to the alignment of the enhancers around the H3K4Me1 peak. Since our data includes all nucleosomes, regardless of histones modification states, we can fairly confidently assume that the random pattern we find in all the other groups is due to different nucleosome positions in inactive enhancers. In conclusion, we derived the first nucleosome map for the relevant cell type for our study of inflammation response. This detailed map provides crucial reference for our study of enhancers and promoters during inflammation response.

Nucleosome downregulation in inflammation and the chromatin protein Hmgb1

HMGB1 is a nuclear protein that interacts with nucleosomes and several transcription factors, including NF-kB. HMGB1 is also a secreted signalling molecule that triggers responses to cell and tissue damage. LPS stimulation of monocytic cells induces HMGB1 secretion, thus leaving the nucleus almost completely devoid of HMGB1. We therefore assessed the consequences of
First, we found that mammalian cells lacking HMGB1 (a proxy for cells that have removed HMGB1 from the nucleus as a response to inflammation) have a reduced amount of histones and nucleosomes.

According to the hypothesis formulated in the corrective action, we asked whether histone genes are transcriptionally down-regulated in macrophages exposed to inflammatory stimuli. MEFs treated with TNF-a for 4 to 24 hrs show a strong up-regulation of inflammatory genes but absolutely no down-regulation of Histone H3 genes.

By Illumina RNA-Seq profiling, we found that the transcription levels of approximately 1400 genes (18%) significantly differ when nucleosomes are fewer in Hmgb1-/- cells. 786 and 564 genes are up- and down-regulated, respectively. Out of these, 141 genes belong to pathways that are involved in the control of the inflammatory response.

Overall, this result suggests that HMGB1 secretion in inflammatory cells might decrease the nucleosome number as a mean to facilitate the expression of genes that are induced by inflammatory stimuli.

Genome-wide identification of cis-regulatory elements controlling inflammatory gene expression and TF binding in vivo.

A major output of the ??DEL-IN consortium was the first characterisation of the genomic repertoire of regulatory elements controlling the expression of inflammatory genes in mouse macrophages. We generated ChIP-Seq data sets for the transcriptional coactivator and histone acetyltransferase p300 in untreated and LPS-stimulated macrophages together with data set for a panel of histone modifications and RNA polymerase II. Computational analyses on the LPS-inducible p300 peaks allowed us to identify a general rule in enhancers controlling inflammatory genes: at these enhancers, binding sites for the macrophage master regulator Pu.1 are combined with binding sites for ubiquitous transcription factors induced by inflammatory stimuli like interferon regulatory factors (IRFs), NF-kB and AP-1. These data provided a paradigm of how inducible responses can be adapted to the specific cellular context / landscape in which they are elicited, thus explaining from a molecular point of view the cell type-specificity of the inflammatory response elicited by a given agonist. ChIP-Seq data sets for the macrophage master regulator Pu.1 also showed that the vast majority of enhancers in macrophages are occupied by this TF, which is necessary and sufficient to induce small nucleosome-depleted sequence stretches bracketed by nucleosomes marked by H3K4me1 (histone H3 monomethylated at K4), an enhancer-specific histone histone modification.

In fact, new TF (NF-kB, IRFs) binding events induced by inflammatory agonists in macrophages and dendritic cells all occur within the basal repertoire of regulatory elements generated by Pu.1 as demonstrated by our own data sets for p65 (RelA) and Irf family proteins.

We have demonstrated a role for IRF5 in determining M1 macrophage lineage commitment. M1 macrophages are characterised by high level of IRF5, expression of which is induced during their differentiation. The induction of M1-specific cytokines is impaired in human M1 macrophages with levels of IRF5 expression reduced by siRNA knock-down or in the peritoneal macrophages of the Irf5-/- mice. Conversely, forced expression of IRF5 in M2 macrophages drives global expression of M1-specific cytokines, chemokines and co-stimulatory molecules and leads to a potent TH1-TH17 response.

