

GeneSpring 解析結果報告書
(JobNo.5021)

2013 年 3 月 21 日



ごあいさつ

この度は当社遺伝子発現受託解析サービス・解析オプションサービスをご用命頂き、誠にありがとうございます。本報告書は、Affymetrix 社製 **GeneChip®**を用いた遺伝子発現測定結果に、GeneSpringGX による基本解析結果をまとめております。

ご報告内容をご確認、精査頂き、今後も本サービスをご活用頂くことができれば幸いです。

株式会社トランスジェニック

1. はじめに(ご注意)

本解析はアジレント社製の解析ソフト GeneSpringGX を用いて行っております。結果に添付されておりますアノテーション情報(annotation)は、Array platform ごとに Affymetrix 社またはアジレント社のサーバーからダウンロードしたものです。

Annotation ファイルが新しいものに更新されている場合もありますので、注目遺伝子を絞り込んだ時や他手法の追加実験を行う前に、必ず配列や ID などでの最終確認を行う事をお勧めします。

2. 解析結果の見方

基本解析で行っている解析方法と、解析結果の見方を説明します。

弊社では基本解析に GeneSpringGX11.5 を使用して、アレイから得られたプローブセットの蛍光シグナル値から Summarization および Normalization を行い、比較できる数値データにしております。

2.1 Normalization

GeneSpring では以下の Normalization を行っています。

特に GeneViewer で Ratio 解析をされる場合にはチップ間による実験誤差をなくするため、Normalized 値を選択するようにして下さい。

●RMA Algorithm

RMA Algorithm では以下の 3 つのステップによりシグナル値を算出します。

1. Background adjustment
2. Quantile normalization
3. Summarization

1. Background adjustment

非特異的なハイブリダイゼーションや非特異的な蛍光シグナル値などによるバックグラウンドの値を考慮してシグナル値を補正します。

2. Quantile normalization

アレイ間のシグナルを標準化(Normalization)します。RNA 抽出・ハイブリ・染色・洗浄・スキニングのムラなどの実験誤差による変動をキャンセルして、生物学的な変動を抽出するために行います。

ホールゲノムマイクロアレイのような大量の雑多な遺伝子が載っているチップを使っていて、大部分の遺伝子が実験的なパラメーターに影響を受けないという仮定にたてば、シグナルの分布はほぼ等しくなると予想されます。この仮定の下、解析する全アレイのヒストグラムの形を揃えるように補正します。

3. Summarization

probe level のシグナル値から probe set level の発現シグナル値へ変換します。

●Baseline Transformation: Baseline to median of all samples

発現パターンから生物学的な意味を抽出するのを助ける目的で行います。発現量の多少に関わらず、とにかく同じ時に増えたり減ったりしている遺伝子は、①機能的に近い関係にある、②同じ転写制御の元にある、③同じカスケードの下流にある、といった仮定に基づいて解析を行う場合に有効です。

遺伝子ごとに全サンプルの Normalized 値(log₂ 値)の中央値を計算し、各サンプルの Normalized 値から引きます。

2.2 遺伝子リスト

GeneSpring による解析結果は遺伝子リストの形で出力されます。弊社が無償でご提供しております解析簡易ソフト *GeneViewer* のほか Excel などの表計算ソフトでもご覧になれます。Figure1 はエクセルで開いた時の遺伝子リストです。

