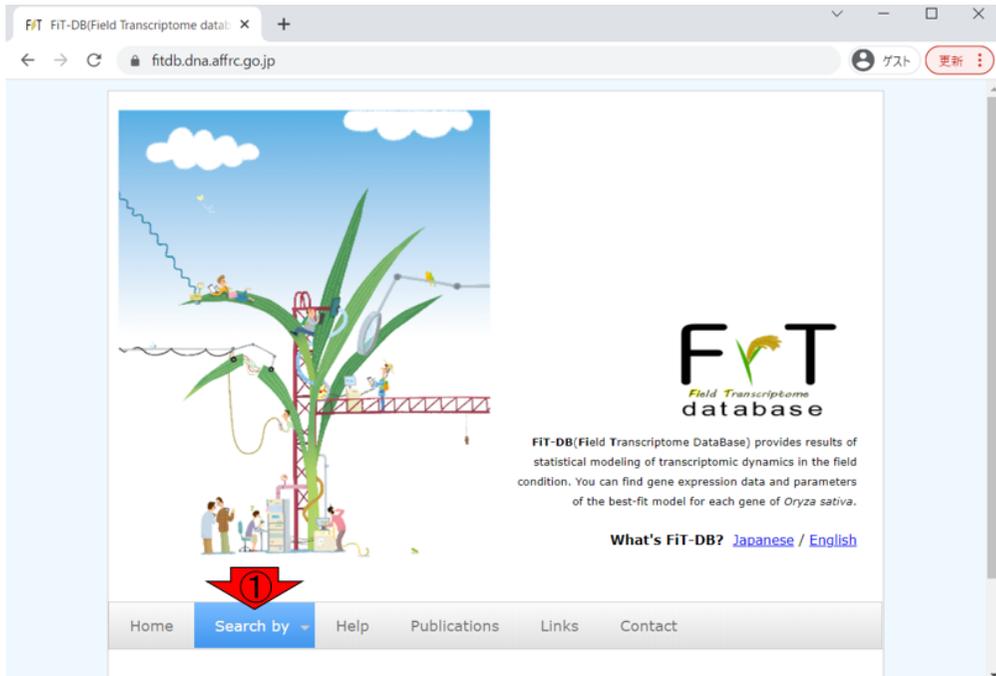


例題 7.2

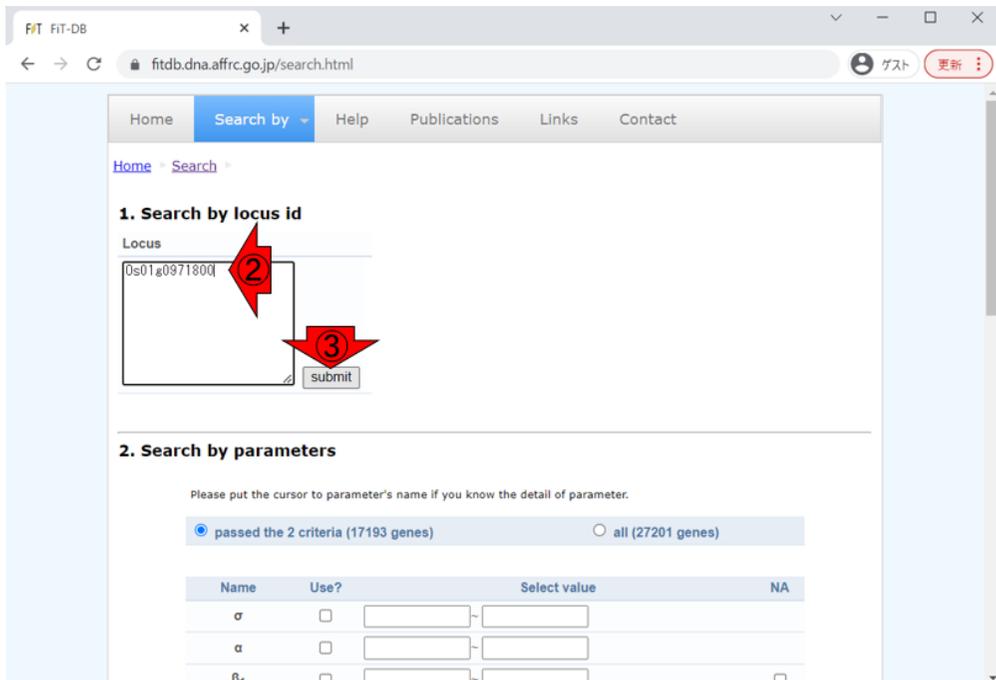
FiT-DB にアクセスし、任意のイネ遺伝子（例：[Os01g0971800](#)）を検索してどのようなフィールド環境に影響を受けやすいか確認せよ。

