

Microarray Data Analysis Tool

複数比較解析の進め方

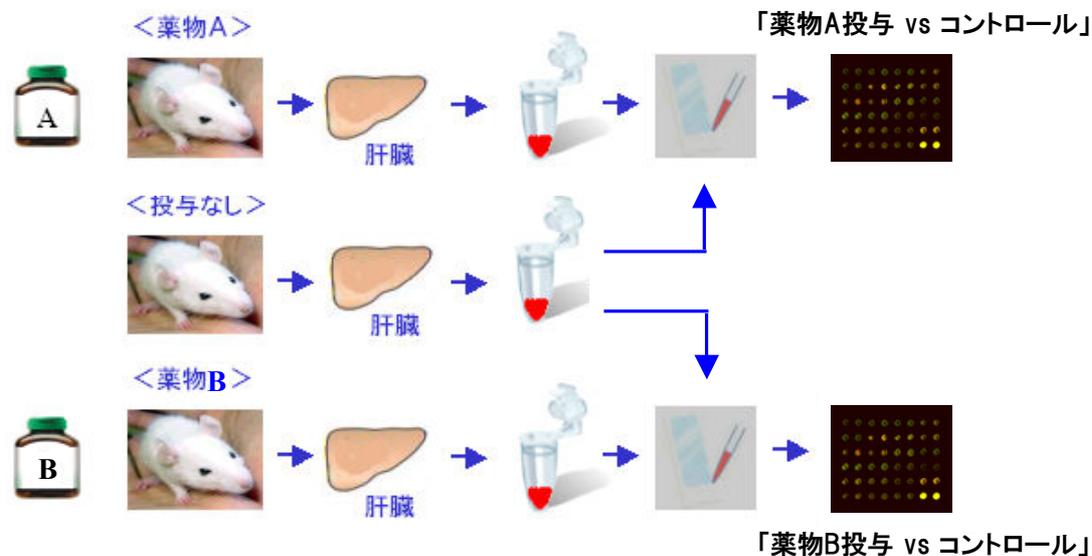
1比較解析に引き続き、複数比較解析についてのソフトウェアの使用方法を、実際の解析例と合わせて説明します。

解析テーマは

「薬物A投与マウスおよび薬物B投与マウスにおける発現変動遺伝子群の抽出と機能解析」です。

薬物Aを投与したマウスおよび薬物Bを投与したマウスをテストサンプル、投与していないマウスをコントロールサンプルとした実験例を紹介します。まず、二色法による「薬物A投与 vs コントロール」および「薬物B投与 vs コントロール」の2比較について、遺伝子発現プロファイルをマイクロアレイ実験で比較します。その後、比較データの中から、コントロールサンプルに対して、薬物A投与マウスおよび薬物B投与マウスで高発現、または低発現している遺伝子を絞り込み、これらの遺伝子がどのような機能をもっているのか調べます。この結果、これらの2種の薬物が生体に対してどのような作用をするのか考察することができます。

<実験概要> サンプル調整 → 組織の抽出 → 蛍光ラベル → ハイブリ → スキャニング



複数比較解析では変動遺伝子群の扱い方が1比較解析と異なります。例として今回の複数比較解析では、変動遺伝子群が次の3つの要素に分けられることを考慮しなければいけません。

- i) 薬剤Aおよび薬剤Bの投与によって発現が共通に変動した遺伝子群
- ii) 薬剤A投与特異的に発現が変動した遺伝子群
- iii) 薬剤B投与特異的に発現が変動した遺伝子群

i)、ii)、iii)のどの変動遺伝子群に着目するかによって解析対象の薬剤が変わってきますので、この点が複数比較解析の主な特徴の一つです。

1. ソフトウェアの起動と実験データの読み込み

1-1. ソフトウェアの準備

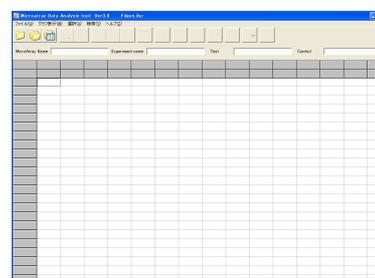
CDに入っている「Microarray Data Analysis Tool」のフォルダを、自分のコンピュータにコピーしてください。

