OmicABEL: Story of a successful interdisciplinary collaboration

Diego Fabregat-Traver and Prof. Paolo Bientinesi
In collaboration with Dr. Yurii Aulchenko

AICES, RWTH Aachen
`fabregat@aices.rwth-aachen.de`

PASC Conference, June 2nd – 3rd, 2014
Zürich, Switzerland

OmicABEL, GenABEL and GWAS

OmicABEL

- A library part of the GenABEL framework for statistical genomics
- High-performance solver(s) for (mixed-models) GWAS
- Growing (work force, functionality)

OmicABEL, GenABEL and GWAS

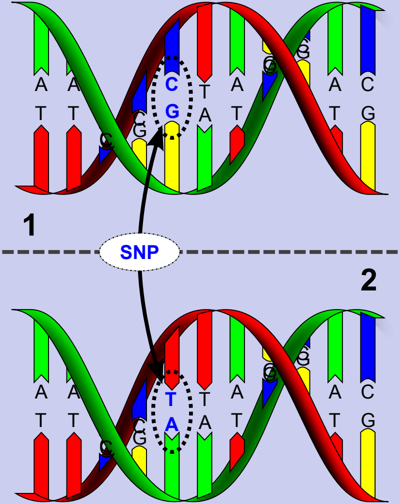
OmicABEL

- A library part of the GenABEL framework for statistical genomics
- High-performance solver(s) for (mixed-models) GWAS
- Growing (work force, functionality)

Genome-Wide Association Studies

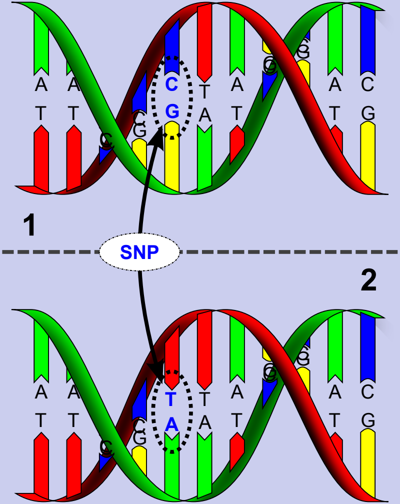
- Aim at identifying association between genetic markers and traits of interest
- Significant association highlights genomic regions involved in the control of a trait

SNP



Source: David Hall

SNP



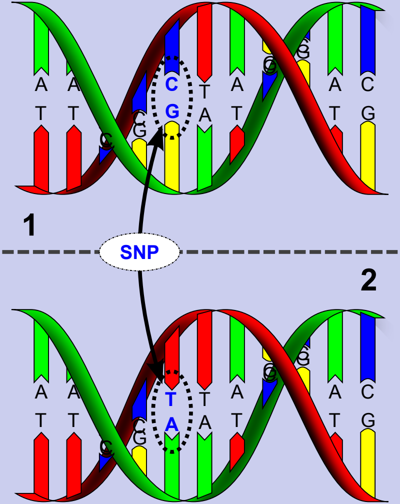
Source: David Hall

Traits

- Hair color, eye color.
- Complex traits: diseases such as coronary artery disease (CAD), . . .

GWAS

SNP



Source: David Hall

Traits

- Hair color, eye color.
- Complex traits: diseases such as coronary artery disease (CAD), . . .

Linear mixed-models

- Generalized Least-Squares: $b = (X^T M^{-1} X)^{-1} X^T M^{-1} y$

Computational problem

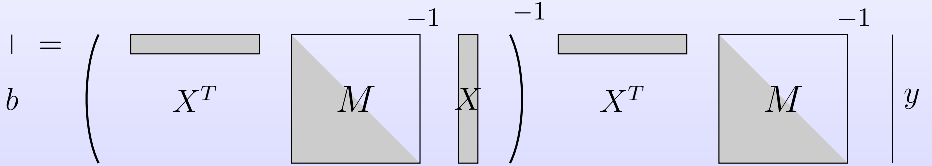
$$\begin{cases} b &= (X^T M^{-1} X)^{-1} X^T M^{-1} y \\ M &= \sigma^2 (h^2 \Phi + (1 - h^2) I) \end{cases}$$

$X \in R^{n \times p}$,	SNP
$y \in R^{n}$,	trait
$b \in R^{p}$,	genetic effect
$h^2, \sigma^2 \in R$,	heritability and res. variance
$\Phi \in R^{n \times n}$,	kinship matrix