解答例

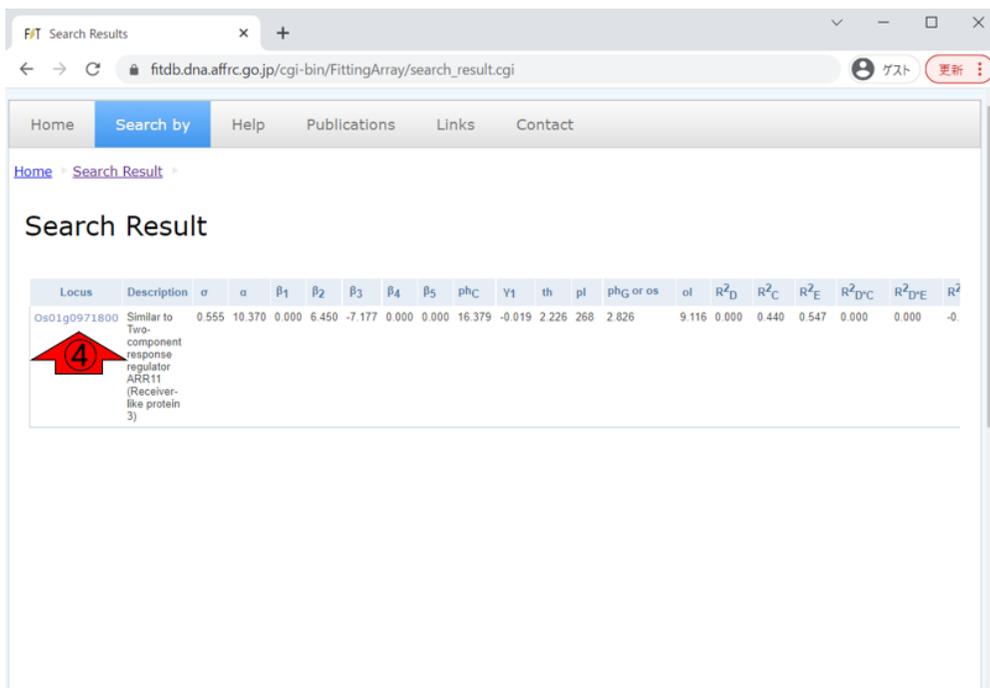
FiT-DB のホームページから「①Search by」をクリックします。



「1. Search by locus id」のボックス内で、②Os01g0971800 と打ち込んで③submit ボタンを押します。



以下のような検索結果の概要が表示されるので、④を押して詳細を眺めます。



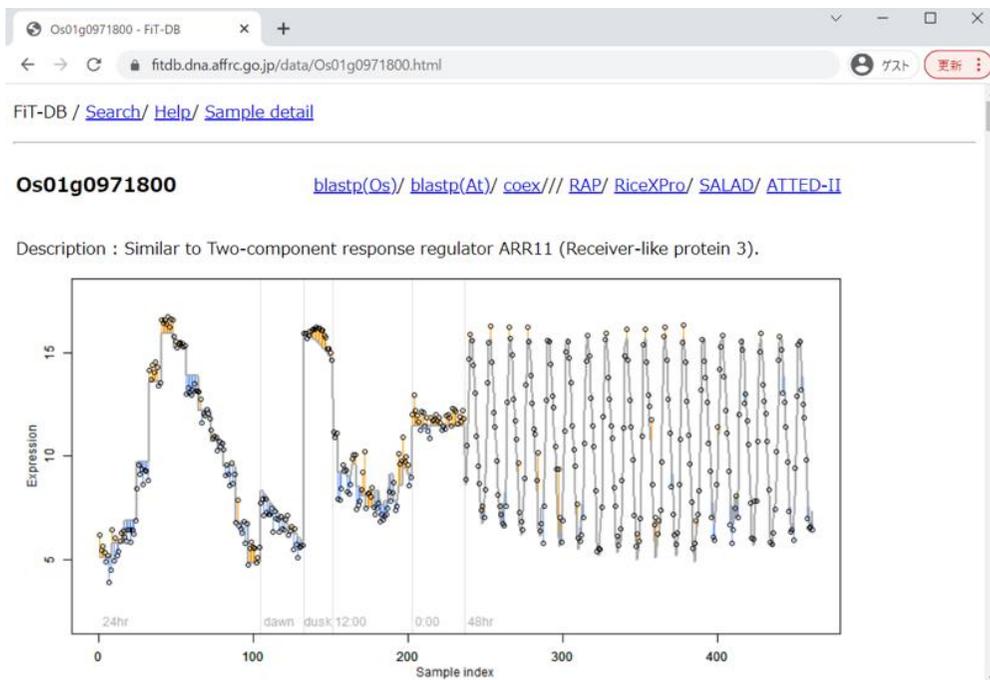
Home Search by Help Publications Links Contact