1-2. ソフトウェアの起動

コピーしたフォルダを開き、
「MAtool.exe」のアイコンをダブルクリックしてください。
右図のタイトル画面が表示され、ソフトウェアが起動します。

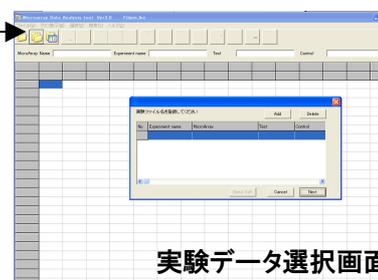


ソフトウェアが起動後は右図のメイン画面が表示されます。



1-3. 実験データ読み込み

アイコンをクリックすると、下記の画面が表示されます。「Add」ボタンをクリックし、対象となる実験データの選択・読み込みを行います。



実験データ選択画面

実験データを読み込むと、右図のように選択した実験データ名・サンプル名等が表示されます。

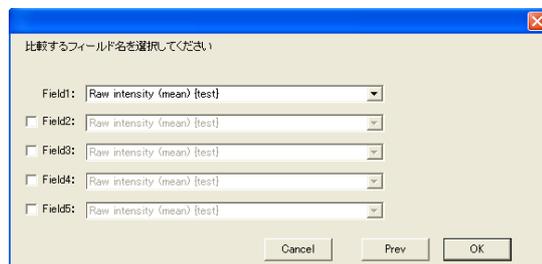
表示の確認が終わったら、「Next」ボタンをクリックして次へ進みます。



「Next」ボタンをクリックすると、比較するフィールド名を選択する画面が現れます。複数比較の場合、選択したフィールド名のみをメイン画面上に表示させます。ここでは最大5つのフィールドを選択・表示することができます

通常、ノーマライズ後の以下のフィールドを選択します。

- 「Normalized intensity (mean) {test}」
- 「Normalized intensity (mean) {control}」
- 「Normalized intensity (mean) {sum}」
- 「Normalized intensity (mean) {ratio}」
- 「Ignored cells」



「OK」ボタンをクリックするとデータがメイン画面に表示されます。

「薬物A投与 vs コントロール」 「薬物B投与 vs コントロール」

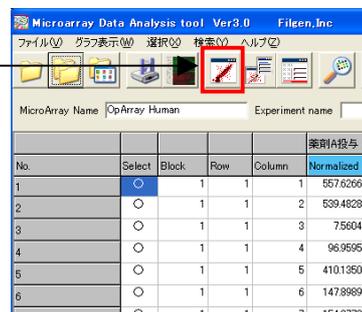
No.	Select	Block	Row	Column	Normalized intensity (mean) {test}	Normalized intensity (mean) {control}	P-value	FDR	配列ID
1	<input type="radio"/>	1	1	1	557.6262	701.3822	1.5582484	0.76112	0
2	<input type="radio"/>	1	1	2	539.48285	881.79059	1421.2454	0.91152	0
3	<input type="radio"/>	1	1	3	7550.619	39742	1732669	0.78501	0
4	<input type="radio"/>	1	1	4	56365658	4459417	14125378	21594	0
5	<input type="radio"/>	1	1	5	41012603	48554061	595.7354	0.84482	0
6	<input type="radio"/>	1	1	6	14739696	1563839	2864828	1.06722	0
7	<input type="radio"/>	1	1	7	15437702	22735635	981.73337	0.67901	0
8	<input type="radio"/>	1	1	8	1819426	1234282	31.03718	1.47688	0
9	<input type="radio"/>	1	1	9	2353979	1721091	40.6507	1.26982	0
10	<input type="radio"/>	1	1	10	4090082	6470209	105.97991	0.62967	0
11	<input type="radio"/>	1	1	11	101.70091	8636468	188.66559	1.16945	0
12	<input type="radio"/>	1	1	12	2628077	1634385	43.11282	1.55995	0
13	<input type="radio"/>	1	1	13	24056162	4641991	289.09152	4.97003	0
14	<input type="radio"/>	1	1	14	10027604	103.76248	204.03182	0.96682	0
15	<input type="radio"/>	1	1	15	1144.69174	695.62213	1043.31387	1.63385	0
16	<input type="radio"/>	1	1	16	30316019	35.03571	651.1959	0.94447	0
17	<input type="radio"/>	1	1	17	74052618	75258458	1493.12076	0.93099	0
18	<input type="radio"/>	1	1	18	23022345	14919986	379.42211	1.54007	0
19	<input type="radio"/>	1	1	19	8222911	7670247	158.29158	1.02192	0
20	<input type="radio"/>	1	1	20	110.05724	131.23096	241.2882	0.83885	0
21	<input type="radio"/>	1	1	21	176.22043	107.63915	283.85968	1.93714	0
22	<input type="radio"/>	1	1	22	12.00685	1728641	29.34726	0.99771	0
23	<input type="radio"/>	1	1	23	132.13164	136.78941	308.93806	1.25842	0

