molecular systems biology

Systems Analysis of Eleven Rodent Disease Models Reveals an Inflammatome Signature and Key Drivers

I-Ming Wang, Bin Zhang, Xia Yang, Jun Zhu, Serguei Stepaniants, Chunsheng Zhang, Qingying Meng, Mette Peters, Yudong He, Chester Ni, Deborah Slipetz, Michael A. Crackower, Hani Houshyar, Christopher M. Tan, Ernest Asante-Appiah, Gary O'Neill, Mingjuan Jane Luo, Rolf Thieringer, Jeffrey Yuan, Chi-Sung Chiu, Pek Yee Lum, John Lamb, Yves Boie, Hilary A. Wilkinson, Eric E. Schadt, Hongyue Dai and Christopher Roberts

Corresponding author: Bin Zhang, Sage Bionetworks
and Ming Wang, Merck Research Laboratories

Review timeline:

Submission date:	26 August 2011
Editorial Decision:	11 October 2011
Resubmission:	15 March 2012
Editorial Decision:	20 April 2012
Revision received:	22 May 2012
Accepted:	25 May 2012

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 11 October 2011

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees whom we asked to evaluate your manuscript. As you will see from the reports below, the referees raise substantial concerns on your work, which, I am afraid to say, preclude its publication.

The most substantial concerns were raised by the second reviewer who clearly indicated that they were not convinced that this work currently presented a decisive advance in our understanding of gene expression patterns associated with inflammation, and the overlap and novelty of this inflammatome expression signature compared to previous works remained to be rigorously demonstrated. The other two reviewers had more specific, but important concerns, regarding the manner in which consensus networks were constructed and key aspects of the expression profiling methodologies. These two reviewers both rated the validity of the conclusions reached in this work and its suitability for publication as "medium," indicating that they felt that these concerns were important enough to cast some doubt on the conclusions reached in these work.

Given these concerns, we feel we have no choice but to return this manuscript with the message that we cannot offer to publish it.

Nevertheless, the reviewers expressed interest in the subject matter, and they make constructive

comments which suggest that their concerns may be addressable with additional analyses and experimentation. Molecular Systems Biology now encourages reviewers to comment on each others reports, and during this process the second reviewer clarified his/her opinion writing, "The authors should thoroughly search the literature for high throughput approaches towards identification of a common inflammatory gene expression signature. I have suggested a couple of papers. The authors should clearly state how the signature they identified overlaps or differs from those identified in the earlier studies. Explanations for differences/overlaps should be provided. Thus, the "novelty" should be worked out." As such, we would like to indicate that we may be willing to reconsider a substantially revised work that addresses these concerns, and conclusively supports and clarifies the novel aspects of the inflammatory gene expression signature reported here, in addition to addressing the other reviewers' concerns. We recognize that this may involve further experimentation and analysis, and we can give no guarantee about its eventual acceptability.

Any resubmitted work would have a new number and receipt date. If you do decide to follow this course then it would be helpful to enclose with your re-submission an account of how the work has been altered in response to the points raised in the present review.

I am sorry that the review of your work did not result in a more favorable outcome on this occasion, but I hope that you will not be discouraged from sending your work to Molecular Systems Biology in the future.

Thank you for the opportunity to examine this work.

Sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Referee reports:

Reviewer #1 (Remarks to the Author):

The authors identified a set of inflamatome genes from twelve expression profiling data sets derived from nine different tissues of eleven rodent inflammatory models. Then they found the inflamatome genes are highly enriched in known drug targets and mouse and human coexpressed gene modules linked to metabolic disorders and cancer. Furthermore, by constructing inflamatome sub-networks from tissue specific Bayesian networks, the authors identified 166 key regulators which appeared to be more biologically important than the non-drivers in terms of the impact on mutant phenotypes. The analysis approach is generally sound and the results are of potential interest to many different fields. However several key analysis methods were not clearly described. The following need to be addressed before the paper is suitable for publication.

Major points:
1. Page 9. Paragraph 2: It is unclear what test is performed here to calculate the GO enrichment. Was Fisher's exact test performed here? Was multi-test adjustment performed to correct the P-values? The authors should make these clear.
2. Page 15. Paragraph 1: The authors claimed they constructed 66 BNs from 11 mouse crosses and 2 human BNs from the two previously described human studies. However, what are the differences between these 66 datasets? Why is there a need to generate all these BNs independently? How many data points were used to generate each BN? The authors should also demonstrate that these BNs were not over-fitted to each dataset.
3. Using edges appeared in >30% of the BNs to define the consensus network is technically too loose. In this definition, even BN structures do not contain directed loops, the consensus network may contain such loops. This is technically unconvincing and the heuristic for removing such loops were not justified. I suggest the authors to use more stringent criteria to define consensus networks (e.g. >60%) and examine whether the results changed.
Minor points:
1. Page 8. Paragraph 2: The authors should consider providing a brief summary on how many cases and controls were tested in each model and how many array data were analyzed in total.

2. Page 38. Paragraph 1, Figure 4 legend: The last sentence is not finished.

Reviewer #2 (Remarks to the Author):

The study by Wang et al. describes the identification of a set of genes that are up- or down-regulated in several disease models, referred to as "rodent inflammatory disease models". This group contains 2,493 genes where approximately 3/5 are up-regulated and 2/5 are down-regulated. These genes were designated as the inflammatome signature. This set of genes is enriched in known drug target genes. Using Bayesian network the authors then derived probability networks to identify genes that are likely to be causally connected with the diseased state. These driver genes were subsequently compared to known mutant phenotypes and a correlation was established. Additional correlation analysis was performed for GWAS-based human traits.

Main concerns:
1) Novelty
The concept of a common or representative inflammatory gene expression pattern, or inflammatory gene signature, is not new. Several studies established that inflammatory conditions (evoked by infection and sterile inflammation) are associated with the expression of a rather stereotypical set of genes. These studies include Jenner & Young (Nat Rev Microbiol. 2005 Apr;3(4):281-94) and Hao & Baltimore (Nat Immunol. 2009 Mar;10(3):281-8.). Although these studies did not propose such a high number of inflammasome-associated genes as the current manuscript there seems to be a great overlap. In addition, the large gene set identified in the current study is in fact not fully shared within the 12 disease models examined. Only 119 genes are consistent in all models (Table 2). While even this number is impressive it does not advances the concept introduced by the previous studies.
As for the cell-type specific inflammatory gene expression and the "key drivers" several other studies investigated expression patterns in various hematopoietic cell lineages including the master regulators of the regulatory circuits controlling the global expression changes: Gilchrist et al. (Nature. 2006 May 11;441(7090):173-8), Nilsson et al. (Genomics. 2006 Aug;88(2):133-42), FANTOM Consortium (Nat Genet. 2009 May;41(5):553-62), Litvak et al. (Nat Immunol. 2009 Apr;10(4):437-43). The authors of the current study may wish to discuss their results in the context of these earlier reports.