The molecular mechanisms of IRF5 function in transcriptional regulation of pro-inflammatory genes are not fully understood but they appear to involve interactions with NF-kB RelA at some genes. We examine whether this is a common mechanism of controlling highly inducible expression of pro-inflammatory genes in M1 macrophages by inspecting a genomic picture of the IRF5 and RelA binding in bone marrow derived GMCSF-polarised M1 macrophages before and after stimulation with LPS. We discover co-occurrence of IRF5 and RelA binding specifically at the promoters of highly induced pro-inflammatory genes, with individual RelA or IRF5 ChIP-seq peaks being associated with more broadly expressed genes.
Moreover, using in vivo and in vitro motif discovery analyses, we demonstrate that RelA and IRF5 aligned peaks are best explained by the consensus NF-kB and ISRE binding sites, whereas the individual IRF5 ChIP-seq peaks may encompass non-consensus IRF5 binding sites. Our data suggest that a predominant mechanism of IRF5 recruitment to DNA is via its interactions with co-factors, e.g. NF-kB RelA.

Modelling inflammatory gene expression

An integrated algorithm for dynamical inflammatory gene expression data.
The precise dynamics of gene expression is often crucial for proper response to stimuli. Time-course gene-expression profiles can provide insights about the dynamics of many cellular responses, but are often noisy and measured at arbitrary intervals, posing a major analysis challenge. IEO developed an algorithm that interleaves clustering time-course gene-expression data with estimation of dynamic models of their response by biologically meaningful parameters. In combining these two tasks, we overcome obstacles posed in each one. Moreover, our approach provides an easy way to compare between responses to different stimuli at the dynamical level.

This mathematical model was applied to gene expression data obtained from mouse bone marrow-derived macrophages and dendritic cells, as well as human monocyte-derived macrophages. Mathematical modelling of the LPS response in primary macrophages indicated eight types of response dynamics, both in wild-type cells and in NF-kB knockouts (RelA), and extract a concise representation of the different dynamical response types. To create a common set of prototypes for both wild-type and knockout cells, we applied our clustering algorithm to both datasets together (each gene was represented twice, once for each cell type). This resulted in a set of eight prototypes that differ from each other in their onset time, expression ratio, and direction (up or down regulation).

We next analysed the similarities and differences between the wild type and knockout cells to identify the involvement of NF-kB (RelA) in the response. Since we applied our method simultaneously to the two different time series data, we can now compare for each gene, its dynamical expression profile, by comparing the cluster it is assigned to in wild-type versus the knockout. Some genes, such as the Nfkbia gene, are not affected at all by the knockout of NF-kB (RelA), indicating that their regulation is independent of RelA or can be compensated by other factors. Other genes are completely dependent on RelA for their expression, such as the gene Cxcl3. In some cases the genes are expressed in response to inflammatory signals also in the knockout cells, however their dynamic expression pattern changes, indicating a partial and important involvement of RelA in their regulation. Two examples for such genes are the Icam1 gene in which the time of the response is dependent on RelA, and Ptgs2 gene in which the amplitude of the response depends on RelA. Interestingly, we hardly find any genes where both the time of the response and the amplitude are dependent on RelA.

Physical models describing NF-kB activity in the context of inflammatory gene expression. Models were developed for the occupancy of NF-kB and its co-factors on target promoters. We first found that the presence of more than one NF-kB binding site is necessary to ensure that the induction profiles of NFKBIA can be fitted by the different model variants. Specifically, we show that a model with only one binding site would not predict correctly the convexity of the observed NFKBIA induction profile.