上図のように、選択したフィールドのデータが実験ごとに表示されます。

2. スキャタープロットの作成

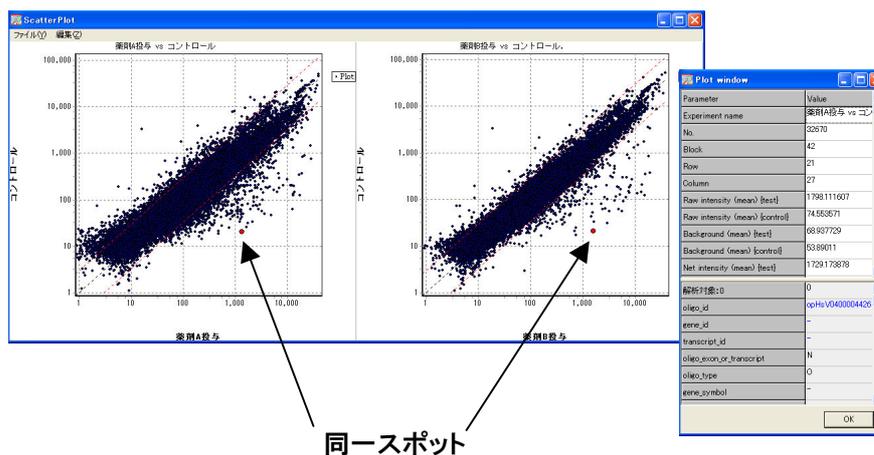
2-1. スキャタープロットの表示

メイン画面のScatterPlot表示ボタンを押してください。あとは、自動でグラフが表示されます。



No.	Select	Block	Row	Column	Normalized Intensity
1	<input type="radio"/>	1	1	1	567.62662
2	<input type="radio"/>	1	1	2	539.48286
3	<input type="radio"/>	1	1	3	7.56049
4	<input type="radio"/>	1	1	4	96.96958
5	<input type="radio"/>	1	1	5	4101.3503
6	<input type="radio"/>	1	1	6	147.89896

Scatter Plotの各プロットをクリックすると、選択されたプロットが赤いプロットに変わります。そして、メイン画面の対応する遺伝子(行)が青色で選択され、行の先頭にスクロール表示されます。また、1つのグラフのプロットをクリックすると、他のグラフの対応するプロットも赤くなります。さらに、メイン画面のPlot windowアイコンをクリックすると、Scatter Plot中の各プロット個別のデータを表示するPlot window画面が現れます。



上記までのステップで スキャタープロットを表示させ、2比較におけるサンプル間の遺伝子発現の全体像を見ることができました。また、注目したいプロットの表示を行うことで、薬剤A投与と薬剤B投与による該当遺伝子の発現パターンの比較を視覚的に確認することができました。

スキャタープロットが中心ライン付近に収束していれば、比較サンプル間の発現差は小さいと解釈が出来ます。ばらつきの幅が広ければサンプル間における発現差が大きいと解釈できます。

次は、変動遺伝子の抽出となります。

3. 発現変動遺伝子の抽出

3-1. Filter Options検索画面の表示

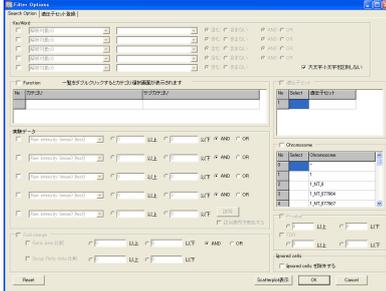
メイン画面のFilter Optionsアイコンをクリックし、Filter Options検索画面を開きます。



No.	Select	Block	Row	Column	薬剤A投与	Normalized
1	<input checked="" type="radio"/>	1	1	1	1	557.62662
2	<input type="radio"/>	1	1	2	2	539.48285
3	<input type="radio"/>	1	1	3	3	7.56049
4	<input type="radio"/>	1	1	4	4	96.95958
5	<input type="radio"/>	1	1	5	5	4101.3503
6	<input type="radio"/>	1	1	6	6	147.89695
7	<input type="radio"/>	1	1	7	7	154.97799