	A	B	C	D	E	F	G	H	I	J	K	L
1	# Notes :	Entity List:All Entities										
2	#Interpretation:	All Samples										
3	#Experiment :	User ID										
4	#Search Condition											
5	#Detection-1	1111111_A <= 0.2										
6	#samples:	at least 1 out of 2 samples satisfies the conditions.										
7	# Technology :	Illumina SingleColor:MouseWG-6_V2_0_R0.11278593_A										
8	# Owner :	gxuser										
9	# Created On	Fri Dec 26 11:27:44 JST 2008										
10	ProbeID	User Fold	Regulation	Number Pa	ID1, Name	ID2, Name	ID1, Name	ID2, Name	Chromosom	Symbol	Definition	Probe_Seq
11	2910427			1	-8.3	0.8	0.254438	0.008868		0610007L01Rik	GAGCCTG	
12	70600			1	0.1	9.4	0.254438	0.008868		0610007L01Rik	GGACTGG	
13	6290215			2	4.2	-3.1	1.200055	0.008868	chr12	0610009D(Mus muscu	TTAATTAC	
14	1450010			1	-11.1	-16.9	-1.9759	0.008868		0610010D24Rik	GCAAATGC	
15	2360348			1	7.5	-7.5	3.161329	0.008868		0610012A05Rik	TCAGGATC	
16	6480593			1	-10.8	-8.2	0.244884	0.008868		0610012A05Rik	TACACTGC	
17	4890082			1	-16.8	-11.1	0.244884	0.008868		0610012K18Rik	AGCCACCC	
18	2570348			2	-27.2	-15.5	0.254438	0.008868	chr10	0610025O(Mus muscu	ATAAAAAT	
19	5860463			5	-1.2	-18.3	0.253751	0.008868		0610033E06Rik	CTACCTTC	

Figure1 遺伝子リスト

Table 1

①	解析対象ファイル
②	解析パラメーター設定
③	プロジェクト名 (ユーザーID)
④	解析条件
⑤	Technology : アレイ名
⑥	ファイルプロパティ
⑦	Probe Set ID : プロブセット固有の ID
⑧	User Fold Difference : 空欄 (<i>GeneViewer</i> で解析を行うと値が入ります)
⑨	User Regulation : 空欄(<i>GeneViewer</i> で解析を行うと値が入ります)
⑩	Number Passed : フィルターを通過したサンプル数を表します。 (フィルタリングファイルのみ) Fold change([sample] vs [control]), Regulation: [control]に対する[sample]の変動倍率です。Regulation が Up で[control]に対して[sample]の発現が増加、Down で減少していることを表します。(Fold Change フィルタリングファイルのみ)
⑪	[Sample ID, Sample Name](raw) : 画像ファイルから算出された蛍光シグナル強度データです。CEL ファイルの値となります。
⑫	[Sample ID, Sample Name] (normalized) : GeneSpring で Normalization を行った Normalized 値
⑬	アノテーション(Table 2)

Table 2

カラム名	内容
Chromosomal Location	遺伝子の染色体番号
Gene Symbol	遺伝子名
Gene Title	遺伝子の詳細な説明
InterPro	InterPro ID
Ensembl	Ensembl ID
Pathway	GenMAPP 上で該当する MAPP file 名
Annotation Description	Transcript Cluster のアレイデザイン
Entrez Gene	Entrez Gene ID
Representative Public ID	GenBank ID
UniGene ID	Unigene ID
RefSeq Transcript ID	RefSeq ID
RefSeq Protein ID	Pfam ID
OMIM	OMIM ID
Gene Ontology Cellular Component	遺伝子機能分類
Gene Ontology Molecular Function	遺伝子機能分類
Gene Ontology Biological Process	遺伝子機能分類

2.3 発現値によるフィルタリング

Raw 値(シグナル値)の信頼性が低い遺伝子を除きます。

一般に、発現レベルが非常に低い遺伝子の測定値は信頼性が低く、その値を分母にして計算される Normalized 値の信頼性も低くなります。

Raw Data を Expression level でフィルタリングし、全てのサンプルでシグナル値が 20th Percentile より小さい遺伝子を除きます。

3. 解析結果

3.1 解析条件

GeneSpringGX version	GX11.5.1
Array platform version	Affymetrix GeneChip® Human Genome U133 Plus 2.0
Annotation file	Affymetrix.GeneChip.HG-U133_Plus_2

3.2 解析目的

遺伝子変異による疾病について、患者、保因者の発現を正常者と比較することで、遺伝子発現の変化を解析することを目的とします。

Sample ID	Sample Name	Cell type	Group
33U1430398-H01	MSC0316	MSC	Normal
33U1430398-H02	MSC0408	MSC	Normal
33U1430398-H03	MSC0423	MSC	Normal
33U1430398-H04	MSC0610	MSC	Normal
33U1430398-H05	MSCSHIMANE1	MSC	Patient
33U1430398-H07	MSCSHIMANE2	MSC	Patient
Shimane_Uni01	HPP01/pQC cl.2#2 iPS-MSC P32(P8)	MSC	iPS-MSC
Shimane_Uni02	HPP02/pQC cl.14#3 iPS-MSC P15(P4)	MSC	iPS-MSC
Shimane_Uni03	HPP01/pQC+ALP cl.2#2 iPS-MSC P35(P11)	MSC	iPS-MSC+ALP
Shimane_Uni04	HPP02/pQC+ALP cl.14#3 iPS-MSC P18(P7)	MSC	iPS-MSC+ALP

