Medical Resequencing: Future Innovations and Sequence-Based Association Analysis

Genotype-Based Warfarin Dose Prediction

Mark Rieder
Department of Genome Sciences
mrieder@u.washington.edu

Pharmacogenomics as a Model for Medical ReSequencing

- Clear genotype-phenotype link
  - intervention → variable response
- Pharmacokinetics - 5x variation
- Quantitative intervention and response
  - drug dose, response time, metabolism rate, etc.
- Target/metabolism of drug generally known
  - gene target that can be tested directly with response
- Prospective testing reduce variability and identify outliers

Medical Resequencing

- Discovery of rare functional variants -
  - Sequencing at the tails of the distribution
- Testing the Common Disease Common Variant (CDCV) hypothesis
  - Candidate genes very feasible
- Whole Genome Association Studies (WGAS)
- Whole Genome Sequencing - Developing Rapidly

Warfarin Pharmacogenetics

1. Background
   - Vitamin K cycle
   - Pharmacokinetics/Pharmacodynamics
   - Discovery of VKORC1
2. VKORC1 - SNP Discovery
3. VKORC1 - SNP Selection (tagSNPs)
4. Clinical Association Study
   - VKORC1 and Warfarin Dose
5. VKORC1 - SNP Replication/Function
Warfarin Dosing - Background

• Commonly prescribed oral anti-coagulant and acts as an inhibitor of the vitamin K cycle
• Prescribed following MI, atrial fibrillation, stroke, venous thrombosis, prosthetic heart valve replacement, and following major surgery
• Warfarin (Coumadin) >20 million US prescriptions (2007)

(-) Difficult to determine effective dosage
- Narrow therapeutic range
- Large inter-individual variation

(-) Major bleeding episodes in 1-2% of all patients
11% of all adverse events (Gurwitz et al. JAMA 2003)

(++) Prevents 20 strokes for each bleeding event

Goal: Use genetics to understand dose requirements to reduce complications

Monitoring Warfarin Dosing

• Measure prothrombin time (PT) - time to clot
Normal times are 10-13 sec

• Also measured as INR (International Normalized Ratio). Range 1.0 - 1.4 (normalized thromboplastin)

• On warfarin therapy, a patient is kept at about 2-3x normal (PT = 20-39 seconds)

<table>
<thead>
<tr>
<th>Target INR Range</th>
<th>Target</th>
<th>Target Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep Ven Thrombosis/PE</td>
<td>2.5</td>
<td>2.0-3.0</td>
</tr>
<tr>
<td>CVL clot prophylaxis</td>
<td>1.5</td>
<td>1.2-2.0</td>
</tr>
<tr>
<td>Mechanical Valves</td>
<td>3.0</td>
<td>2.5-3.5</td>
</tr>
<tr>
<td>Cardiomyopathy</td>
<td>2.5</td>
<td>2.0-3.0</td>
</tr>
<tr>
<td>Ischemic Stroke</td>
<td>2.0</td>
<td>1.5-2.0</td>
</tr>
</tbody>
</table>

Ave: 5.2 mg/d
n = 186
European-American
30x dose variability

✓ Patient/Clinical/Environmenal Factors
✓ Pharmacokinetic/Pharmacodynamic - Genetic
**Vitamin K Cycle**

- Vitamin K synthesized by plants and bacteria
  - e.g., leafy green vegetables and intestinal flora
- Vitamin K - discovered from defects in blood "koagulation"
- Vitamin K - required coenzyme for the conversion of glutamic acid (Glu) to carboxyglutamic acid (Gla)
- **Glu** $\rightarrow$ **Gla** modification needed for Ca$^{2+}$ binding, clot formation
- Vitamin K administration is the antidote for warfarin toxicity

**Warfarin drug metabolism**

- Major pathway for termination of pharmacologic effect is through metabolism of S-warfarin in the liver by CYP2C9
- CYP2C9 SNPs alter warfarin metabolism:
  - CYP2C9*1 (WT) - normal
  - CYP2C9*2 (Arg144Cys) - intermediate metabolizer
  - CYP2C9*3 (Ile359Leu) - poor metabolizer
- CYP2C9 alleles occur at:
  - European: *2 - 10.7% *3 - 8.5%
  - Asian: *2 - 0% *3 - 1-2%
  - African-American: *2 - 2.9% *3 - 0.8%

**Effect of CYP2C9 Genotype on Anticoagulation-Related Outcomes**

(Higashi et al., JAMA 2002)

**Warfarin inhibits the vitamin K cycle**

- Warfarin inhibits the vitamin K cycle via CYP2C9
- Epoxide Reductase
- Vitamin K-dependent clotting factors (FII, FVII, FIX, FX, Protein C/S/Z)
- Vitamin K administration is the antidote for warfarin toxicity

**Effect of CYP2C9 Genotype on Anticoagulation-Related Outcomes**

(Higashi et al., JAMA 2002)

- Still large variability in warfarin dose (15-fold) in *1/*1 "controls"?