- $n \in [1{,}000, ..., 10{,}000]$
- $p \in [1, ..., 20]$

Computational problem

$$\begin{cases} b & = (X^T M^{-1} X)^{-1} X^T M^{-1} y \\ M & = \sigma^2 (h^2 \Phi + (1 - h^2) I) \end{cases}$$

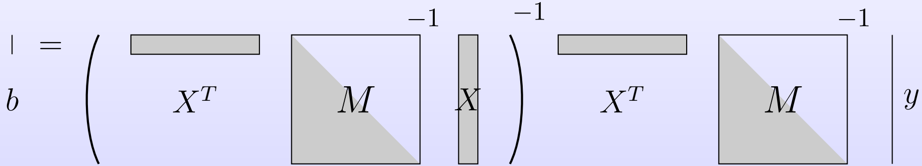


- $X \in R^{n \times p}$, SNP
- $y \in R^n$, trait
- $b \in R^p$, genetic effect
- $h^2, \sigma^2 \in R$, heritability and res. variance
- $\Phi \in R^{n \times n}$, kinship matrix

- $n \in [1{,}000, ..., 10{,}000]$
- $p \in [1, ..., 20]$

Computational problem

$$\begin{cases} b_i &= ({X_i}^T M^{-1} X_i)^{-1} {X_i}^T M^{-1} y \\ M &= \sigma^2(h^2\Phi + (1-h^2)I) \end{cases} \quad \text{with } 1 \leq i \leq m$$

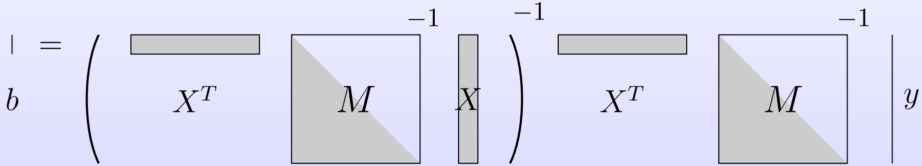


- $X \in R^{n \times p}$, SNP
- $y \in R^n$, trait
- $b \in R^p$, genetic effect
- $h^2, \sigma^2 \in R$, heritability and res. variance
- $\Phi \in R^{n \times n}$, kinship matrix

- $n \in [1{,}000, ..., 10{,}000]$
- $p \in [1, ..., 20]$
- $m \in [10^6, ..., 10^7]$

Computational problem

$$\begin{cases} b_{ij} = (X_i{}^T M_j{}^{-1} X_i)^{-1} X_i{}^T M_j{}^{-1} y_j & \text{with } 1 \le i \le m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \le j \le t. \end{cases}$$



- $X \in R^{n \times p}$, SNP
- $y \in R^n$, trait
- $b \in R^p$, genetic effect
- $h^2, \sigma^2 \in R$, heritability and res. variance
- $\Phi \in R^{n \times n}$, kinship matrix

- $n \in [1{,}000, ..., 10{,}000]$
- $p \in [1, ..., 20]$
- $m \in [10^6, ..., 10^7]$
- t is 1 or $\in [10^3, ..., 10^5]$

The challenge

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) \end{cases} \quad \begin{array}{l} \text{with } 1 \leq i \leq m \\ \text{and } 1 \leq j \leq t. \end{array}$$

Back in 2010...

- $n = 1{,}500$, $p = 4$, $m = 220$k: 4 hours.
- $n = 1{,}500$, $p = 4$, $m = 2.5$m: ≈43 hours.

The challenge

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j & \text{with } 1 \leq i \leq m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \leq j \leq t. \end{cases}$$

Scenario 1

- n: 10,000
- p: 4
- m: 36,000,000
- y: 1
- Data set: ≈ 3 TB

The challenge

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) \end{cases} \quad \begin{array}{l} \text{with } 1 \leq i \leq m \\ \text{and } 1 \leq j \leq t. \end{array}$$

Scenario 1

- n: 10,000
- p: 4
- m: 36,000,000
- y: 1
- Data set: ≈ 3 TB

Tool	Time
GWFGLS	20 days
FaST-LMM	20 hours

The challenge

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j & \text{with } 1 \leq i \leq m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \leq j \leq t. \end{cases}$$

Scenario 1

- n: $10{,}000$
- p: 4
- m: $36{,}000{,}000$
- y: 1

- Data set: ≈ 3 TB

Tool	**Time**
GWFGLS	20 days
FaST-LMM	20 hours

Scenario 2

- n: $1{,}000$
- p: 4
- m: $1{,}000{,}000$
- t: $100{,}000$

- Data set: ≈ 3 TB

The challenge

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j & \text{with } 1 \leq i \leq m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \leq j \leq t. \end{cases}$$