2) Concept
The authors attempted to identify genes/gene families that display dysregulated expression patterns in diseased tissues/cells. This group of genes was then regarded as representative for diseased conditions and proposed to contain drug target genes. The aim of this approach was to facilitate the development of drugs that would be beneficial in many inflammatory diseases rather than only in a limited spectrum of diseases. This appears somewhat inconsistent with the current efforts of both the pharma industry, medical doctors and public health authorities to develop individualized therapies that are tailored to the needs of individual patients. Thus, rather than looking for common signatures a search for genes differently expressed in each of the disease models appears useful. Alternatively, association of the representative expression patterns with the severeness of a particular disease appears important.

Other issues:
1) Introduction part
The authors use TNF as example of a potential general drug target since it commonly over-expressed under conditions of chronic inflammation but also in adipose tissues of obese humans. However, this example in fact shows that TNF may not be a proper target in these pathological conditions since TNF is also known to cause cachexia and loss of adipose tissue. Thus, TNF overexpression may be a rather a correlation than cause.

2) Page 10, last sentence of the paragraph before the section "Comparisons with current drug targets..."
Table 4 is referenced to contain down-regulated and their links to disease. Table 4 howeverdoes not contain such data.
3) Page 10, section "Comparisons with current drug targets and disease associated genes"
The authors compare the inflammatory gene signature with known drug targets and GWAS genes.

Subsequently, the presence (in percentage with p values) of drug targets and GWAS genes in the inflammatory signature is calculated ("The results indicate that 168 out of 803 drug target genes (22%, p<1.0e-35) and 346 out of 2883 GWAS genes (12.00%, p=1.07E-2) are included in the inflammatome signature, respectively". It is not clear how the p-values (that should indicate enrichment under assumption of "null hypothesis") were obtained since the parameters needed for the enrichment test are not provided.

4) Page 15, first sentence of the last paragraph
Supplement Table 3 does not contain the information as referred to.

5) Page 18, last paragraph before Discussion part
Supplementary Information Table 6 is missing.

6) Typos throughout the manuscript e.g. Agilent is often misspelled.

Reviewer #3 (Remarks to the Author):

1)Throughout the manuscript: It would be very helpful for reviewers if the manuscript had page numbers.

2)Figure 2: Although the authors make considerable efforts both in the text and legend to explain this figure, it is still very difficult to understand. For example, the legend states, "The symmetric heat map with rows and columns as genes represents the network connection strength (indicated by the blackness, where black means two nodes are tightly connected) between any pair of nodes (genes) in the network." To this reviewer's eyes, there is no blackness in the matrix itself only in the column to the right (mouse male adipose, mouse mail liver, mouse male muscle, mouse female adipose, mouse female liver, mouse female muscle, human breast cancer, human normal liver, human cancer liver) or in the column to the right and the row at the bottom (human male adipose, human female adipose). Likewise, the black box in each panel is not amenable to ready interpretation. One would think that a given module would always contain the inflammatome signature and this appears to be true in some cases, especially for the turquoise color block (mouse female muscle, human male adipose, human female adipose, human liver cancer), but not for the other tissues.

3)In the results for Table 5: It is stated "On the other hand, the purple module in NKI network and the black modules in the HCC network, enriched for cell cycle genes, are also significantly enriched for the inflammatome signature with Fisher's exact test p-values < 2.6e-37 and <1.6e-31 respectively." In Table 5 itself there is no "purple" module listed anywhere and the only "black" module listed is for NKI not HCC. Two lines later the non-listed "purple" and "black" modules in the NKI network are again evoked in terms of patient survival time.

4)Methods: In the mouse model for asthma, it is stated "mice were dosed". Please clarify. In the mouse model for fibrosis, both days 2 and 14 are mentioned, but the day of expression profiling is not given. In the mouse model for atherosclerosis, please indicate the HFD stands for high fat diet. Also, in this same model weeks 0, 8 and 16 are mentioned, but the week of expression profiling is not given. In the mouse model of diabetes, please define eWAT. In the mouse model for obesity (islet), weeks 6 and 9 are mentioned, but the day of expression profiling is not given. In the rat model for stroke, hours 4, 8, and 24 are mentioned, but the hour of expression profiling is not given. In the rat model for neuropathic pain, 4 time points are mentioned, but the time of expression profiling is not given. In the rat model for inflammation pain, please define mpk. Finally, there is no indication as to how RNA was prepared from the various tissues and how labeling was done prior to hybridization. Since the 12 datasets were presumably generated at different times and methodologies change over time, this could be a significant issue, especially with the newer amplification methods now in wide use compared to the older methods of not too many years ago.

We are pleased to submit our revised manuscript, *"Systems Biology Analysis of Eleven Rodent Disease Models Reveals An Inflammatome Signature and Its Key Regulators"* for consideration as a research article in *Molecular Systems Biology* under the Subject Category of *Molecular Biology of Disease*. We are sorry for the delay of the submission of our revision to this manuscript originally reviewed in October 2011 since several key authors of the paper experienced career changes and thus a prolonged relocation in the past several months.

We have now carried out an extensive revision of our original manuscript, given the reviewers' thoughtful and constructive review. We have addressed all of the major and minor issues raised by the reviewers. In particular, we have systematically collected seventeen inflammatory response gene and transcription factor signatures from a number of previous publications and then compared them with our inflammatome signature and predicted key drivers. The results show that

1) A vast majority of these seventeen inflammatory response gene signatures are not only significantly enriched in our inflammmatome signature and driver list;

2) A vast majority of these tested signatures are more likely in the inferred driver list than the whole inflammatome signature;

3) Among all the enrichment tests, the driver list has the highest likelihood to harbor two groups of transcription factors that are regulated in the early phase of temporal activation of macrophages and are likely to control gene expression in the intermediate and late phases, suggesting the higher regulatory power of the inferred drivers as oppose to the non-drivers in the inflammatome signature.

We have included point-by-point responses to all reviewer comments. If there is any additional information we can provide to facilitate your review of this manuscript, please let us know.

We look forward to hearing from you.

The authors appreciate the thoughtful and constructive comments from the three reviewers pertaining to our manuscript *"Systems Biology Analysis of Eleven Rodent Disease Models Reveals an Inflammatome Signature and Its Key Regulators"*. We have addressed all reviewer comments and revised the manuscript accordingly. The reviewer comments have served to significantly strengthen the manuscript.

Responses to all of the reviewer comments are provided below with our responses given in blue. All page numbers and other such references given are with respect to the revised manuscript unless otherwise stated.

Reviewer #1 (Remarks to the Author):

The authors identified a set of inflamatome genes from twelve expression profiling data sets derived from nine different tissues of eleven rodent inflammatory models. Then they found the inflamatome genes are highly enriched in known drug targets and mouse and human coexpressed gene modules linked to metabolic disorders and cancer. Furthermore, by constructing inflamatome sub-networks from tissue specific Bayesian networks, the authors identified 166 key regulators which appeared to be more biologically important than the non-drivers in terms of the impact on mutant phenotypes. The analysis approach is generally sound and the results are of potential interest to many different fields. However several key analysis methods were not clearly described. The following need to be addressed before the paper is suitable for publication.