3.3 解析方針

- 1) Normalization を行います。
- 2) Expression Level によるフィルタリングを行い、シグナル値による遺伝子(Raw Data で 20th Percentile より大きい遺伝子)の絞り込みを行います。
- 3) 変動倍率によるフィルタリングを行い、Patient に対して iPS-MSC で 2 倍以上変動している遺伝子を絞り込みます。
- 4) 3)で絞り込んだ遺伝子リストにおける階層クラスタリング解析を行い、その結果から選択された遺伝子群のリスト化を行います。
- 5) 変動倍率によるフィルタリングを行い、Normal に対して iPS-MSC+ALP で 2 倍以上変動している遺伝子を絞り込みます。
- 6) 5)で絞り込んだ遺伝子リストにおける階層クラスタリング解析を行い、その結果から選択された遺伝子群のリスト化を行います。

結果 1 全遺伝子の解析結果

全遺伝子の発現値とアノテーション情報を出力します。

Parameter	Value
Gene List	All Entities
File name	01_All Entities_5021.txt
Output number of genes	54675

結果 2 シグナル値による遺伝子の絞り込み

全サンプルのうち少なくとも 1 サンプルが、Raw Data で 20th Percentile より大きい遺伝子に絞ります。

Parameter	Value
Gene list	01_All Entities_5021
Value	Lower cut-off: 20 th Percentile
Pass	10 サンプル中 1 サンプル
File name	02_Filtered on Expression (20.0 - 100.0)th Percentile in the Raw Data_5021.txt
Output number of genes	48615