Scenario 1

- n: 10,000
- p: 4
- m: 36,000,000
- y: 1
- Data set: ≈ 3 TB

Tool	Time
GWFGLS	20 days
FaST-LMM	20 hours

Scenario 2

- n: 1,000
- p: 4
- m: 1,000,000
- t: 100,000
- Data set: ≈ 3 TB

Tool	Time
GWFGLS	1.5 years
FaST-LMM	7.5 months

The challenge

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j & \text{with } 1 \le i \le m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \le j \le t. \end{cases}$$

Scenario 1

- n: 10,000
- p: 4
- m: 36,000,000
- y: 1

- Data set: ≈ 3 TB

Tool	Time
GWFGLS	20 days
FaST-LMM	20 hours
OmicABEL-Chol	6 hours

Scenario 2

- n: 1,000
- p: 4
- m: 1,000,000
- t: 100,000

- Data set: ≈ 3 TB

Tool	Time
GWFGLS	1.5 years
FaST-LMM	7.5 months
OmicABEL-Eig	12 hours

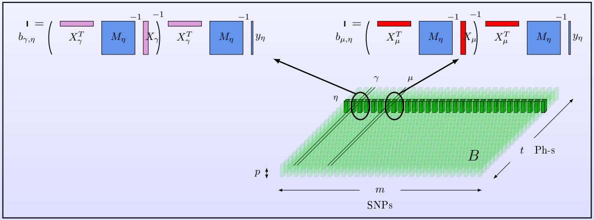
Systematic/Engineering approach

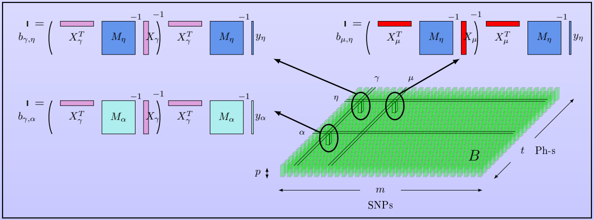
How to attain OmicABEL-EIG's speedup

- Complexity: consider the problem in its entirety
- Data management: analyze data movement to reduce I/O impact (tiling)
- Performance and Scalability: blocking + hybrid parallelism (MT-BLAS + OpenMP)

Reducing complexity

Consider the grid in its entirety





Reducing complexity

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j & \text{with } 1 \leq i \leq m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \leq j \leq t. \end{cases}$$

$$\begin{aligned} M_j &= \sigma_j (h_j \Phi + (1 - h_j) I) \\ &\downarrow \alpha_j := \sigma_j * h_j, \beta_j := \sigma_j * (1 - h_j) \\ M_j &= \alpha_j \Phi + \beta_j I \end{aligned}$$

Reducing complexity

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j & \text{with } 1 \le i \le m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \le j \le t. \end{cases}$$

$$\begin{aligned} M_j &= \sigma_j (h_j \Phi + (1 - h_j) I) \\ &\downarrow \alpha_j := \sigma_j * h_j, \beta_j := \sigma_j * (1 - h_j) \\ M_j &= \alpha_j \Phi + \beta_j I \end{aligned}$$

$$\begin{aligned} \Phi &= Z W Z^T \\ M_j &= Z(\alpha_j W + \beta_j I) Z^T \end{aligned}$$

Reducing complexity

$$\begin{cases} b_{ij} = (X_i^T M_j^{-1} X_i)^{-1} X_i^T M_j^{-1} y_j & \text{with } 1 \leq i \leq m \\ M_j = \sigma_j^2 (h_j^2 \Phi + (1 - h_j^2) I) & \text{and } 1 \leq j \leq t. \end{cases}$$