Home > Search Result >

Search Result

Locus	Description	σ	α	β_1	β_2	β_3	β_4	β_5	ph_C	γ_1	th	pl	ph _G or os	ol	R^2_D	R^2_C	R^2_E	$R^2_{D^*C}$	$R^2_{D^*E}$	R^2
Os01g0971800	Similar to Two-component response regulator ARR11 (Receiver-like protein 3)	0.555	10.370	0.000	6.450	-7.177	0.000	0.000	16.379	-0.019	2.226	268	2.826	9.116	0.000	0.440	0.547	0.000	0.000	-0.

以下のスクショが、ページ遷移後の状態です。



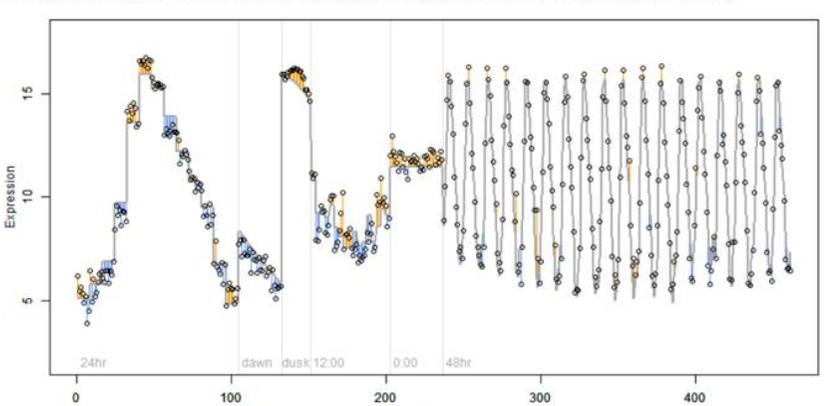
Os01g0971800 - FIT-DB

fitdb.dna.affrc.go.jp/data/Os01g0971800.html

FIT-DB / [Search/](#) [Help/](#) [Sample detail](#)

Os01g0971800 [blastp\(Os\)/](#) [blastp\(At\)/](#) [coex///](#) [RAP/](#) [RiceXPro/](#) [SALAD/](#) [ATTED-II](#)

Description : Similar to Two-component response regulator ARR11 (Receiver-like protein 3).