結果 3 遺伝子の変動倍率による絞り込み

あるサンプルに対して、一方のサンプルの発現が 2 倍以上増加及び減少している遺伝子を絞るため、変動倍率：fold change を使用します。

Fold change = Condition1/Condition2 を計算します。

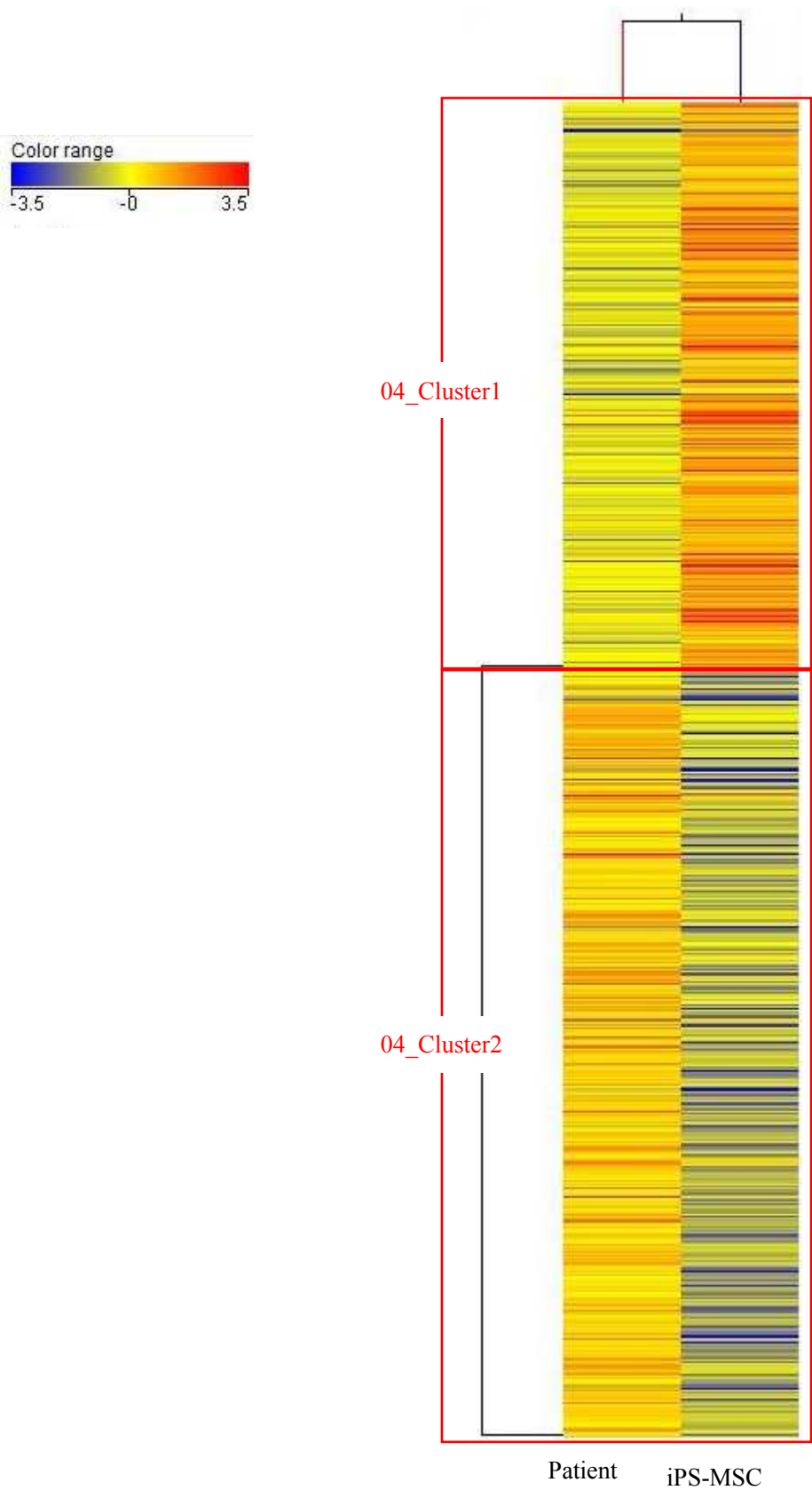
Patient に対して iPS-MSC が 2 倍以上発現変動している遺伝子

Parameter	Value	Number of genes
Input gene list	02_Filtered on Expression (20.0 - 100.0)th Percentile in the Raw Data_5021.txt	48615
Condition1	iPS-MSC	
Condition2	Patient	
Fold Difference	2.0	
Output gene list	03_FC2_iPS-MSC vs Patient_5021.txt	5287

結果 4 クラスタリング解析

結果 3 の遺伝子リストを用い、Hierarchical Clustering(階層型クラスタリング)により発現パターンの視覚化を行います。その結果からお客様が選択された遺伝子群のリストを作成します。

Parameter	Value	Number of genes
Input gene list	03_FC2_iPS-MSC vs Patient_5021.txt	5287
Clustering Algorithm	Hierarchical	
Cluster on	Both Entities and conditions	
Distance metric	Pearson Centered	
Linkage rule	Average	
Output gene list	04_Cluster1_5021.txt	2248
	04_Cluster2_5021.txt	3039