$$M_j = \sigma_j (h_j \Phi + (1 - h_j) I)$$

$$\downarrow \alpha_j := \sigma_j * h_j, \beta_j := \sigma_j * (1 - h_j)$$

$$M_j = \alpha_j \Phi + \beta_j I$$

$$\Phi = Z W Z^T$$

$$M_j = Z(\alpha_j W + \beta_j I) Z^T$$

$$M_j^{-1} = Z(\alpha_j W + \beta_j I)^{-1} Z^T$$

$$b_{ij} = (X_i^T Z D_j^{-1} Z^T X_i)^{-1} X_i^T Z D_j^{-1} Z^T y_j$$

$$b_{ij} = ({X'}_i^T D_j^{-1} {X'}_i)^{-1} {X'}_i^T D_j^{-1} y'_j$$

OmicABEL-EIG algorithm.

```
$ZWZ^T = \Phi$
for 1 ≤ i ≤ m
    $X'_i := Z^T X_i$
for 1 ≤ j ≤ t
    $y'_j := Z^T y_j$
for 1 ≤ j ≤ t
    $D_j := \sigma_j^2 (h_j^2 W + (1 - h_j^2) I)$
    $K_j K_j^T = D_j^{-\frac{1}{2}}$
    $v_j := K_j^T y'_j$
    for 1 ≤ i ≤ m
        $W_{ij} := K_j^T X'_i$
        $S_{ij} := W_{ij}^T W_{ij}$
        $b_{ij} := W_{ij}^T v_j$
        $b_{ij} := S_{ij}^{-1} b_{ij}$
```

Reducing complexity

Approach	**1 GLS**	**Single-trait**	**Multi-trait**	**Flops**
Naive	$O(n^3)$	$O(mn^3)$	$O(tmn^3)$	$O(10^{20})$ - 100 ExaF
GWFGLS		$O(mn^2)$	$O(tmn^2)$	$O(10^{17})$ - 100 PetaF
Omic-EIG			$O(tmn)$	$O(10^{14})$ - 100 TeraF

Setup: $n = 1{,}000, m = 1{,}000{,}000$, $t = 100{,}000$.

Data management

Naive data streaming

```
[...]
for 1 ≤ j ≤ t
    load $y_j$
    [...]
    for 1 ≤ i ≤ m
        load $X_i$
        [...]
        store $b_{ij}$
```

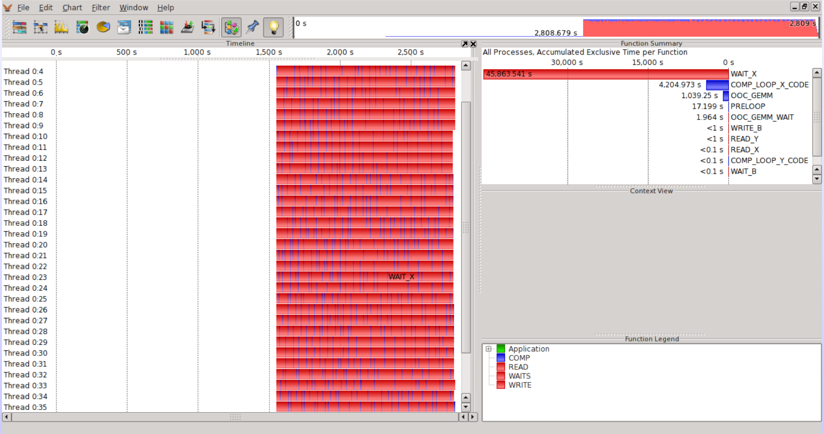
Data management

Naive data streaming

```
[...]
for 1 ≤ j ≤ t
    load $y_j$
    [...]
    for 1 ≤ i ≤ m
        load $X_i$
        [...]
        store $b_{ij}$
```

Loads $m \times t$ times a vector X_i (of size n): 10^{14} doubles $\approx$ **1PB**.

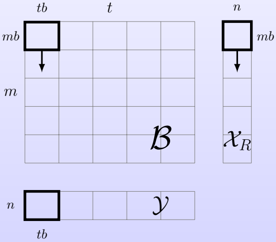
Data management



Data management

Tiling

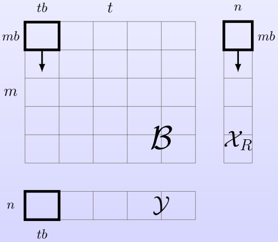
```
[...]
for 1 ≤ j ≤ $t/t_b$
   load block of $y_j$'s
   [...]
   for 1 ≤ i ≤ $m/m_b$
      load block of $X_i$'s
      [...]
      for 1 ≤ j ≤ $t_b$
         for 1 ≤ i ≤ $m_b$
            [...]
      store block of $b_{ij}$'s
```



