JWAS:

Just another Whole-genome Analysis Software

Hao Cheng, Rohan Fernando, Dorian Garrick
JWAS
Julia for Whole-genome Analysis Software

**Bayesian Regression**
- MCMC
- Bayesian Alphabet

**Basic Analysis**
- Linear Mixed Models
- Multivariate Analysis

**Genomic Data**
- Genomic Prediction
- Genome-wide Association Studies

Julia  Linear Mixed Models  Multivariate  Univariate  Missing Phenotypes
Whole-genome Data  Bayesian Alphabet  Pedigree  Incomplete Genomic Data
1. completely open-source
2. routine data analysis
3. "reproducible research"
   • participate, maintain, extend
## Bayesian Linear Mixed Models (BLMM)

<table>
<thead>
<tr>
<th>BLMM</th>
<th>Marker Effect Model</th>
<th>Breeding Value Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>incomplete Genomic Data</strong></td>
<td>Non-Mixture Priors</td>
<td>RR-BLUP (^{ss\ mt})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BayesA (^{ss\ mt})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>Mixture Priors</td>
<td>BayesB (^{ss\ mt})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BayesC(\pi) (^{ss\ mt})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BayesR (^{ss\ mt})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Genomic BLUP (^{ss\ mt})</td>
</tr>
</tbody>
</table>
user interface of JWAS

- In cell 2 below, the JWAS package is loaded, as well as the DataFrames package for reading text files.

```python
In [2]:
1 using JWAS, DataFrames
```

- In cell 3 below, the data and pedigree information is loaded on line 1-2. The first several rows of data is shown on line 3.

```python
In [3]:
1 data = readable("phenotypes.txt")
2 pedigree = get_pedigree("pedigree.txt");
3 head(data)
```

Finished!

```
Out[3]:
   animal trait1  trait2 age sex  litter  dam
 0     O1  150.0  13.0   3  M     b    D1
 1     O3  40.0  5.0   4  F     a    D1
```
Out[3]:

<table>
<thead>
<tr>
<th>animal</th>
<th>trait1</th>
<th>trait2</th>
<th>age</th>
<th>sex</th>
<th>litter</th>
<th>dam</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1</td>
<td>150.0</td>
<td>13.0</td>
<td>3</td>
<td>M</td>
<td>b</td>
<td>D1</td>
</tr>
<tr>
<td>O3</td>
<td>40.0</td>
<td>5.0</td>
<td>4</td>
<td>F</td>
<td>a</td>
<td>D1</td>
</tr>
</tbody>
</table>

- In cell 4 below, the non-genomic part of the model equation for a 2-trait analysis is defined.
  - The effects fitted for trait 1 are the intercept, sex, direct genetic effects and maternal genetic effects.
  - The effects fitted for trait 2 are the intercept, sex, age, the interaction between sex and age and direct genetic effects.

In [4]:

```
model_equations = "trait1 = intercept + sex + litter + animal + dam
trait2 = intercept + sex + age + age*sex + animal";
```
In [4]:

```python
model_equations = "trait1 = intercept + sex + litter + animal + dam
                   trait2 = intercept + sex + age + age*sex + animal";
```

- In **cell 5** below, the model is built given the model equation in cell 3 and residual variance R. **By default, all effects in model are fixed and factors.** On line 2, the effect **age** is defined as covariates. On line 3, the effect **litter** is defined as random with variance **G1**. On line 4, direct genetic effects and maternal genetic effects are fitted as **animal** and **dam** with variance **G2** and numerator relationship matrix from **pedigree**. (Codes to define G1,G2, G3, R are skipped for demonstration purpose)

In [5]:

```python
model = build_model(model_equations,R)
set_covariate(model,"age")
set_random(model,"litter",G1)
set_random(model,"animal dam", pedigree,G2)
```
- In **cell 6** below, genomic part of the model is defined with genotypes file and variance **G3**. The format of genotypes file is:

  01,1,2,0,1,0  
  03,0,0,2,0,1

**In [6]:**

```
add_genotype(model,"genotypes.txt",G3)
```

- In **cell 7** below, a multi-trait BayesC analysis is performed with **model** and **data** defined in cell 1 to 5.

**In [7]:**

```
runMCMC(model,data,methods="BayesC")
```

More information is available [here](#). For help on JWAS, type `?JWAS` and press enter.
## Future plan

<table>
<thead>
<tr>
<th>BLMM</th>
<th>Marker Effect Model</th>
<th>Breeding Value Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Mixture Priors</td>
<td>RR-BLUP $\text{as mt}$</td>
<td>Gonomio BLUP $\text{as mt}$</td>
</tr>
<tr>
<td></td>
<td>BayesA $\text{as mt}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BayesB $\text{as mt}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BayesC $\text{as mt}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BayesD $\text{as mt}$</td>
<td></td>
</tr>
</tbody>
</table>

---

### Workshops

- **Run Julia from the Browser. No setup.**
  - The Julia community is doing amazing things. **We want you in on it!**

- **JuliaBox:**
  - **Julia**
    - Create Julia Notebooks and share them.
  - **Console**
    - Use in-browser terminal to fully control your Docker instance.
  - **Google Drive**
    - Collaborate with others.
  - **Sync & Share**
    - Setup folders to sync with remote git repositories.
Questions?

– http://QTL.rocks
– My group is Recruiting!