Filter Options検索画面では各種検索条件でデータの抽出を行うことができます。

keyword
機能カテゴリー
実験データ
染色体番号
遺伝子セット



3-2. 検索条件を設定

ここでは下記の目的を前提に条件設定を行います。

- ①コントロール遺伝子などを除去し、解析対象遺伝子のみにする。
- ②スポット上にごみ等があり、シグナルに影響するデータを除外する。
- ③発現差のあるデータを抽出する。
- ④ノイズレベルの信頼性の低いデータを除外する。

条件設定の方法

①コントロール遺伝子などを除去し、解析対象遺伝子のみにする。

ここでは、keyword検索を使用します。事前にアレイに搭載されているプローブは「0、2、3」という数字で識別をしています。

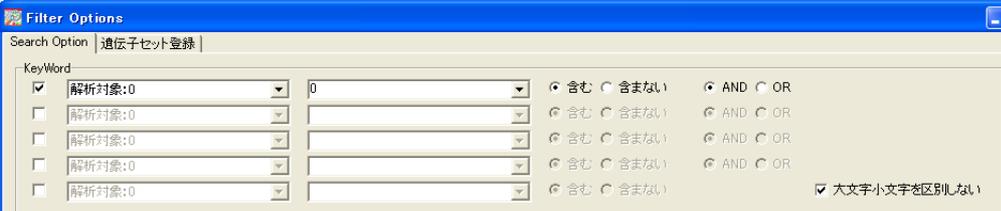
「0」は解析対象遺伝子

「2」は解析対象外遺伝子(NegativeControl,Blank,QCコントロールなど)

「3」はデータベースの更新により対応プローブの情報が削除された遺伝子

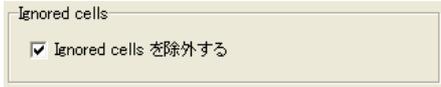
***アレイの種類により解析対象の条件設定が必要ない場合もあります。**

まずは、左のチェックボックスをクリックし、該当項目で「解析対象:0」と、「0」を選択してください。これで、解析対象遺伝子のみを抽出する条件設定ができました。



②スポット上にごみ等があり、シグナルに影響するデータを除外する。

ここでは、Ignore cellの設定を使用します。
事前にアレイに搭載されているデータの中でごみなどの影響を受け、正確なシグナルを算出できていない可能性があるデータはメイン画面で表示されているIgnored cellsの項目が「1」になっています。上図、チェックボックスをクリックすることで、Ignored cellsを除外することができます。

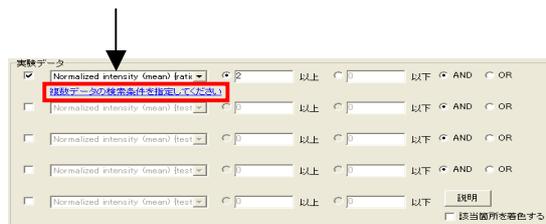


③発現差のあるデータを抽出する。

ここでは、実験データ検索を使用します。実験の内容や結果によっても変わってきますが、シグナルの差が任意の倍数以上あるデータを差のあるデータとして抽出します。

左のチェックボックスをクリックし、項目「Normalized intensity (mean) {ratio}」で条件を設定します。(下図は、「Normalized intensity (mean) {ratio}」を2以上にした例です。)