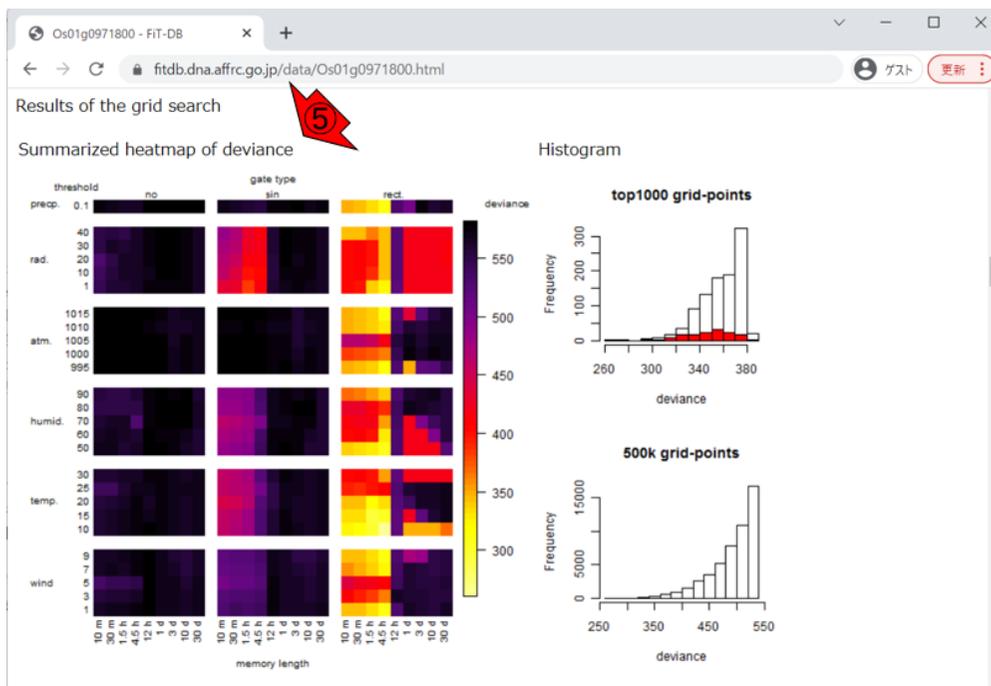


Expression

Sample index

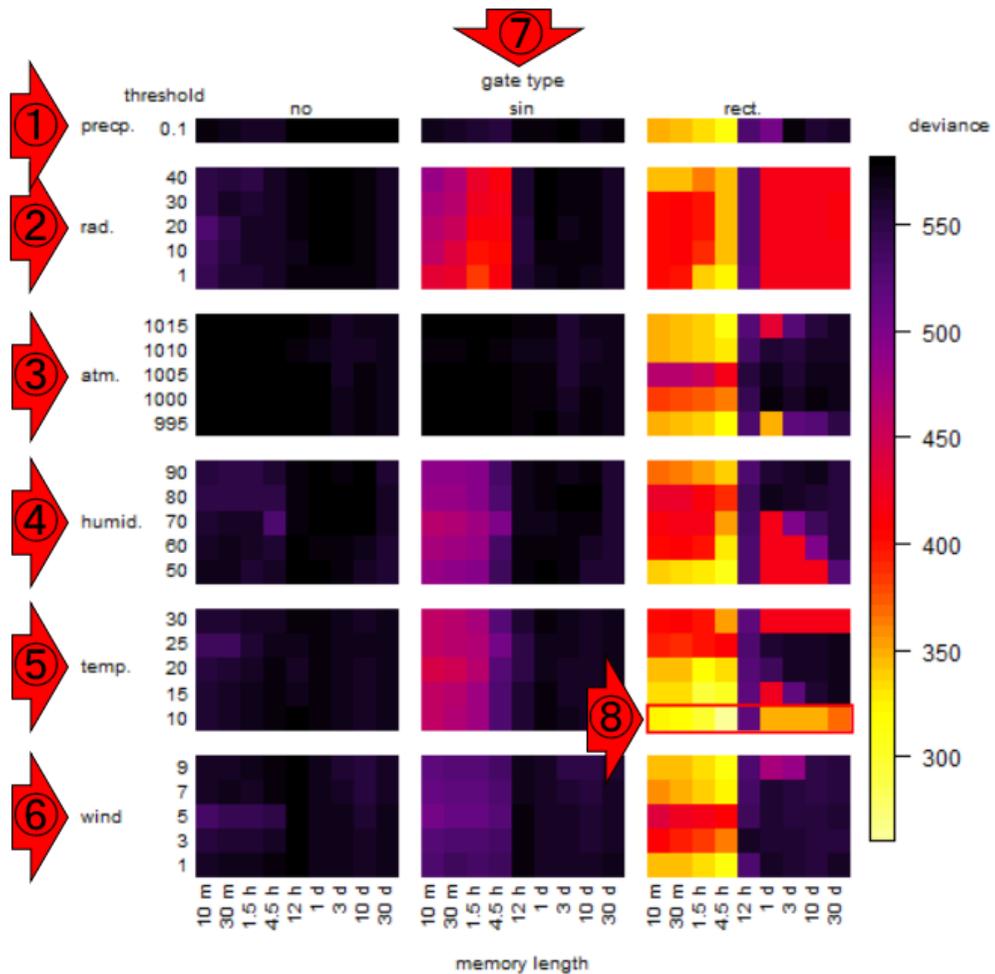
24hr dawn dusk:12:00 0:00 48hr

「⑤Summarized heatmap of deviance」が見えるところまで、ページ下部に移動した状態です。



このヒートマップは、実データへの統計モデルの当てはまりの程度を示したものであり、逸脱度 (deviance) が低いものが当てはまりの良い統計モデルだと解釈します。