Data management

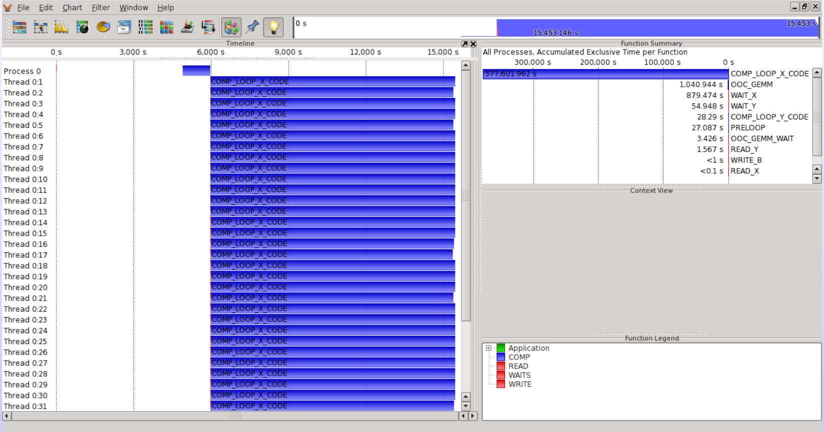
Tiling

```
[...]
for 1 ≤ j ≤ $t/t_b$
  load block of $y_j$'s
  [...]
  for 1 ≤ i ≤ $m/m_b$
    load block of $X_i$'s
    [...]
    for 1 ≤ j ≤ $t_b$
      for 1 ≤ i ≤ $m_b$
        [...]
    store block of $b_{ij}$'s
```



+ Overlap of I/O with computation!

Data management



All Processes, Accumulated Exclusive Time per Function

Time	Function
377,601.962 s	COMP_LOOP_X_CODE
1,040.944 s	OOC_GEMM
879.474 s	WAIT_X
54.948 s	WAIT_Y
28.29 s	COMP_LOOP_Y_CODE
27.087 s	PRELOOP
3.426 s	OOC_GEMM_WAIT
1.567 s	READ_Y
<1 s	WRITE_B
<0.1 s	READ_X

- 🟩 Application
- 🟦 COMP
- 🟥 READ
- 🟥 WAITS
- 🟥 WRITE

Performance and scalability

GEMV-based

```
for 1 ≤ i ≤ m
    $X_i' := Z^T X_i$ (GEMM)
for 1 ≤ j ≤ t
    $y_j' := Z^T y_j$ (GEMV)
for 1 ≤ j ≤ t
    [...]
    for 1 ≤ i ≤ m
        [...]
```

GEMM-based

```
$\mathcal{X}' := Z^T \mathcal{X}$ (GEMM)
$\mathcal{Y}' := Z^T \mathcal{Y}$ (GEMM)
for 1 ≤ j ≤ t
    [...]
    for 1 ≤ i ≤ m
        [...]
```

Multiple mat-vec (GEMV) as a single mat-mat (GEMM).

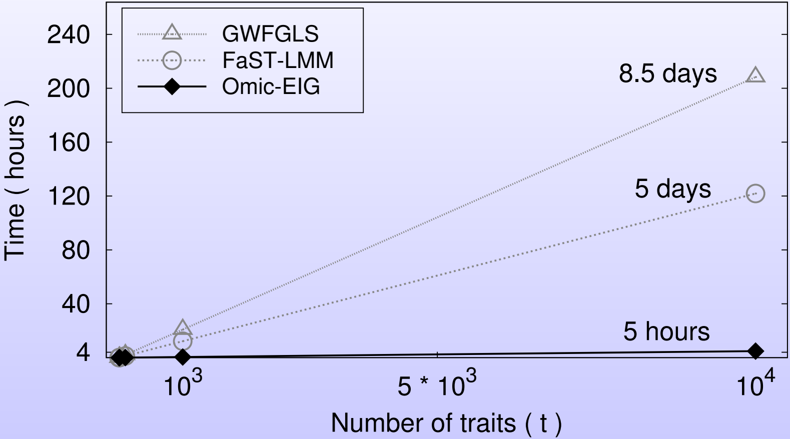
Performance and scalability

Blocking innermost loops

```
[...]
for 1 ≤ j ≤ t
    [...]
    for 1 ≤ i ≤ m
        $W_{ij} := K_j^T X'_i$     (SCAL)
        $S_{ij} := W_{ij}^T W_{ij}$  (GEMM)
        $b_{ij} := W_{ij}^T v_j$     (GEMV)
        $b_{ij} := S_{ij}^{-1} b_{ij}$  (POSV)
```