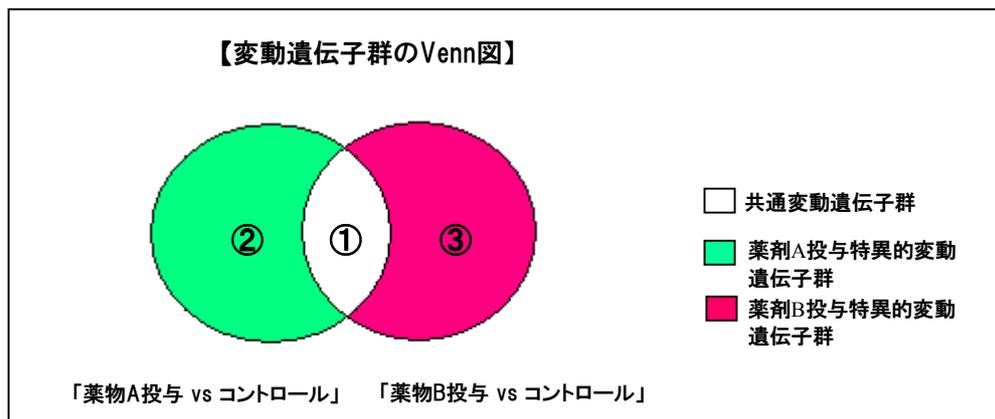
続いて、「複数データの検索条件を指定してください」の項目をクリックします。



指定した検索に対して、どのデータを対象として実行するか選択し、チェックボックスをクリックします。

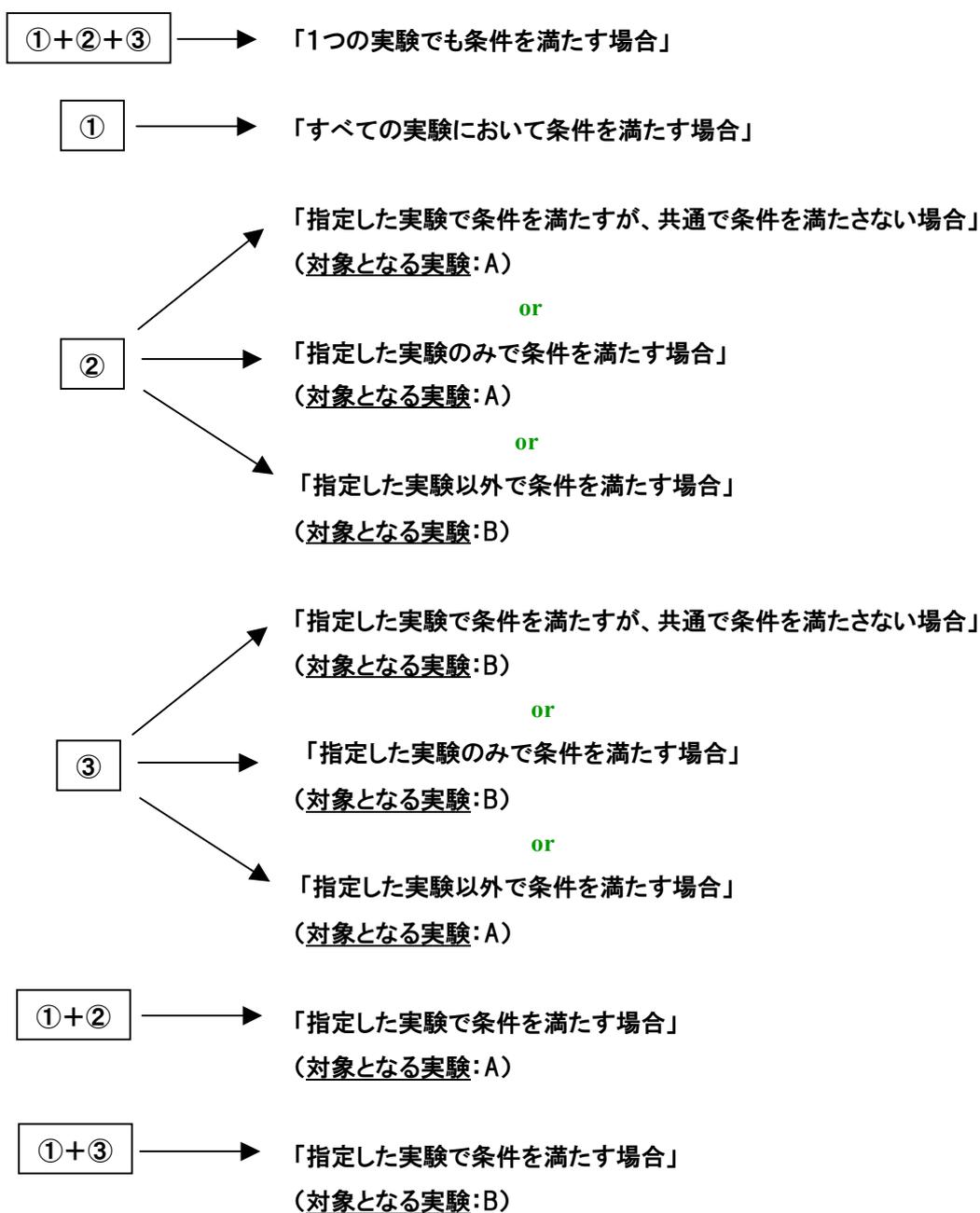


今回の2比較(「薬物A投与 vs コントロール」、「薬物B投与 vs コントロール」)では変動遺伝子群が3つの要素に分けられることを先ほど触れましたが、検索したい遺伝子群と実験条件検索の組み合わせを次ページで図示します。