Major points:

1. Page 9. Paragraph 2: It is unclear what test is performed here to calculate the GO enrichment. Was Fisher's exact test performed here? Was multi-test adjustment performed to correct the P-values? The authors should make these clear.

Response: We clarified the statistical test and multiple testing adjustment involved on Page 9 (Lines 12-14):

We then queried the biological functions associated with the up- and down-regulated inflammatome signature by using GO biological process enrichment analysis based on the Fisher Exact Test (P values were Bonferroni-corrected for multiple testing)

2. Page 15. Paragraph 1: The authors claimed they constructed 66 BNs from 11 mouse crosses and 2 human BNs from the two previously described human studies. However, what are the differences between these 66 datasets? Why is there a need to generate all these BNs independently? How many data points were used to generate each BN? The authors should also demonstrate that these BNs were not over-fitted to each dataset.

Response: The reviewer brought up an important question about the mouse and human studies. On Page 15 (Lines 8-22), we provided more detailed information about these studies:

The 66 BNs include 25 for liver, 23 for adipose and 18 for muscle from mouse and human studies. There are between 100 and 1000 samples in each data set used for network reconstruction. All Mouse studies were in genetic F2 settings. As different F2 crosses have different genetic architectures, so that there are different sets of causal reactive relationships. Thus, we constructed a network for each data set independently. Each BN

was gender and tissue-specific and was constructed using the genetic and gene expression data generated from each population (Zhu, Lum et al. 2004; Zhu, Wiener et al. 2007; Zhu, Zhang et al. 2008). Our previous studies showed that predictive BN can be constructed based on genetic and gene expression data with over 100 samples (Zhu, Lum et al. 2004; Zhu, Wiener et al. 2007; Zhu, Zhang et al. 2008). As construction of BN is NP-hard and we used a MCMC method to construct a network structure. We generally simultaneously run one thousand MCMC chains and result 1000 networks (Friedman, Linial et al. 2000; Zhu, Wiener et al. 2007). Our final network is the common features among these 1000 structures. Over-fitting the data is not an issue here. At a final step, we derived a tissue-specific consensus BN as the union of all the BNs for each of these three tissues.

3. Using edges appeared in >30% of the BNs to define the consensus network is technically too loose. In this definition, even BN structures do not contain directed loops, the consensus network may contain such loops. This is technically unconvincing and the heuristic for removing such loops were not justified. I suggest the authors to use more stringent criteria to define consensus networks (e.g. >60%) and examine whether the results changed.

Response: Previously we conducted a comprehensive simulation study (Zhu, Wiener et al. 2007) where 30% cutoff was determined to be optimal. We addressed this concern in detail on Page 34 (Lines 16-22):

The 30% cutoff threshold for edge inclusion is based on our previous simulation study (Zhu, Wiener et al. 2007), where 30% cutoff yields the best tradeoff between recall rate and precision. The consensus network resulted by averaging may not be a Bayesian network (a directed acyclic graph) any more. To make the consensus network structure into a directed acyclic graph, edges in this consensus network were removed if and only if 1) the edge was involved in a loop, and 2) the edge was the most weakly supported of all edges making up the loop.

Minor points:

1. Page 8. Paragraph 2: The authors should consider providing a brief summary on how many cases and controls were tested in each model and how many array data were analyzed in total.

Response: We have summarized the detailed information as suggested by the reviewer into Table 1.

Table 1. *12 rodent inflammatory disease models and the number of cases and controls used in the current analysis.*

Disease	Model	Species	Tissue profiled	# of Cases	# of Controls	# of Total Arrays
Asthma	OVA	Mouse	Lung	5	4	9
COPD	IL-1b Tg	Mouse	Lung	5	3	8
Fibrosis	TGFb Tg	Mouse	Lung	4	4	8
Atherosclerosis	ApoE KO HFD	Mouse	Aorta	3	3	6
Diabetes	db/db	Mouse	Adipose	3	3	6
Diabetes	db/db	Mouse	Islet	5	5	10
Obesity	ob/ob	Mouse	Adipose	3	3	6
Multiple	LPS	Rat	Liver	4	4	8
Stroke	MCAO	Rat	Brain	4	4	8
Neuropathic pain	Chung	Rat	DRG	4	4	8
Inflammation pain	CGN	Rat	Skin	4	5	9
Sarcopenia	Aged vs. Young	Rat	Muscle	5	5	10

2. Page 38. Paragraph 1, Figure 4 legend: The last sentence is not finished.
Response: We apologize for the oversight. This has been corrected as follows (Page 40, Lines 3-4):
The red nodes highlighted are the predicted driver genes.

Reviewer #2 (Remarks to the Author):
The study by Wang et al. describes the identification of a set of genes that are up- or down-regulated in several disease models, referred to as "rodent inflammatory disease models". This group contains 2,493 genes where approximately 3/5 are up-regulated and 2/5 are down-regulated. These genes were designated as the inflammatome signature. This set of genes is enriched in known drug target genes. Using Bayesian network the authors then derived probability networks to identify genes that are likely to be causally connected with the diseased state. These driver genes were subsequently compared to known mutant phenotypes and a correlation was established. Additional correlation analysis was performed for GWAS-based human traits.
Main concerns:
1) Novelty
The concept of a common or representative inflammatory gene expression pattern, or inflammatory gene signature, is not new. Several studies established that inflammatory conditions (evoked by infection and sterile inflammation) are associated with the expression of a rather stereotypical set of genes. These studies include Jenner & Young (Nat Rev Microbiol. 2005 Apr;3(4):281-94) and Hao & Baltimore (Nat Immunol. 2009 Mar;10(3):281-8.). Although these studies did not propose such a high number of inflammasome-associated genes as the current manuscript there seems to be a great overlap. In addition, the large gene set identified in the current study is in fact not fully shared within the 12 disease models examined. Only 119 genes are consistent in all models (Table 2). While even this number is impressive it does not advances the concept introduced by the previous studies.
Response: Although we agree with the reviewer that the concept of identifying consensus inflammatome gene signatures is not new, the study we carried out differ significantly from the previous studies and went far beyond the identification of a common set of genes. First of all, the disease models investigated by our study cover far

more diverse disease types (12 diseases) than the pathogen- or TNF-centric inflammatory conditions described in the earlier studies. The gene signatures identified, therefore, are more representative of a consensus signature of inflammatory diseases. Second, the systems biology approaches taken are far more integrative than ever considered or achieved by any of the previous studies. The integration of gene expression data with network analysis helped identify not only the underlying multi-tissue network architecture of inflammotome but also the key driver genes. The obtained network structure offers a better understanding of the interactions among the signature genes and the differentiation between driver genes and non-driver genes in the inflammotome signature yields putative causal genes that are more critical for disease pathogenesis. Indeed, many key drivers identified in the current manuscript have been validated in the available literature. None of these insights have been obtained from the previous studies. We emphasize this in the section of Discussion (Page 21, Lines 20-23; Page 22, Lines 1-6):

To our knowledge, this is the first study that systematically investigates multiple (9) tissues of multiple (12) disease models with inflammatory component. The integration of gene expression profiling data with knowledge-based databases and data-driven networks not only help identify a common inflammatome signature and relate the signature to the diseases under investigation, but also help uncover the network architecture and key genes that drive the inflammatome signature. Since the inflammatome was derived from multiple disease models and tissues, it points to a central role that infiltrating inflammatory cells such as macrophages play in all major disease areas. Several genes of macrophage origin, when perturbed, have been shown to impact multiple disease outcomes. It is conceivable that further mining and validation of the inflammatome signature, especially the key drivers identified, could result in additional high value targets.