**TIME TO STABLE ANTICOAGULATION**

- CYP2C9-WT ~90 days
- CYP2C9-Variant ~180 days

- Variant alleles have significant clinical impact
- *2 or *3 carriers take longer to reach stable antiocoagulation
Warfarin acts as a vitamin K antagonist

Vitamin K-dependent clotting factors (FII, FVII, FIX, FX, Protein C/S/Z)

Epoxide Reductase
γ-Carboxylase (GGCX)

Warfarin acts as a vitamin K antagonist

Inactivation

CYP2C9

γ-Carboxylase (GGCX)

Clotting Factors (FII, FVII, FIX, FX, Protein C/S/Z)

Warfarin resistance VKORC1 polymorphisms

• Rare non-synonymous mutations in VKORC1 causative for warfarin resistance (15-35 mg/d)
• No non-synonymous mutations found in 'control' chromosomes (n ~400)

Inter-Individual Variability in Warfarin Dose: Genetic Liabilities

SENSITIVITY
CYP2C9 coding SNPs: *3/*3
Common VKORC1 non-coding SNPs?

RESISTANCE
VKORC1 nonsynonymous coding SNPs

Warfarin maintenance dose (mg/day)

5 kb - Chromosome 16

New Target Protein for Warfarin

SNP Discovery: Resequencing VKORC1

- PCR amplicons → Resequencing of the complete genomic region
- 5 Kb upstream and each of the 3 exons and intronic segments: ~11 Kb
- Warfarin treated clinical patients (UWMC): 186 European
- Other populations: 96 European, 96 African-Am., 120 Asian

Rieder et al, NEJM 352, 2285-2293, 2005

SNP Selection: VKORC1 tagSNPs

SNP Testing: VKORC1 tagSNPs

SNP Discovery: Resequencing Results

VKORC1 - PGA samples (European, n = 23)
Total: 13 SNPs identified
10 common/3 rare (<5% MAF)

VKORC1 - Clinical Samples (European patients n = 186)
Total: 28 SNPs identified
10 common/18 rare (<5% MAF)

15 - intronic/regulatory
7 - promoter SNPs
2 - 3' UTR SNPs
3 - synonymous SNPs
1 - nonsynonymous
  - single heterozygous indiv. - highest warfarin dose = 15.5 mg/d

None of the previously identified VKORC1 warfarin-resistance SNPs were present (Rost, et al.)

Do common SNPs associate with warfarin dose?
### SNP Testing: VKORC1 tagSNPs

**Five Bins to Test**

1. 381, 3673, 6484, 6853, 7566
   - Bin 1: \( p < 0.001 \) \( r^2 = 21\% \)
2. 2653, 6009
   - Bin 2: \( p < 0.02 \) \( r^2 = 3\% \)
3. 861
   - Bin 3: \( p < 0.01 \) \( r^2 = 4.5\% \)
4. 5808
   - Bin 4: \( p < 0.001 \) \( r^2 = 12\% \)
5. 9041
   - Bin 5: \( p < 0.001 \) \( r^2 = 11\% \)

Each tagSNP was adjusted for other confounders such as age, sex, medication, and CYP2C9.

### Multi-SNP testing: Haplotypes

**Five tagSNPs (10 total SNPs)**

- Adjusted for other confounders such as age, sex, medication, and CYP2C9.