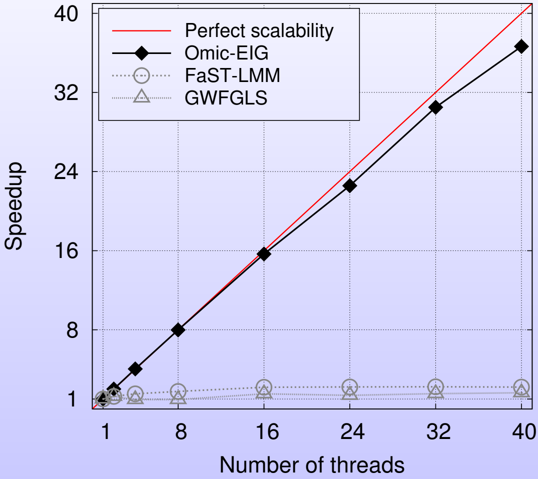
- Memory bound operations, but compute bound as a block
- Blocking increases performance
- Explicit parallelism with OpenMP

Experimental results for challenge *Scenario 2*



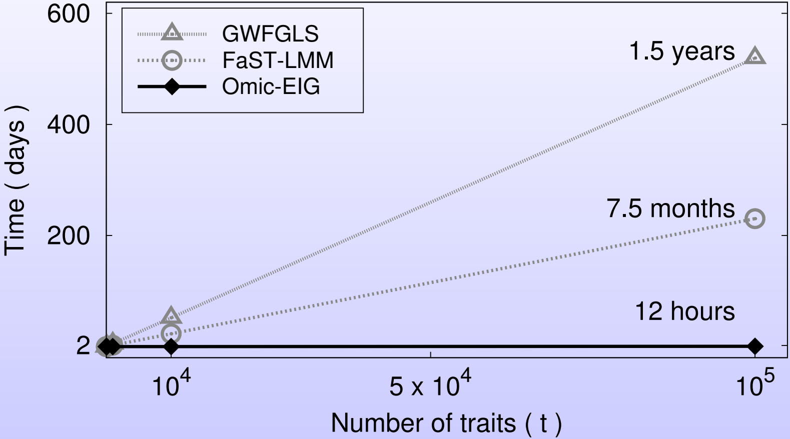
Setup: $n = 1{,}000, p = 4, m = 100{,}000$. Single core.

Experimental results for challenge *Scenario 2*



Setup: $n = 1{,}000, p = 4, m = 100{,}000, t = 20{,}000$.

Experimental results for challenge *Scenario 2*



Setup: $n = 1{,}000, p = 4, m = 1{,}000{,}000$. Full 40-core node.

Interdisciplinary collaboration hurdles

Remarkable results...

... at the expense of a long development time

Interdisciplinary collaboration hurdles

Remarkable results...

... at the expense of a long development time

Why?

- Initial overhead in finding a common language

Interdisciplinary collaboration hurdles

Remarkable results...

... at the expense of a long development time

Why?

- Initial overhead in finding a common language
- Continuous refinement of the problem definition

Interdisciplinary collaboration hurdles

Remarkable results...

... at the expense of a long development time

Why?

- Initial overhead in finding a common language
- Continuous refinement of the problem definition
- Different priorities

It is worth it!

Once the initial overhead is overcome...

- MPI: single- and multi-trait (Elmar Peise)
- GPU: single trait (Lucas Beyer)
- Related problem (OLS) (Alvaro Frank)

- OmicABEL + OLS-based used by a third research group
- Joint proposals, PhD

It is worth it!

Once the initial overhead is overcome...

- MPI: single- and multi-trait (Elmar Peise)
- GPU: single trait (Lucas Beyer)
- Related problem (OLS) (Alvaro Frank)

- OmicABEL + OLS-based used by a third research group
- Joint proposals, PhD

In a broader sense

- There is a gap between applications and HPC

It is worth it!

Once the initial overhead is overcome...

- MPI: single- and multi-trait (Elmar Peise)
- GPU: single trait (Lucas Beyer)
- Related problem (OLS) (Alvaro Frank)

- OmicABEL + OLS-based used by a third research group
- Joint proposals, PhD

In a broader sense

- There is a gap between applications and HPC
- If...
  - Application experts are patient and get involved
  - HPC experts focus on application's priorities
- Then, experience shows that results can be remarkable

Links:

- `http://hpac.rwth-aachen.de/`
- `http://www.genabel.org/packages/OmicABEL`
- Computing Petaflops over Terabytes of Data: The Case of Genome-Wide Association Studies. *Fabregat et al.* ACM TOMS. In press.

Financial support from the **Deutsche Forschungsgemeinschaft** (German Research Association) through grant GSC 111 is gratefully acknowledged.

Deutsche
Forschungsgemeinschaft
DFG