Next, we note that the three models can be parameterised to fit each other. However, the regions of parameter space that return fits with similar quality are clearly separated. We plotted the rmsq value as a function of the imposed binding constant KA and the best estimate of parameter N (setting the strength of binding cooperativity) obtained in the fit. Parameter sets that return fits with similar rmsq distances to the data have significantly different values of the NF-kB binding constant KA (30 nM in the case of model 1 versus 300 nM in the case of model 3) and the strength of binding cooperativity (N = 0.8 kBT for model 1 versus N= -0.2 kBT for model 3). These two parameter sets are sufficiently separated for the two models to produce qualitatively and quantitatively different predictions for the profile of NF-kB recruitment to the cluster, which we directly challenged in an independent experiment. In the case of the two parameter sets, the root mean square distances of model
predictions against the ChiP experiment were significantly different, and supporting model 3 with a rmsq=0.0012 versus 0.0021 for model 1.

We next examined the effect of the number of NF-κB binding sites on the induction curves and on the dynamic range and transcriptional noise of the target genes.

Finally, we show the experimental predictions of model number 3 in comparison to the measured transcriptional output.

Challenging physical models with genetics in the mouse

As discussed above, NF-κB binding sites occur in highly conserved homotypic clusters in the cis-regulatory regions of many validated NF-κB-dependent genes. Among the genes containing clustered NF-κB binding sites, we selected the nfkbia gene encoding IkBa - which is the main inhibitor of the pathway and is under NF-κB transcriptional control providing negative feedback to the system - for mutagenesis by using a knock-in approach. The nfkbia gene contains the highest number of highly conserved kB sites organised in two clusters of six sites each (promoter proximal, 6 kB sites; first intron 6 kB sites). To study the role of clustered NF-κB binding sites in the transcriptional regulation of the nfkbia gene, we developed a targeting strategy designed to introduce point mutations to the NF-κB binding sites of the endogenous nfkbia gene.

Four DNA constructs for the targeting of the two NF-κB binding site clusters of the nfkbia gene (1-pBSIkBa1-6Ex5, 2-pBSIkBa7-12Ex5, 3-pBSIkBa1-12Ex5, 4-pBSIkBaEx5) were generated. In construct 1 all NF-κB sites on the promoter cluster are mutated, in construct 2 all NF-κB sites in the intronic cluster were mutated, while in construct 3 all 12 sites in both clusters were mutated. Construct 4 is the wild type control that is designed to introduce a silent mutation in exon 5 of the nfkbia gene but no mutations in NF-κB clusters. This mutation was introduced into all targeting vectors and is needed to distinguish the product of the targeted allele of the gene from the product of the endogenous allele in heterozygous cells.

Mice bearing homozygous mutations in all NF-κB binding sites of cluster 1, cluster 2 or both clusters in the endogenous nfkbia gene are viable and do not show any macroscopic phenotypical abnormalities, indicating that mutation of the NF-κB site clusters of the nfkbia gene does not affect mouse development and physiology, at least in young age.

To study the effect of the NF-κB site cluster mutations in the transcription of the nfkbia gene, we stimulated primary bone marrow derived macrophages (BMDM) prepared from mice carrying homozygous NF-κB site mutations in the nfkbia gene with LPS or TNF. Cell carrying homozygous mutations of NF-κB sites in both clusters (IkBa 1-12) showed dramatically reduced upregulation of IkBa expression after stimulation with either LPS or TNF. Therefore, mutation of both NF-κB site clusters very strongly inhibited, but did not completely prevent, the upregulation of IkBa gene transcription in response to these stimuli. Cells carrying mutations in the NF-κB sites of cluster 1 (IkBa 1-6) also showed strongly reduced expression of IkBa upon stimulation with LPS or TNF, but the extent of inhibition was not as strong as in the case of the IkBa 1-12 cells. Cells carrying mutations of the NF-κB sites of cluster 2 (IkBa 7-12), located in the first intron, showed only mildly reduced expression of IkBa after LPS stimulation. However, upon TNF stimulation the mutation of cluster 2 inhibited IkBa gene transcription to an extent similar to the inhibition observed in cells with mutation of cluster 1.