As for the cell-type specific inflammatory gene expression and the "key drivers" several other studies investigated expression patterns in various hematopoietic cell lineages including the master regulators of the regulatory circuits controlling the global expression changes: Gilchrist et al. (Nature. 2006 May 11;441(7090):173-8), Nilsson et al. (Genomics. 2006 Aug;88(2):133-42), FANTOM Consortium (Nat Genet. 2009 May;41(5):553-62), Litvak et al. (Nat Immunol. 2009 Apr;10(4):437-43). The authors of the current study may wish to discuss their results in the context of these earlier reports.

Response: We systematically collected the genes and the transcription factors from six previous publications and then compared them with our inflammatome signature and key drivers. The results were summarized in Table 7 (Page 49) and incorporated in the Results section as follows (Page 19, Lines 14-23; Page 20 Lines 1-7):

Previous studies investigated expression patterns in various hematopoietic cell lineages and their potential regulatory transcription factors (Jenner and Young 2005; Gilchrist, Thorsson et al. 2006; Nilsson, Bajic et al. 2006; Hao and Baltimore 2009; Litvak, Ramsey et al. 2009; Suzuki, Forrest et al. 2009). From these publications, we collected 17 inflammatory response gene signatures. As shown in Table 7, our inflammatome signature and its driver list significantly overlap with 11 and 13 of the 17 signatures. Among the 15 signatures that show different

enrichment (the difference between fold enrichments is greater 0.5) in the inflammatome signature and the driver list, 12 are more likely in the inferred driver list than the whole inflammatome signature. Among all the enrichment tests, the driver list has the highest likelihood to harbor two groups of macrophage induced transcription factors, Cluster 6 and LPS-TF-Cluster1, with fold enrichment as 15.4 and 19.8, respectively. Both groups are comprised of many transcription factors regulated in the early phase of temporal activation of macrophages and are likely to control gene expression in the intermediate and late phases (Gilchrist, Thorsson et al. 2006; Litvak, Ramsey et al. 2009), indicating the regulatory power of the inferred drivers as oppose to the non-drivers in the inflammatome signature. Although there is an over-representation of these tested regulatory TFs in our driver list, a majority of the candidate drivers are new, suggesting that our approach is complementary to the previous studies.

Table 7. Enrichment test of our inflammatome signature and its drivers for the inflammatory signatures from the literature.

			FET p value		Overlap		Fold Enrichment	
signature	source	size	Drivers	inflammatome	Drivers (166)	inflammatome (2496)	Drivers	inflammatome
Cluster1	Gilchrist et al. 2006	137	1	0.717	0	12	0	0.818609728
Cluster10	Gilchrist et al. 2006	215	5.64E-08	1.22E-08	11	51	7.08882	2.216909368
Cluster11	Gilchrist et al. 2006	61	1	0.000521	0	15	0	2.298146162
Cluster2	Gilchrist et al. 2006	167	0.00137	0.00000515	5	37	4.14833	2.070625105
Cluster3	Gilchrist et al. 2006	64	0.0781	0.00092	1	15	2.16491	2.190420561
Cluster4	Gilchrist et al. 2006	140	8.35E-07	0.000000138	8	36	7.917384	2.403204272
Cluster5	Gilchrist et al. 2006	42	0.0369	0.00000307	1	15	3.29891	3.337783712
Cluster6	Gilchrist et al. 2006	18	0.000278	0.119	2	3	15.39491	1.557632399
Cluster7	Gilchrist et al. 2006	178	0.00189	0.0627	5	25	3.891972	1.312611572
Cluster8	Gilchrist et al. 2006	146	7.15E-10	9.98E-11	11	43	10.43902	2.752528485
Cluster9	Gilchrist et al. 2006	36	0.0278	0.0838	1	6	3.848728	1.557632399
FANTON-TF	FANTON, 2008	47	0.0453	0.00321	1	11	2.947962	2.187313581
HostResponse	Jenner & Young 2005	511	1.41E-10	1.77E-24	20	135	5.422866	2.469045485
LPS-TF-cluster1	Litvak et al. 2009	21	0.0000142	0.00477	3	6	19.79346	2.670226969
LPS-TF-cluster2	Litvak et al. 2009	57	1	0.151	0	8	0	1.311690441
macrophage-regulated	Nilson et al. 2006	1552	2.14E-13	1.81E-23	40	292	3.570985	1.758358223
TNF-signature	Hao & Baltimore 2009	89	0.0000457	1.26E-15	5	38	7.783945	3.990339179

2) Concept

The authors attempted to identify genes/gene families that display dysregulated expression patterns in diseased tissues/cells. This group of genes was then regarded as representative for diseased conditions and proposed to contain drug target genes. The aim of this approach was to facilitate the development of drugs that would be beneficial in many inflammatory diseases rather than only in a limited spectrum of diseases. This appears somewhat inconsistent with the current efforts of both the pharma industry, medical doctors and public health authorities to develop individualized therapies that are tailored to the needs of individual patients. Thus, rather than looking for common signatures a search for genes differently expressed in each of the disease models appears useful. Alternatively, association of the representative expression patterns with the severeness of a particular disease appears important.

Response: We agree with the reviewer that disease-specific signatures can be of significant value for personalized medicine, as commonly practiced in today's pharmaceutical industry and medical system. However, such approach may in part reflect the lack of understanding of the relationships among diseases, especially the commonalities among diseases. We envision that in light of a comprehensive understanding of the shared molecular mechanisms, an opposite approach in which common features across diseases are of focus can be equally valuable. As such approach allows the development of drugs that can be used for targeting multiple diseases, the current study could potentially help point the drug discovery process to a complementary direction. The most effective and personalized treatment of a particular disease could be the combination of a drug targeting the core processes common to multiple diseases and a drug that is more specific to each disease type.
In order to address the reviewer's concern over the missing aspect of signatures for individual diseases, we have supplied Supplementary Table S6 to show the membership of each signature gene in each disease model. On average ~70 genes show differential expression in only one disease model but not the others, supporting disease-specific alternations in gene expression. However, this number is relatively small compared to the common inflammatome identified.