結果 5 遺伝子の変動倍率による絞り込み

あるサンプルに対して、一方のサンプルの発現が 2 倍以上増加及び減少している遺伝子を絞るため、変動倍率：fold change を使用します。

Fold change = Condition1/Condition2 を計算します。

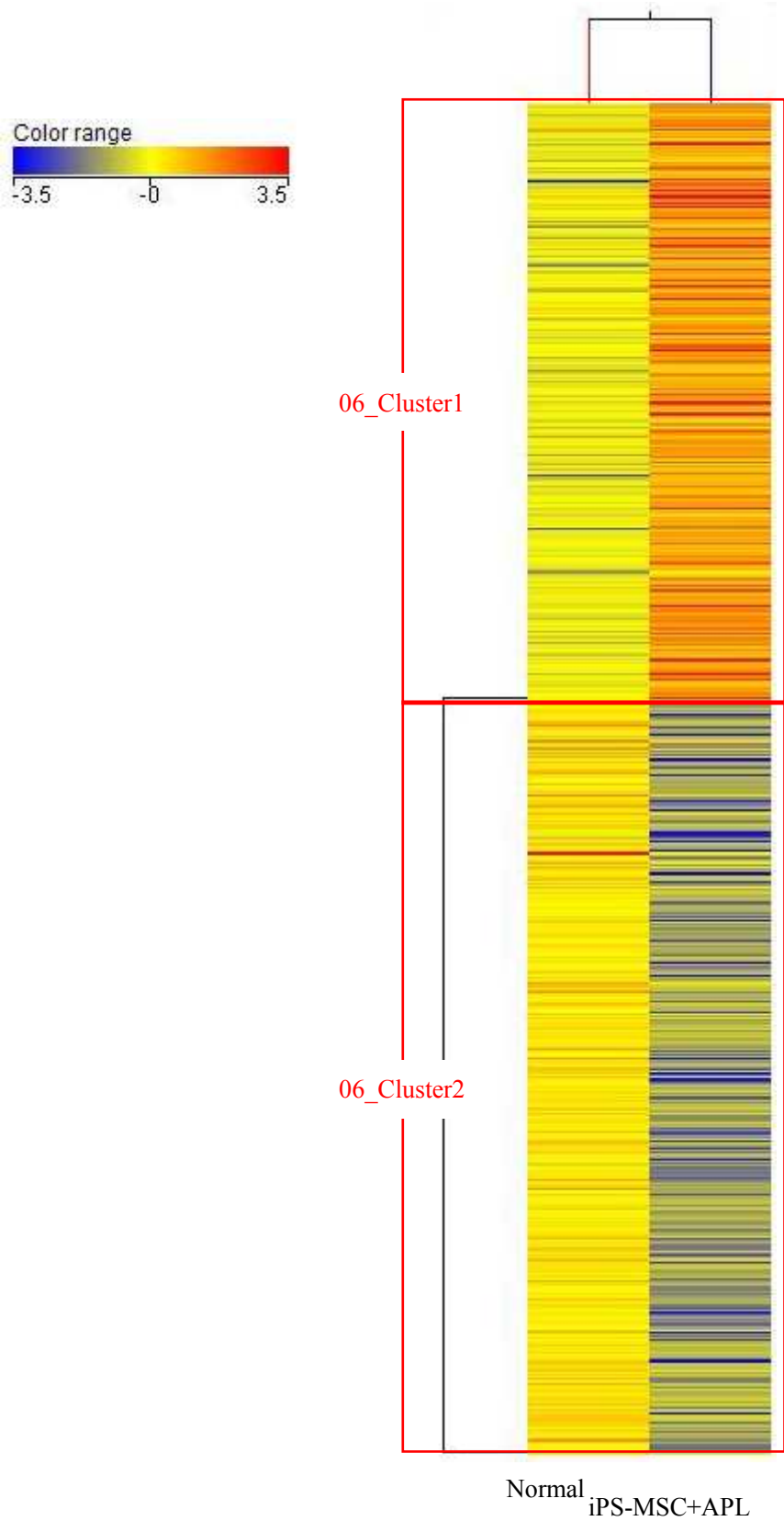
Normal に対して iPS-MSC+ALP で 2 倍以上変動している遺伝子

Parameter	Value	Number of genes
Input gene list	02_Filtered on Expression (20.0 - 100.0)th Percentile in the Raw Data_5021.txt	48615
Condition1	iPS-MSC+ALP	
Condition2	Normal	
Fold Difference	2.0	
Output gene list	05_FC2_iPS-MSC+ALP vs Normal_5021.txt	4645

結果 6 クラスタリング解析

結果 5 の遺伝子リストを用い、Hierarchical Clustering(階層型クラスタリング)により発現パターンの視覚化を行います。その結果からお客様が選択された遺伝子群のリストを作成します。

Parameter	Value	Number of genes
Input gene list	05_FC2_iPS-MSC+ALP vs Normal_5021.txt	4645
Clustering Algorithm	Hierarchical	
Cluster on	Both Entities and conditions	
Distance metric	Pearson Centered	
Linkage rule	Average	
Output gene list	06_Cluster1_5021.txt	2058
	06_Cluster2_5021.txt	2587



<お問い合わせ>

株式会社トランスジェニック 研究開発部

〒650-0047 神戸市中央区港島南町 7-1-14

TEL: 078-306-0590 FAX: 078-306-0589