These results show that the promoter cluster (sites 1-6) has a primary role in regulating IkBa gene transcription in response to the pro-inflammatory stimulus LPS in BMDMs, while the intronic cluster (sites 7-12) is dispensable for IkBa transcription under this condition. In addition, our results show that both cluster 1 and cluster 2 participate to a similar extent to the transcriptional induction of the gene upon TNF stimulation, suggesting that cluster 2 is as important as cluster 1 in driving IkBa gene transcription in response to TNF. Therefore, cluster 1 appears to confer strong transcriptional control to the IkBa gene in response to different stimuli, while cluster 2 appears to be important for IkBa gene transcription in response to specific stimuli. Collectively, our results suggest that the maximal level of transcriptional upregulation is controlled by the additive action of both clusters. Mice carrying mutations in both NF-κB site clusters on the IkBa promoter develop normally and
do not show any abnormalities at least in young age. We hypothesised that the NF-κB site clusters could become important under conditions of experimental challenge, where NF-κB signalling is induced systemically in response to microbial components and cytokines. Therefore, to analyse the in vivo relevance of the two NF-κB site clusters in the IkBa locus especially on the regulation of immune responses, we applied the model of endotoxin induced septic shock. This model relies on the administration of endotoxin (LPS) to the mice, which causes the strong systemic activation of NF-κB and other signalling pathways resulting in the release of many proinflammatory cytokines and chemokines in the circulation. Here we injected a sublethal dose of LPS (25 μg/g bodyweight) IP into age matched female mice of the four different genotypes and monitored survival closely for a timespan of 72 hours.

About 70 % of the IkBa Ex5 control mice survived this endotoxin challenge, showing that under our experimental conditions this dose of endotoxin caused 30 % lethality in control mice. However, all mice carrying homozygous IkBa 1-12 mutations died within 30 hours after injection, demonstrating that mutation of both clusters strongly sensitises mice to LPS-induced septic shock. In addition, all mice carrying IkBa 1-6 mutations died within 50 hours after injection of the same dose of LPS, demonstrating that the promoter NF-κB site cluster has a major role in controlling IkBa expression levels that are important to prevent an overactivation of the immune system and cause lethal septic shock. Interestingly, although our previous in vitro experiments suggested, that mutation of the intronic cluster only had a minor effect on the inducibility of IkBa upon LPS treatment, mice carrying homoyzgous mutation of the intronic cluster 1 (IkBa 7-12) were also highly sensitive to LPS-induced septic shock, with about 90 % of these animals dying after administration of a similar dose of LPS. Thus, while the promoter cluster appears to be primarily responsible for the regulation of IkBa gene transcription in response to LPS in macrophages, both cluster 1 and cluster 2 are required for optimal control of the systemic immune response and mutations in either cluster sensitise mice to a systemic challenge in an endotoxin-induced model of septic shock. These in vivo results highlight the importance of an accurate control of the NF-κB signalling system by the action of its negative inhibitor IkBa and show that both binding-site clusters are vital for the induction of sufficient mRNA and subsequently protein amounts to prevent overactivation of the immune system upon challenge with pro-inflammatory mediators.

Overall, data obtained in this mouse model are fully consistent with the proposed analogue (as opposed to a digital, i.e. on-off) model of function of κB sites within the cluster, in that genetic ablation of a subset of six sites reduced Nfkbia activation and elimination of all of them was required to completely shut-off the gene. However, our thermodynamic model was not deigned to take into account a qualitative aspect of this analogue behaviour, namely the existence of an obvious hierarchy between the two κB site clusters, in which the promoter cluster gave a higher contribution than the intronic cluster to the final gene activation. The definitive refinement of the model on a physical-mathematical basis will require the generation of mice in which individual sites from each clusters are ablated.

In summary, the ??DEL-IN consortium has achieved the key objective of developing mechanism-based quantitative (physical and mathematical) models linking genomic determinants to transcriptional control of the inflammatory response.