We further clarify this point on Pages 22 (Lines 8-23) and 23 (Lines 1-8):

These results will facilitate the development of drugs targeting many inflammatory diseases rather than only in a limited spectrum of diseases. However, the common inflammatome genes and their drivers don't deny the significant value of disease-specific signatures for personalized medicine, commonly practiced in today's pharmaceutical industry and medical system. Since such personalized medicine approach may in part reflect the lack of understanding of the relationships among diseases, especially the commonalities among diseases, we envision that in light of a comprehensive understanding of the shared molecular mechanisms, an opposite approach in which common features across diseases are of focus can be equally valuable. As such approach allows the development of drugs that can be used for targeting multiple diseases, the current study could potentially help point the drug discovery process to a complementary direction. The most effective and personalized treatment of a particular disease could be the combination of a drug targeting the core processes common to multiple diseases and a drug that are more specific to each disease type. Supplementary Table S6 shows the membership of each signature gene in each disease model. On average ~70 genes show differential expression in only one disease model but not the others, supporting disease-specific alternations in gene expression. However, this number is relatively small compared to the common inflammatome we identified.

The approach we used to come up with the inflammatome can be modified to derive disease model-specific gene sets with the potential of identifying disease-specific biomarkers and mechanisms. For instance, when we included human osteoarthritis cartilage tissue with 6 mouse disease models (OVA/asthma, IL-1β-Tg/emphysema, TGFβ-Tg/lung fibrosis, ApoE KO/atherosclerosis, db/db T2DM and ob/ob obesity) for an integrated analysis, we were able to identify OA-specific disease markers such as asporin

(Kizawa, Kou et al. 2005) and matrilin 3 (van der Weyden, Wei et al. 2006; Vincourt, Vignaud et al. 2008). Such analysis can certainly be extended to other disease models in the future.

Other issues:
1) Introduction Part
The authors use TNF as example of a potential general drug target since it commonly over-expressed under conditions of chronic inflammation but also in adipose tissues of obese humans. However, this example in fact shows that TNF may not be a proper target in these pathological conditions since TNF is also known to cause cachexia and loss of adipose tissue. Thus, TNF overexpression may be a rather a correlation than cause.

Response: We used TNF as an example to associate multiple diseases with inflammation. Anti-TNF therapeutics has been shown to be efficacious in several autoimmune diseases including rheumatoid arthritis, Crohn's disease and psoriasis (Van Hauwermeiren et al. Cytokine Growth Factor Rev. 22:311, [2011]) but we are not claiming that it could also be applied to other inflammatory diseases as a therapeutic target. However, it is worth mentioning that inflammatory stress caused by TNF or IL-6 has a similar effect as energy stress caused by over-nutrition or inactivity; both target stress-induced serine kinases in liver and result in insulin resistance (Muoio et al. Nat. Rev. Mol. Cell Bio. 9:193 [2008]). This line of evidence points to a causal role of TNF in obesity onset rather than merely a correlation or a reactive response.

2) Page 10, last sentence of the paragraph before the section "Comparisons with current drug targets..." Table 4 is referenced to contain down-regulated and their links to disease. Table 4 however does not contain such data.
Response: We apologize for the typo. It should refer to Table 3 instead. We have corrected this on page 10 (Lines 6-8):

*The down-regulated genes are significantly associated with nerve impulse transmission, energy generation and all major metabolic processes involving amino acid, fatty acid and carbohydrate **(Table 3)**.*

3) Page 10, section "Comparisons with current drug targets and disease associated genes" The authors compare the inflammatory gene signature with known drug targets and GWAS genes. Subsequently, the presence (in percentage with p values) of drug targets and GWAS genes in the inflammatory signature is calculated ("The results indicate that 168 out of 803 drug target genes (22%, p=1.28e-21) and 346 out of 2883 GWAS genes (12.00%, p=5.00e-5) are included in the inflammatome signature, respectively". It is not clear how the p-values (that should indicate enrichment under assumption of "null hypothesis") were obtained since the parameters needed for the enrichment test are not provided.
Response: The null hypothesis is that our inflammatome signature is a random sampling of genes and is not enriched for the drug targets or GWAS genes. To clarify the details of the statistic tests, we have revised the paragraph as follows on page 10, lines 15-21:

The results indicate that 168 out of 803 drug target genes and 346 out of 2883 GWAS genes are included in the inflammatome signature ***(Supplementary Information Table 2),*** *representing significant enrichment of inflammatome genes in these gene sets with enrichment p values 1.28e-21 and 5.00e-5 as assessed by Fisher's exact test, respectively (compared to 2483 inflammatome genes among 25k common genes across the arrays under the null hypothesis that our inflammatome signature is a random combination and is not enriched for the drug targets or GWAS genes.)*

4) Page 15, first sentence of the last paragraph
Supplement Table 3 does not contain the information as referred to.
Response: This has been corrected in the revised manuscript.

5) Page 18, last paragraph before Discussion part. Supplementary Information Table 6 is missing.
Response: This was a typo. It should refer to Supplementary Information Table 5. We have made a correction.

6) Typos throughout the manuscript e.g. Agilent is often misspelled.
Response: We have gone through careful edits and corrected all typos we could identify.

Reviewer #3 (Remarks to the Author):
1)Throughout the manuscript: It would be very helpful for reviewers if the manuscript had page numbers.
Response: All the pages and lines have been numbered in the current version.

2)Figure 2: Although the authors make considerable efforts both in the text and legend to explain this figure, it is still very difficult to understand. For example, the legend states, "The symmetric heat map with rows and columns as genes represents the network connection strength (indicated by the blackness, where black means two nodes are tightly connected) between any pair of nodes (genes) in the network." To this reviewer's eyes, there is no blackness in the matrix itself only in the column to the right (mouse male adipose, mouse mail liver, mouse male muscle, mouse female adipose, mouse female liver, mouse female muscle, human breast cancer, human normal liver, human cancer liver) or in the column to the right and the row at the bottom (human male adipose, human female adipose). Likewise, the black box in each panel is not amenable to ready interpretation. One would think that a given module would always contain the inflammatome signature and this appears to be true in some cases, especially for the turquoise color block (mouse female muscle, human male adipose, human female adipose, human liver cancer), but not for the other tissues.
Response: The black squares were corrupted in the process of copy-and-paste. We apologize for this and have corrected the figures and rewritten the legend as follows (Page 39 Lines 9-18):

Figure 2. *Topological overlap matrix (TOM) plots of weighted, gene coexpression networks constructed from several studies discussed in the text. The symmetric heat map with rows and columns as genes represents the network connection strength (as indicated*

by the different shades of color- from white signifying not significantly correlated to red signifying highly significantly correlated) between any pair of nodes (genes) in the network. The network connection strength is measured as the topological overlap between genes. The network modules highlighted as color block along the rows and columns (each color block represents a module) were identified via an average linkage hierarchical clustering algorithm using topological overlap as the dissimilarity metric. In each network, the module highlighted with a black box is most enriched with the inflammatome signature.