検索したい遺伝子群

検索条件



④ノイズレベルの信頼性の低いデータを除外する。

ここでは、③と同様、実験データ検索を使用します。複数比較の場合、各比較ごとにCutoffを設けるか、共通のCutoffを設ける場合が多いです。一例としてエクセルで作成された「薬物A投与 vs コントロール」および「薬物B投与 vs コントロール」の解析データのCutoffの条件設定を参考に設定をします。

エクセルで作成された解析データのCutoffの条件設定
「薬物A投与 vs コントロール」のCutoff → Normalized intensity (mean)がnegative-controlのNet intensity(sum)の平均値(=300)、
「薬物B投与 vs コントロール」のCutoff → Normalized intensity (mean)がnegative-controlのNet intensity(sum)の平均値(=200)

例1. 各比較ごとにCutoffを設定する場合

まず、「薬物A投与 vs コントロール」のCutoffの条件設定を行います。左のチェックボックスをクリックし、項目「Normalized intensity (mean) {sum}」で「300」以上、かつ、複数データの検索条件の選択において、「指定した実験で条件を満たす場合」(対象となる実験:「薬物A投与 vs コントロール」)を選びます。

同様に、「薬物B投与 vs コントロール」のCutoffの条件設定を行います。左のチェックボックスをクリックし、項目「Normalized intensity (mean) {sum}」で「200」以上、かつ、複数データの検索条件の選択において、「指定した実験で条件を満たす場合」(対象となる実験:「薬物B投与 vs コントロール」)を選びます。

下図は、コントロールに対して2倍以上のUp共通変動遺伝子の抽出で、例1の条件設定を行った抽出ノイズ除去の設定例です。



例2. 共通のCutoffを設定する場合

まず、共通のCutoffの値を決めます。ここでは、共通のCutoffとして、2比較の平均値のうち、値が小さい平均値(=200; 「薬物B投与 vs コントロール」)を採用した設定を以下に示します。(共通のCutoffの値は便宜、検索結果と照らし合わせて最終的に決めてください。)

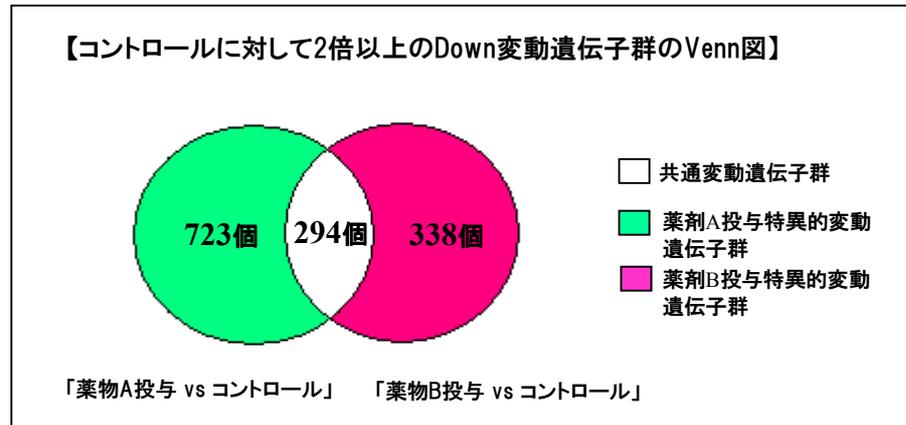
左のチェックボックスをクリックし、項目「Normalized intensity (mean) {sum}」で「200」以上、かつ、複数データの検索条件の選択において、「すべての実験において条件を満たす場合」を選びます。

下図は、コントロールに対して2倍以上のUp共通変動遺伝子の抽出で、例2の条件設定を行った抽出ノイズ除去の設定例です。



②検索後の、コントロールに対して2倍以上のDown変動した遺伝子総数:1355個

- ・共通変動遺伝子群:294個
- ・薬剤A投与特異的変動遺伝子群:723個
- ・薬剤B投与特異的変動遺伝子群:338個



①および②の検索後の変動遺伝子は、AおよびBの薬剤投与によって発現が誘導され、AおよびBの薬剤の作用を規定するうえで重要な遺伝子だと考えることが出来ます。

この後、各変動遺伝子群において、Gene Ontologyによる生物学的機能の検索と、GenMAPPによるパスウェイの検索を行い、薬剤Aおよび薬剤Bの作用に関連する遺伝子の機能を調べていくことになります。

以下の手順は、1比較解析と同じ流れになりますので、1比較解析の「[4.遺伝子セットの登録方法](#)」以降を参照してください。