### VKORC1 Haplotypes Associate with Dose

<table>
<thead>
<tr>
<th>Haplotype Identification Code</th>
<th>Haplotype Sequence</th>
<th>Frequency of Haplotype in Primary Patient Sample (n)</th>
<th>Average Maintenance Dose for Homozygous Patients (mg/d)</th>
<th>( p )-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>CCGATCTCTTG</td>
<td>0.12 (43)</td>
<td>2.9 (2.2 – 3.7)</td>
<td>0.0001</td>
</tr>
<tr>
<td>H2</td>
<td>CCGAGCTCTTG</td>
<td>0.24 (88)</td>
<td>3.0 (2.5 – 3.6)</td>
<td>0.001</td>
</tr>
<tr>
<td>H3</td>
<td>CCGGTCCCCCG</td>
<td>0.01 (2)</td>
<td>NA</td>
<td>NS</td>
</tr>
<tr>
<td>H4</td>
<td>CCGGTCCGTG</td>
<td>0.00 (1)</td>
<td>NA</td>
<td>NS</td>
</tr>
<tr>
<td>H5</td>
<td>TCGAGCTCTTG</td>
<td>0.00 (1)</td>
<td>NA</td>
<td>NS</td>
</tr>
<tr>
<td>H6</td>
<td>TCGGTCCCGCG</td>
<td>0.00 (0)</td>
<td>NA</td>
<td>NS</td>
</tr>
<tr>
<td>H7</td>
<td>TCGGGTTCCGCA</td>
<td>0.35 (132)</td>
<td>6.0 (5.2 – 6.9)</td>
<td>0.0001</td>
</tr>
<tr>
<td>H8</td>
<td>TAGGGTCCGCA</td>
<td>0.08 (28)</td>
<td>4.8 (3.4 – 6.7)</td>
<td>0.76</td>
</tr>
<tr>
<td>H9</td>
<td>TACGGTCCGCG</td>
<td>0.21 (77)</td>
<td>5.5 (4.5 – 6.7)</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Adjusted for all significant covariates: age, sex, amiodarone, CYP2C9 genotype

**25% variance in dose explained**

### Multi-SNP testing: Haplotypes

**Explore the evolutionary relationship across haplotypes**

- **VKORC1** haplotypes cluster into divergent clades

- Patients can be assigned a clade diplotype:
  - e.g., Patient 1 - H1/H2 = A/A
  - Patient 2 - H1/H7 = A/B
  - Patient 3 - H7/H9 = B/B
**VKORC1 clade diplotypes show a strong association with warfarin dose**

Univ. of Washington
n = 185

Washington University
n = 386

Brian Gage
Howard McCleod
Charles Eby

21% variance in dose explained

---

**SNP Function: VKORC1 Expression**

Expression in human liver tissue (n = 53) shows a graded change in expression.

---

**SNP Testing: VKORC1 tagSNPs**

**Five Bins to Test**

<table>
<thead>
<tr>
<th>Bin</th>
<th>p-value</th>
<th>r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;0.001</td>
<td>21%</td>
</tr>
<tr>
<td>2</td>
<td>&lt;0.02</td>
<td>3%</td>
</tr>
<tr>
<td>3</td>
<td>&lt;0.01</td>
<td>4.5%</td>
</tr>
<tr>
<td>4</td>
<td>&lt;0.001</td>
<td>12%</td>
</tr>
<tr>
<td>5</td>
<td>&lt;0.01</td>
<td>11%</td>
</tr>
</tbody>
</table>

1. 381, 3673, 6484, 6853, 7566
2. 2653, 6009
3. 861
4. 5808
5. 9041

---

**SNP Function: VKORC1 Expression**

Expression in human liver tissue (n = 53) shows a graded change in expression.

---

**SNP Testing: VKORC1 tagSNPs**

**Five Bins to Test**

<table>
<thead>
<tr>
<th>Bin</th>
<th>p-value</th>
<th>r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;0.001</td>
<td>21%</td>
</tr>
<tr>
<td>2</td>
<td>&lt;0.02</td>
<td>3%</td>
</tr>
<tr>
<td>3</td>
<td>&lt;0.01</td>
<td>4.5%</td>
</tr>
<tr>
<td>4</td>
<td>&lt;0.001</td>
<td>12%</td>
</tr>
<tr>
<td>5</td>
<td>&lt;0.01</td>
<td>11%</td>
</tr>
</tbody>
</table>

1. 381, 3673, 6484, 6853, 7566
2. 2653, 6009
3. 861
4. 5808
5. 9041

---

**SNP Function: VKORC1 Expression**

Expression in human liver tissue (n = 53) shows a graded change in expression.