3)In the results for Table 5: It is stated "On the other hand, the purple module in NKI network and the black modules in the HCC network, enriched for cell cycle genes, are also significantly enriched for the inflammatome signature with Fisher's exact test p-values < 2.6e-37 and <1.6e-31 respectively." In Table 5 itself there is no "purple" module listed anywhere and the only "black" module listed is for NKI not HCC. Two lines later the non-listed "purple" and "black" modules in the NKI network are again evoked in terms of patient survival time.

Response: We are sorry for the confusion since two modules in the NKI network are highly enriched for the inflammatome signature but the purple module information was missing in Table 5 and the module name for HCC was supposed to be turquoise. We have added the purple module into Table 5 and made a correction in the main text as follows (Page 12, Line 23; Page 13 Lines 1-11):

About one third of the black module in NKI breast cancer network overlap with the inflammatome signature and this represents a 3 fold enrichment ($P<2e^{-51}$). All these modules are significantly enriched for genes in inflammatory and immune response pathways. On the other hand, the purple module in NKI network and the turquoise modules in the HCC network, are also significantly enriched for the inflammatome signature with Fisher's exact test p-values < $2.6e^{-37}$ and <$1.6e^{-31}$ respectively. Both modules include many typical cell cycle genes such as TOP2A, CHEK1, E2F1 and EZH2, just name a few. Interestingly, the purple module the NKI network, is the most predictive of the patient's survival time (Cox model p-value < e^{-12}) while the black module is moderately predictive of survival time as well. Thus these results show that the inflammatome signature is highly conserved not only in multiple tissues (liver, adipose, muscle and breast) but also in multiple diseases (cancer, obesity and diabetes) in datasets independent of those from which the signature was derived.

Table 5. *Network gene modules most enriched with the inflammatome signature*

tissue	gender	module	network size	signature size	signature in network	module size	signature in module	fold enrichment	Enrichment p-value
Adipose	Female	blue	21936	2505	2258	1991	672	3.2789252	1.13E-203
Adipose	Male	brown	21936	2505	2258	1604	597	3.6157922	2.58E-203
Muscle	Male	blue	21836	2505	2249	2803	721	2.4974442	1.16E-143
Liver	Female	red	21936	2505	2258	605	291	4.6727346	2.81E-129
Liver	Male	yellow	21936	2505	2258	1206	395	3.1818763	1.29E-108
Muscle	Female	turquoise	21842	2505	2250	4518	858	1.8435331	4.40E-91
Breast (cancer)	All	purple	19570	2276	1995	387	130	3.2951888	2.56E-37
Breast (cancer)	All	black	19570	2276	1995	644	201	3.0616681	1.97E-51
Liver (cancer)	All	turquoise	14878	2276	1835	2405	510	1.7193472	2.47E-42
Adipose	Male	turquoise	5580	2276	824	1123	316	1.90552352	1.50E-40
Adipose	Female	turquoise	5561	2276	842	1696	411	1.6005041	2.10E-34
Liver	All	yellow	4408	2276	623	180	84	3.3018727	7.63E-28

4)Methods: In the mouse model for asthma, it is stated "mice were dosed". Please clarify. In the mouse model for fibrosis, both days 2 and 14 are mentioned, but the day of expression profiling is not given. In the mouse model for atherosclerosis, please indicate the HFD stands for high fat diet. Also, in this same model weeks 0, 8 and 16 are mentioned, but the week of expression profiling is not given. In the mouse model of diabetes, please define eWAT. In the mouse model for obesity (islet), weeks 6 and 9 are mentioned, but the day of expression profiling is not given. In the rat model for stroke, hours 4, 8, and 24 are mentioned, but the hour of expression profiling is not given. In the rat model for neuropathic pain, 4 time points are mentioned, but the time of expression profiling is not given. In the rat model for inflammation pain, please define mpk. Finally, there is no indication as to how RNA was prepared from the various tissues and how labeling was done prior to hybridization. Since the 12 datasets were presumably generated at different times and methodologies change over time, this could be a significant issue, especially with the newer amplification methods now in wide use compared to the older methods of not too many years ago.

Response: We thank the reviewers for pointing out the missing information. Methods for the 12 sample sets assessed in this manuscript have been revised to clarify the time point issues raised by the reviewer. Original studies from which these samples were derived contained multiple time points and most were also treated by compounds. Since our objective in this study was to identify a representative inflammatory signature, we took only a subset of samples at one fixed time point (when inflammation is eminent) without any compound treatment from each individual model. We made the corrections in the "Methods" to include only the info relevant to the current analyses. We also explained the acronyms used and provided details about the RNA purification and microarray protocols.

2nd Editorial Decision 20 April 2012

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the two referees who agreed to evaluate this resubmitted study. As you will see, the opinion of the referees was somewhat split, with one reviewer fully supportive of publication and the other still concerned about the depth of the novel biological insights. After consultation with an additional expert, we have decided to request a final revision of this work to address some key remaining issues, and to improve the suitability of this work for publication at Molecular Systems Biology.

Overall, all the reviewers appreciated the potential value of the underlying datasets, with the additional expert writing "I agree with all the referees that the extensive data set will be valuable to the field and if Merck has not deposited these data due to prior publications in particular areas, then the aggregated data will be quite useful to have in the public domain." Given that this was an essential point for the reviewers, and in line with Molecular Systems Biology policies, any revised work will need to fully release the underlying data for *all* unpublished datasets used in this work. I note that the data for the C3ar1-/- experiments, and the Mouse BXH Cross are already present in GEO (GSE11995 and GSE2008), but the expression profiles for the 11 rodent disease models, and the Alox5 knockout apparently still need to be deposited. Please contact us if you have any questions regarding our data release policies.

The expert adviser also felt that some additional analysis could increase the depth of the insights derived from these data, writing: "Given the diversity of data sets available, they could have done not just the binary k means clustering based on inflammation vs. control, but also created finer gene sets that might have given overlapping yet distinct signatures for inflammatory genes in distinct disease settings, thus helping the reader to understand what is shared by these processes and what is unique, because the unique aspects may be an or more important for choosing targets as the common features. See the Chaussabel work in mention below" [Pankla et al, 2009]. The expert also noted the works by Banchereau, Chaussabel and colleagues as potentially deserving of mention and citation.

Lastly, the expert adviser felt that the manuscript "needs substantial editorial work to clean up poorly constructed and often confusing sentences. I also find the main figures of limited value - some zoomed in maps of key drivers and their (proximate) downstream targets would be useful as an adjunct to the tables, for example."

The editor feels that collectively addressing these three points (full data release, some additional analysis of disease-specific inflammatory signatures, and a thorough revision for writing and figure clarity) would help to support the value of this work, and its appropriateness for publication in Molecular Systems Biology.

Thank you for submitting this paper to Molecular Systems Biology.

Sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Referee reports:

Reviewer #2 (Remarks to the Author):

The re-submitted study puts forward the idea that that it has identified a gene expression signature which is common to inflammatory diseases regardless of tissue. This signature is called "inflammatome".