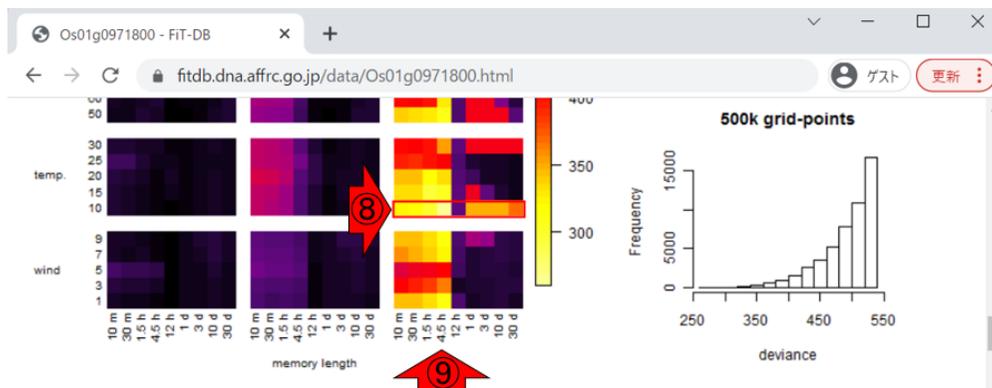
上の図は全体的に小さいので、拡大したのが下図です。



行方向は全部で 6 つのカテゴリに分けられており、それぞれ①降水量 (precipitation)、②日射量 (radiation)、③大気圧 (atmospheric pressure)、④湿度 (humidity)、⑤気温 (temperature)、⑥風速 (wind velocity) です。それぞれのカテゴリ中で、さらに (①以外は) 5 種類の閾値 (threshold) で分かれています。列方向は、⑦ゲート効果 (gate type) が示されています。これは「特定の時刻帯にのみ環境応答を示す現象」のことで、3 タイプ (no : ゲートなし、sin : サイン型ゲート、rect. : 矩形ゲート) に分けられています。このゲート効果はさらに、図の下部に示されているように 9 種類の memory length で分けられている。これはどの程度前の時間に①～⑥のパラメータの値であった場合に当該遺伝子 (つまり [Os01g0971800](#)) の発現に影響を及ぼすかを示すものです。

今我々が知りたいのは、当該遺伝子がどのような条件のもとで発現変動するかどうかであり、その条件を探るために様々なパラメータで探索した結果を疑似カラーで大まかに眺めているということになります。今全体として着目すべき箇所は、右側の deviance の値が低いことを表す黄色っぽいセルになります。全体的に最も黄

色っぽいのは、⑦ゲート効果が一番右側の矩形ゲート (rect.)、⑤温度の閾値が 10 度の⑧で示したあたりだということがわかります。この疑似カラーで示したものを実際の数値として、当てはまりのよい上位の統計モデル (つまりセル) をリストアップしたものが、下図の真ん中あたりに書かれている「Local optima within top 1000 grid-points」になります。最も当てはまりのよいモデルは、deviance が 260.32 のものであり、⑧の行と⑨の列の交点に相当するセルだということが分かります。⑨4.5h (つまり 270 分) 前までの温度がこの遺伝子の発現に影響しているのだと読み解きます。



Local optima within top1000 grid-points

rank	deviance	weather	threshold	memory length	response mode	dose dependency	type of G	peak or start time of G	open length of G
1	260.32	temperature	10.0	270	> th	dose dependent	rect.	3	9
6	299.45	temperature	10.0	90	> th	dose dependent	rect.	19	18
13	308.61	humidity	50.0	270	> th	dose independent	rect.	17	18

補足 (2023 年 5 月 31 日追加)

イネのゲノムアノテーションに採用されている遺伝子座 (locus) には、以下の 2 つの DB が提供する 2 種類の id があります。

- RAP-DB (<https://rapdb.dna.affrc.go.jp/>)
- MSU Rice Genome Annotation Project (RGAP) Database (<http://rice.uga.edu/>)

前者は日本の農研機構が、後者はアメリカのミシガン州立大学が提供しています。Fit-DB で利用する locus id は、前者の id のみです。従って、イネの遺伝子を Fit-DB で検索するためには、例えば以下に示すようなステップを踏むことに

なります。

1. まず、論文等を読み、興味のある遺伝子を見つける。
2. 論文中に記載されている遺伝子座 (locus id)を見つける。
3. Fit-DB で採用している locus id であることを確認する。
4. Fit-DB の Search by locus id のボックスに入力して submit する。

論文中には、遺伝子名だけではなく locus id も記載されています。つまり、論文を読むと locus id がわかります。Fit-DB では、RAP-DB の locus id が採用されていますので、RAP-DB の locus id であるかを確認せねばなりません。RAP-DB の locus id は Os から始まり、MSU の locus id は LOC_Os から始まります。したがって、論文に記載してある id がどちらから始まっているかを確認する、ということになります。Os から始まる locus id (例題 7.2 にある Os01g0971800 など) であれば、Fit-DB で利用することができます。

謝辞

Fit-DB は農研機構で開発されたデータベースです。