Potential impact:

The dissemination actions were undertaken at several levels:

i) publications in per reviewed journals;
ii) presentations at conferences and workshops;
iii) organisation of workshops; and
iv) disseminations to the public.

In terms of peer reviewed papers, the consortium produced so far 18 publications in the most prestigious journals (such as Nature Biotechnology, Nature Immunology, Molecular Cell, PNAS and others). As such the project produced key papers and technologies on aspects of gene regulation that contributed to the advancement not only of the immunology and genomics fields but are of interest to a wide audience and have implications for biomedical sciences overall. The project had an impact
in promoting cooperation between academic and industrial sector, as it aimed to explore the nanopore technology for
detection of protein-DNA interactions at single molecule level. Although the cost of creating nanopore chips was prohibitive in
terms of funding available within ??DEL-IN, the inroads were made to improve the cost throughout the commercialisation
process, allowing for the future development of a sensor device for protein-DNA or even protein-protein interactions.

Aspects of the work were also presented in 19 international conferences and workshops including high profile conferences
organised at Keystone and the Cold Spring Harbour Laboratory.

The consortium organised two workshops on the 'genomic determinants of inflammation'.

The first workshop was organised in Oxford, United Kingdom (UK). It lasted one full day, included 32 participants and 10
speakers. The second workshop was organised in Athens and included eight international speakers as well as presentations
form postdoctoral researchers and students of the two days. The workshop gave opportunity for junior and senior researchers,
located in Greece, as well as in several European countries, to meet and exchange ideas. It also promoted discussions in
separate sessions and supported young researchers by providing travel funds and free attendance.

Two public releases were made to inform lay audience about the activities of the consortium, as well as targeting an academic
oriented audience. Leaflets were distributed at all academic institutions taking part, in order to disseminate the information to
students and other principal investigators.

The project had a major impact in academic training of young researchers. 11 PhD students have received their training within
the framework of MODEL-IN. In particular the students had the opportunity to interact with top level scientists from different
European Countries and receive mentoring and advice accordingly. MODEL-IN also provided an ideal networking opportunity
for students and early career researchers through direct collaborations between laboratories and participation in workshops.
Particularly, pleasing was the engagement of women in the project, which has maintained a very good gender balance for
early career researches (post doctoral and PhD) throughout (15 women to 18 men). For the PIs and their laboratories it also
allowed further advancement both in applying the latest technologies, but also through discussions and brain storming
sessions. These interactions are reflected in the resulting publications where three - four different laboratories have typically
collaborated to produce the work, as demonstrated by the author's affiliations.

The project had an impact in promoting cooperation between academic and industrial sector, as it aimed to explore the
nanopore technology for detection of protein-DNA interactions at single molecule level. Although the cost of creating nanopore
chips was prohibitive in terms of funding available within ??DEL-IN, the inroads were made to improve the cost throughout the
commercialisation process, allowing for the future development of a sensor device for protein-DNA or even protein-protein
interactions. A further outcome of this proposal is that the network is currently planning follow up grant applications both at
national and EU level.

Overall, the impact of the project was very high in terms of training young and early carrier researchers and contributing to
the wider scientific community with landmark, high quality publications that are contributing to the overall progress in science
and research.

Public website address: http://www.model-in.org

**Related information**

<table>
<thead>
<tr>
<th>Result In Brief</th>
<th>The genomic determinants of inflammation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documents and Publications</td>
<td>Final Report - MODEL-IN (Genomic determinants of inflammation: from physical measurements to system perturbation and mathematical modeling)</td>
</tr>
</tbody>
</table>
Reported by
THE CHANCELLOR, MASTERS AND SCHOLARS OF THE UNIVERSITY OF OXFORD
Roosevelt Drive
OX3 7BN OXFORD
United Kingdom
See on map

Subjects
Biotechnology - Medicine and Health - Scientific Research

Last updated on 2013-02-01
Retrieved on 2019-07-26

© European Union, 2019