Throughout the study multiple statistical and probability analyses are conducted with the aim to prove that the identified signature correlates with already known inflammation-associated expression data and/or genes linked (based on known animal models, disease association studies or SNP analyses) to inflammatory diseases.

Altogether, the study represents a thorough analysis which however does not advance the already well established knowledge about the existence of a common inflammation-associated gene expression pattern. Although the study is unique with respect of the large dataset analyzed and, consequently, the identification of new inflammation-associated genes (be it key drivers or bystanders), altogether the findings represent only incremental advance over previous information. This is best documented by the fact that 11 out of 17 signatures identified in previously published reports significantly overlap with the total inflammation signature in this study, and 13 out of 17 known signatures significantly overlap with the drive signature in this study (page 19, lines 17 - 19).

As for the association with drug targets, it is difficult to imagine how the data will be used for drug design and development: e.g. VEGF is found here as member of the inflammation signature, yet the role of VEGF in diseases such as atherosclerosis or diabetes is controversial with both beneficial and harmful effects described in numerous publications.

Reviewer #3 (Remarks to the Author):

The authors have adequately responded to this referee's comments and made appropriate revisions to the manuscript in accordance with those comments.

1st Revision - authors' response 22 May 2012

We are pleased to submit our revised manuscript, *"A Systems Biology Analysis of Eleven Rodent Disease Models Reveals An Inflammatome Signature and Its Key Regulators"* for consideration as a research article in *Molecular Systems Biology* under the Subject Category of *Molecular Biology of Disease*.

We have now carried out a comprehensive revision of our previous submission, given the thoughtful and constructive review by you and the reviewers. We have addressed all the comments from the 2nd round review. In particular, we have made following major additions:

1) The revised manuscript has gone through substantial editing from two native speakers. The major network figure (Figure 4) was replaced with two high-resolution subnetworks highlighting the key drivers and their local subnetworks. The node and edge information of the overall inflammatome causal networks is provided as Supplementary Table 7 to allow readers to examine details of the network structures.
2) We have systematically updated the network analysis and all the comparisons using the latest key driver analysis package and knowledge databases (MGI and GWAS). An improved performance has been observed.
3) We have added a section (Page 10 Line 12 to Page 11 Lines 9) to describe our analysis and results on disease-specific genes. The detailed information about the disease-specific signatures is provided as two supplementary tables.
4) We have added the inflammatory signature identified from septicemic melioidosis by Pankla et al, 2009 into the literature-based inflammation signature set to evaluate the cross-study consistency of our inflammatome signature and its driver gene sets.
5) All the mouse model datasets as well as the Alox5 knockout data have been uploaded into GEO under accession no. GSE31559.

We have included point-by-point responses to all the comments. If there is any additional information we can provide to facilitate your review of this manuscript, please let us know.

We look forward to hearing from you.

Response to MSB-12-3648 Decision Letter

The authors appreciate the thoughtful and constructive comments from the editor and reviewers pertaining to our manuscript *"A Systems Biology Analysis of Eleven Rodent Disease Models Reveals an Inflammatome Signature and Its Key Regulators"*. We have addressed all comments and revised the manuscript accordingly.

Responses to all of the reviewer/editor comments are provided below with our responses given in blue. All page numbers and other such references given are with respect to the revised manuscript unless otherwise stated.

MSB Editor:

Overall, all the reviewers appreciated the potential value of the underlying datasets, with the additional expert writing "I agree with all the referees that the extensive data set will be valuable to the field and if Merck has not deposited these data due to prior publications in particular areas, then the aggregated data will be quite useful to have in the public domain." Given that this was an essential point for the reviewers, and in line with Molecular Systems Biology policies, any revised work will need to fully release the underlying data for *all* unpublished datasets used in this work. I note that the data for the C3ar1-/- experiments, and the Mouse BXH Cross are already present in GEO (GSE11995 and GSE2008), but the expression profiles for the 11 rodent disease models, and the Alox5 knockout apparently still need to be deposited. Please contact us if you have any questions regarding our data release policies.

Response:
All the mouse model datasets as well as the Alox5 knockout data have been uploaded into GEO under accession no. GSE31559.

The expert adviser also felt that some additional analysis could increase the depth of the insights derived from these data, writing: "Given the diversity of data sets available, they could have done not just the binary k means clustering based on inflammation vs. control, but also created finer gene sets that might have given overlapping yet distinct signatures for inflammatory genes in distinct disease settings, thus helping the reader to understand what is shared by these processes and what is unique, because the unique aspects may be an or more important for choosing targets as the common features. See the Chaussabel work in mention below" [Pankla et al, 2009].

Response:
We agree with the adviser and the other reviewers that defining disease-specific signatures is a very important aspect. We have added a section (Page 10 Line 12 to Page 11 Lines 10) to describe our analysis and results on disease-specific genes. The detailed information about the disease-specific signatures can be found in the Supplementary Tables 2 and 3.

"**Identification of disease-specific signatures"**

We also identified disease-specific gene signatures, defined herein as genes that reach statistical significance at false discovery rate (FDR) <5% only in one of the rodent models between the disease group and the control group using one-way ANOVA. At FDR<5%, 1175, 26, 1120, 284, 177, 28, 782, 1123, 1208, 476, 292, and 222 genes were identified as disease model-specific signatures for Apoe KO (aorta), *db/db* adipose, *db/db* islet, Chung neuropathic pain (DRG), CGN-induced pain (skin), IL-1β Tg emphysema model (lung), LPS treated acute injury (liver), ob/ob (adipose), OVA-challenged asthma (lung), sarcopenia (muscle), MCAO stroke (brain), and pulmonary fibrosis model (lung), respectively (**Supplementary Table 2)**. On average 576 genes show disease-specific differential expression, supporting the presence of disease-specific alternations in gene expression. However, compared to the ~2500 inflammatome genes shared among the disease models as described above, the numbers of the disease-specific genes are much smaller. Moreover, a GO biological process enrichment analysis based on the Fisher Exact Test only revealed enrichment of a limited set of functional categories (such as sensory perception, ion homeostasis, and neuron development) for four of the disease model-specific signatures – ApoE KO, db/db islet, ob/ob, and MCAO stroke - at *Bonferroni*-corrected p<0.05 (**Supplementary Table 3**).

Although disease-specific signature genes can help identify disease-specific mechanisms, targets, and biomarkers, the smaller signature sizes limit our power to identify robust disease-specific signatures, as indicated by the poor coherence in the functionalities of the disease-specific signature genes in comparison to the common inflammatome signature. We therefore conducted a more in-depth analysis of the inflammatome signature in subsequent sections.

The expert also noted the works by Banchereau, Chaussabel and colleagues as potentially deserving of mention and citation.