---

**SNP Testing: VKORC1 tagSNPs**

**Five Bins to Test**

<table>
<thead>
<tr>
<th>Bin</th>
<th>p-value</th>
<th>r²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;0.001</td>
<td>21%</td>
</tr>
<tr>
<td>2</td>
<td>&lt;0.02</td>
<td>3%</td>
</tr>
<tr>
<td>3</td>
<td>&lt;0.01</td>
<td>4.5%</td>
</tr>
<tr>
<td>4</td>
<td>&lt;0.001</td>
<td>12%</td>
</tr>
<tr>
<td>5</td>
<td>&lt;0.01</td>
<td>11%</td>
</tr>
</tbody>
</table>

1. 381, 3673, 6484, 6853, 7566
2. 2653, 6009
3. 861
4. 5808
5. 9041

---
**General approach for candidate gene association study**

1. Establish baseline genetic diversity in a small discovery population (n=23) (e.g. find all common SNPs via resequencing)
2. Calculate correlation between SNPs to find informative SNPs (tagSNPs)
3. Genotype tagSNPs in clinical population (n=186)
4. Perform statistical test for association
   - Association Results: **YES** - Multivariate Regression
     - Determine effect size - variability in dose explained
     - Adjusted for all confounding factors (Age, Sex, VKORC1, CYP2C9)
   - Association Results: **NO** - Because all common SNPs have been tested, it is very likely no association exists

**GWAS approach for association studies**

1. Establish baseline genetic diversity in a discovery population (HapMap) (e.g. find all common SNPs)
2. Calculate correlation between SNPs to find informative SNPs (tagSNPs)
3. Genotype tagSNPs in population
   - Use commercially available whole genome chips (e.g. Illumina, Affy)
   - QC genotype data
4. Perform statistical test for association
   - Association Results: Multivariate Regression
     - Establish p-value cutoff (1E-7)
     - Replicate in similar populations
   - If replication then association can be considered
   - NO - Association Results: All common SNPs have been tested, it is very likely no association exists assuming sufficient power

**Clinical Adoption of Dosing Algorithms**

**NHLBI Clinical Warfarin Trial:**
- Randomized trial of prospective genotype-guided dosing
- Multi-center, double-blind randomized trial (n=2,000)
  - "standard of care" vs. clinical alg. vs. clinical + genetic alg.
  - multiple outcomes (e.g. time within therapeutic range)

```
Total warfarin variance:
```

**Design - WGAS Warfarin Dose**

550 K Illumina
181 samples tested (Higashi, et al., Rieder, et al.)

- Call rates 99%
- 100% concordance - VKORC1-3673 (rs9923231) with rs10871454 (LD)
- 100% concordance - CYP2C9*3 (rs1057910)

**Models Examined:**

**Independent SNP effects (univariate):** ln(dose) = SNP

**Full Model:** Genetic + Clinical (multivariate):
ln(dose) = Age + Sex + Amiodarone + Losartan + CYP2C9 (*2 or *3) + VKORC1-3673
Warfarin Dose - Detection Power

Ave = 5.2 ± 2.5 mg/d
Additive effect, quantitative trait
P = 1x10⁻⁷

Univariate Results - SNP Associations

CYP2C9
VKORC1

Univariate Additive - log(p-value)

Range for large effect genes for warfarin dosing

Large real effects (VKORC1)
Non-replicated SNPs

Warfarin Dose GWA Replication Results
**VKORC1 Pharmacogenetics Summary**

1. *VKORC1* haplotypes (tagSNPs) are the major contributor to warfarin dose variability (21–25%). Overall variance described by clinical and genetic factors is 50-60%.

2. *VKORC1* haplotypes are correlated with mRNA expression in the liver.

3. Whole genome studies have power to detect large effect size (20–25% variance) with limited sample size. Power is limited to detect small/moderate effects.

4. No other SNPs have a large effect similar to *VKORC1*.

5. Prospective clinical trials in progress should be definitive.

6. Prospective genotyping may lead to more accurate warfarin dosing and have impacts on the overall clinical treatment time.

**Acknowledgments**

Greg Cooper
Allan Rettie
Alex Reiner
Dave Veenstra
Dave Blough
Ken Thummel
Debbie Nickerson

Josh Smith
Chris Baier
Peggy Dyer-Robertson

**Washington University**
Brian Gage
Howard McLeod
Charles Eby

**Replication Studies**
Julie Johnson, UF
Dan Roden, VU