Response:
We appreciate that the expert pointed out this valuable work. We have added the relevant references. In addition, we added the inflammatory signature identified from septicemic melioidosis by Pankla et al, 2009 into the literature-based inflammation signature set to evaluate the cross-study consistency of our inflammatome signature and its driver gene sets. As shown in Table 7, our inflammatome signature significantly overlaps with majority of the literature-based signatures, including the septicemic melioidosis signature from Pankla et al. Moreover, the key drivers have higher likelihood to overlap with the literature signatures than the non-drivers of the inflammatome signature (4.5 versus 2.9 fold enrichment in the case of the septicemic melioidosis signature). These results suggest that our signature is highly consistent with previous findings.

Lastly, the expert adviser felt that the manuscript "needs substantial editorial work to clean up poorly constructed and often confusing sentences. I also find the main figures of limited value - some zoomed in maps of key drivers and their (proximate) downstream targets would be useful as an adjunct to the tables, for example."
Response: The revised manuscript has gone through substantial editing from two native speakers. The major network figure (Figure 4) was replaced with two high-resolution subnetworks highlighting the key drivers and their local subnetworks. The node and edge information of the

overall inflammatome causal networks has now been provided as Supplementary Tables 8 and 9 to allow readers to examine details of the network structures.

The editor feels that collectively addressing these three points (full data release, some additional analysis of disease-specific inflammatory signatures, and a thorough revision for writing and figure clarity) would help to support the value of this work, and its appropriateness for publication in Molecular Systems Biology.

Response:
All three points have been fully addressed as detailed above.

Also, when preparing your revised work, please address the following format and content issues:

1. Please upload all of the supplementary tables as individual excel or tab-delimited text files, rather than as a single pdf.

Response:
All supplementary tables have been individually uploaded as excel files.

2. Please provide three to four 'bullet points' highlighting the main findings of your study.

Response:
Three bullet points highlighting the main findings of the study have been added to the main text on Page 4:

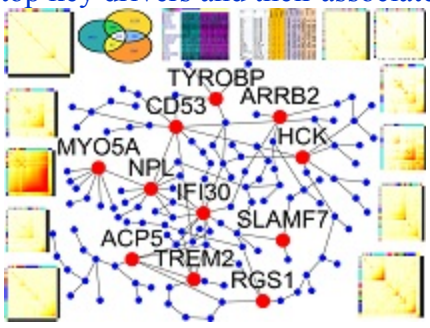
- Representative inflammatome gene signatures as well as eleven disease model-specific gene signatures were identified from twelve gene expression profiling data sets derived from nine different tissues isolated from eleven rodent inflammatory disease models.
- The inflammatome signature is highly enriched for immune response-related genes, disease causal genes, and drug targets.
- Regulatory relationships among the inflammatome signature genes were examined in over 70 causal networks derived from a number of large scale genetic studies of multiple diseases and the potential key drivers were uncovered and validated prospectively.
- Over 70% of the inflammatome signature genes and over 50% of the key driver genes have not been reported in previous studies of common signatures in inflammatory conditions.

3. Please provide a 'standfirst text' summarizing in two sentences the study (approx. 250 characters).

Response:
A standfirst text has been added to the main text on Page 4:
“Common inflammatome signatures as well as model-specific gene signatures were identified from 11 rodent inflammatory disease models. The causal regulatory networks and the drivers of the inflammatome signature were uncovered and retrospectively validated.”

4. Please provide a "thumbnail image" (width=211 x height=157 pixels, jpeg format), which can be used to highlight your paper on our homepage.

Response:
The following thumbnail image summarizes the key findings of the paper and highlights a few top key drivers and their associated subnetwork.



5. Please include an author contributions statement after the Acknowledgements section (see http://www.nature.com/msb/authors/index.html#Submission).

Response:
Author contributions and Acknowledgement sections have been added.

Reviewer #2 (Remarks to the Author):

The re-submitted study puts forward the idea that that it has identified a gene expression signature which is common to inflammatory diseases regardless of tissue. This signature is called "inflammatome".

Throughout the study multiple statistical and probability analyses are conducted with the aim to prove that the identified signature correlates with already known inflammation-associated expression data and/or genes linked (based on known animal models, disease association studies or SNP analyses) to inflammatory diseases.

Altogether, the study represents a thorough analysis which however does not advance the already well established knowledge about the existence of a common inflammation-associated gene expression pattern. Although the study is unique with respect of the large dataset analyzed and, consequently, the identification of new inflammation-associated genes (be it key drivers or bystanders), altogether the findings represent only incremental advance over previous information. This is best documented by the fact that 11 out of 17 signatures identified in previously published reports significantly overlap with the total inflammation signature in this study, and 13 out of 17 known signatures significantly overlap with the drive signature in this study (page 19, lines 17 - 19).

Response:

Although we agree with the reviewer that our study supports the previous notion that a common inflammatome signature exist among inflammatory diseases, the previous studies mainly focused on transcriptional profiling of blood or hematopoietic cell lineages in a limited set of disease conditions such as pathogen-induced host responses, autoimmune diseases, or lung inflammatory diseases and the inflammatome signatures were identified within each disease category but not across disease categories (Jenner and Young 2005; Gilchrist, Thorsson et al. 2006; Nilsson, Bajic et al. 2006; Pennings, Kimman et al. 2008; Pankla, Buddhisa et al. 2009; O'Hanlon, Rider et al. 2011). Many non-traditional inflammatory diseases such as obesity, diabetes, atherosclerosis, pain, and sarcopenia had not been included in such analysis. Therefore, it was not clear 1) whether the common signatures are shared across different tissue types and across different types of inflammatory diseases/conditions, 2) whether the common signature genes have causal relationships with the diseases and if so, whether they have therapeutic potentials, and 3) whether coherent gene-gene interaction networks and regulatory mechanisms underlying the common signature can be derived. We feel that our study has adequately addressed all these critical questions and thus represents a major advance in the field.

The reviewer may have also misunderstood the meaning of the significant overlaps. Although there is a statistically significant over-representation of these previously identified signatures in our driver and inflammatome lists, 51% or 77 of the predicted key drivers and 74% or 1831 of the inflammatome signature genes from our study are novel, suggesting that our approach is complementary to all the previous studies. Moreover, none of the previous studies provides a framework to systematically infer causal relationships among the signature genes and to further predict and validate potential drivers of inflammation. Therefore, there are a number of novel findings based on this integrative network approach and the work will serve as a blueprint for future studies of inflammation in many diseases.

As for the association with drug targets, it is difficult to imagine how the data will be used for drug design and development: e.g. VEGF is found here as member of the inflammation signature, yet the role of VEGF in diseases such as atherosclerosis or diabetes is controversial with both beneficial and harmful effects described in numerous publications.

Response:
Our study only establishes the importance of these genes in multiple inflammatory diseases and the fact that they are enriched for current drug targets under testing substantiates their therapeutic potential. We agree with the reviewer that not every single gene in our inflammatome signature or driver list is suitable as drug targets. More in-depth investigation of individual genes is needed before proving their therapeutic utility. We have now removed the VEGF discussion to avoid misinterpretation.