

The structural origin of metabolic quantitative diversity

Seizo Koshiha, Ikuko Motoike, Kaname Kojima, Takanori Hasegawa, Matsuyuki Shirota, Tomo Saito, Daisuke Saigusa, Inaho Danjoh, Fumiki Katsuoka, Soichi Ogishima, Yosuke Kawai, Yumi Yamaguchi-Kabata, Miyuki Sakurai, Sachiko Hirano, Junichi Nakata, Hozumi Motohashi, Atsushi Hozawa, Shinichi Kuriyama, Naoko Minegishi, Masao Nagasaki, Takako Takai-Igarashi, Nobuo Fuse, Hideyasu Kiyomoto, Junichi Sugawara, Yoichi Suzuki, Shigeo Kure, Nobuo Yaegashi, Osamu Tanabe, Kengo Kinoshita, Jun Yasuda, & Masayuki Yamamoto

Supplementary Materials

CONTENTS

SUPPLEMENTARY TABLE

Supplementary Table S1.

Supplementary Table S2.

Supplementary Table S3.

SUPPLEMENTARY FIGURE

Supplementary Figure S1.

Supplementary Figure S2.

Supplementary Figure S3.

Supplementary Figure S4.

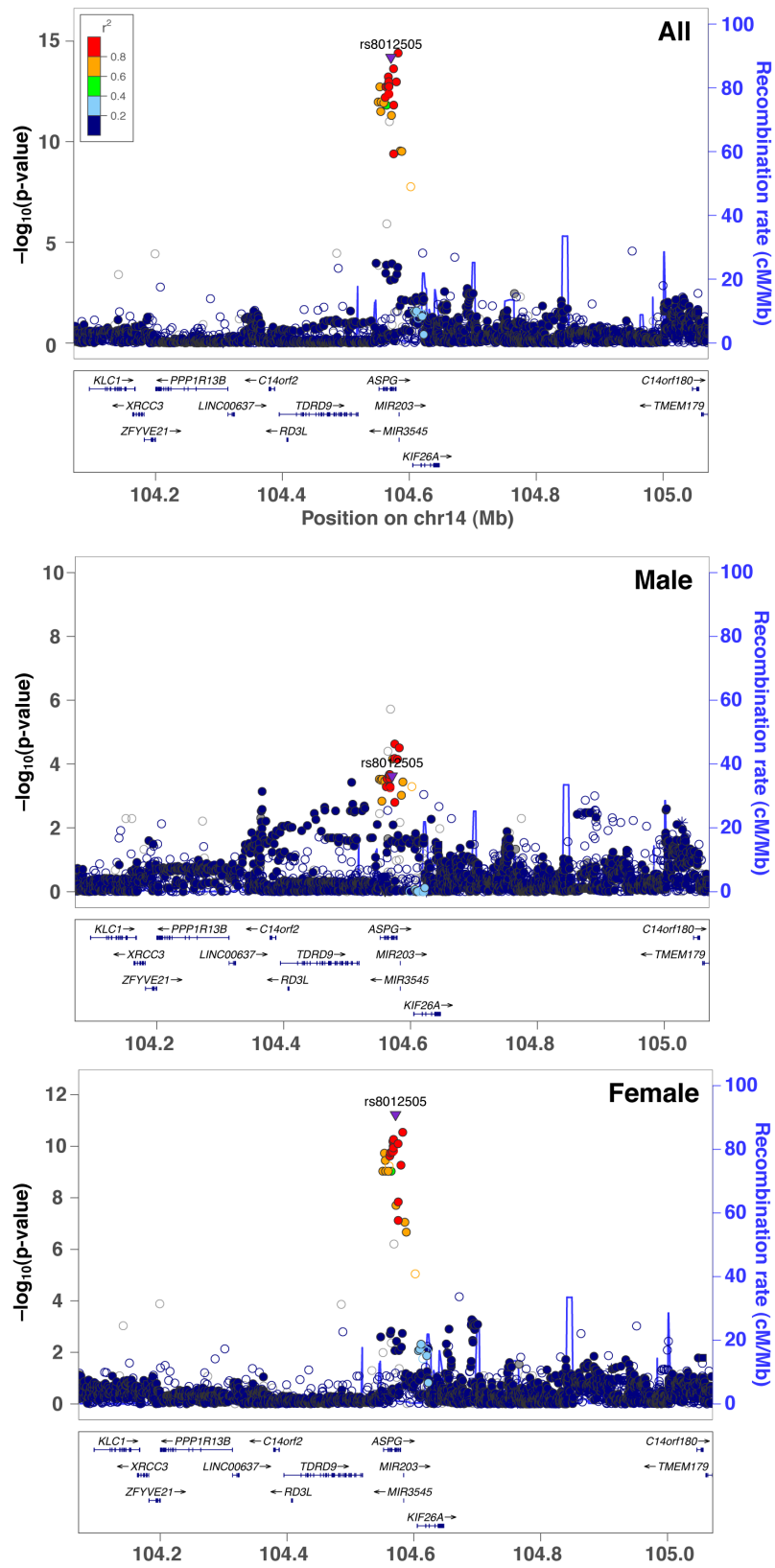
Supplementary Table S1: Plasma metabolites identified and quantified in this study

| Metabolite | mean \pm s.d. (μM) | Metabolite | mean \pm s.d. (μM) |
|------------------------|---|-------------------|---|
| 2-Aminobutyrate | 16.7 \pm 5.7 | Glycerol | 78.4 \pm 29.0 |
| 2-Hydroxybutyrate | 34.7 \pm 15.8 | Glycine | 220.2 \pm 68.4 |
| 2-Oxoisocaproate | 24.4 \pm 6.2 | Histidine | 79.1 \pm 10.6 |
| 3-Hydroxybutyrate | 95.2 \pm 125.0 | Isoleucine | 57.9 \pm 17.7 |
| 3-Hydroxyisobutyrate | 12.1 \pm 3.8 | Lactate | 2222.7 \pm 605.6 |
| 3-Methyl-2-oxovalerate | 18.4 \pm 4.3 | Leucine | 109.3 \pm 25.8 |
| Acetate | 37.3 \pm 9.0 | Lysine | 119.9 \pm 23.7 |
| Acetone | 7.9 \pm 7.0 | Methionine | 23.5 \pm 6.1 |
| Alanine | 331.8 \pm 80.7 | Ornithine | 63.2 \pm 15.8 |
| Arginine | 62.2 \pm 15.6 | Phenylalanine | 58.2 \pm 10.5 |
| Asparagine | 56.5 \pm 11.5 | Proline | 158.5 \pm 48.4 |
| Carnitine | 49.6 \pm 10.3 | Pyruvate | 73.8 \pm 17.1 |
| Creatine | 33.3 \pm 17.3 | Serine | 114.7 \pm 21.5 |
| Creatinine | 55.1 \pm 19.6 | Succinate | 6.9 \pm 1.3 |
| Cysteine | 52.2 \pm 17.3 | Threonine | 132.3 \pm 29.3 |
| Formate | 16.6 \pm 3.3 | Tryptophan | 49.0 \pm 9.1 |
| Glucose | 4844.2 \pm 994.2 | Tyrosine | 64.8 \pm 14.9 |
| Glutamate | 53.6 \pm 16.6 | Valine | 209.2 \pm 43.5 |
| Glutamine | 528.6 \pm 65.4 | | |

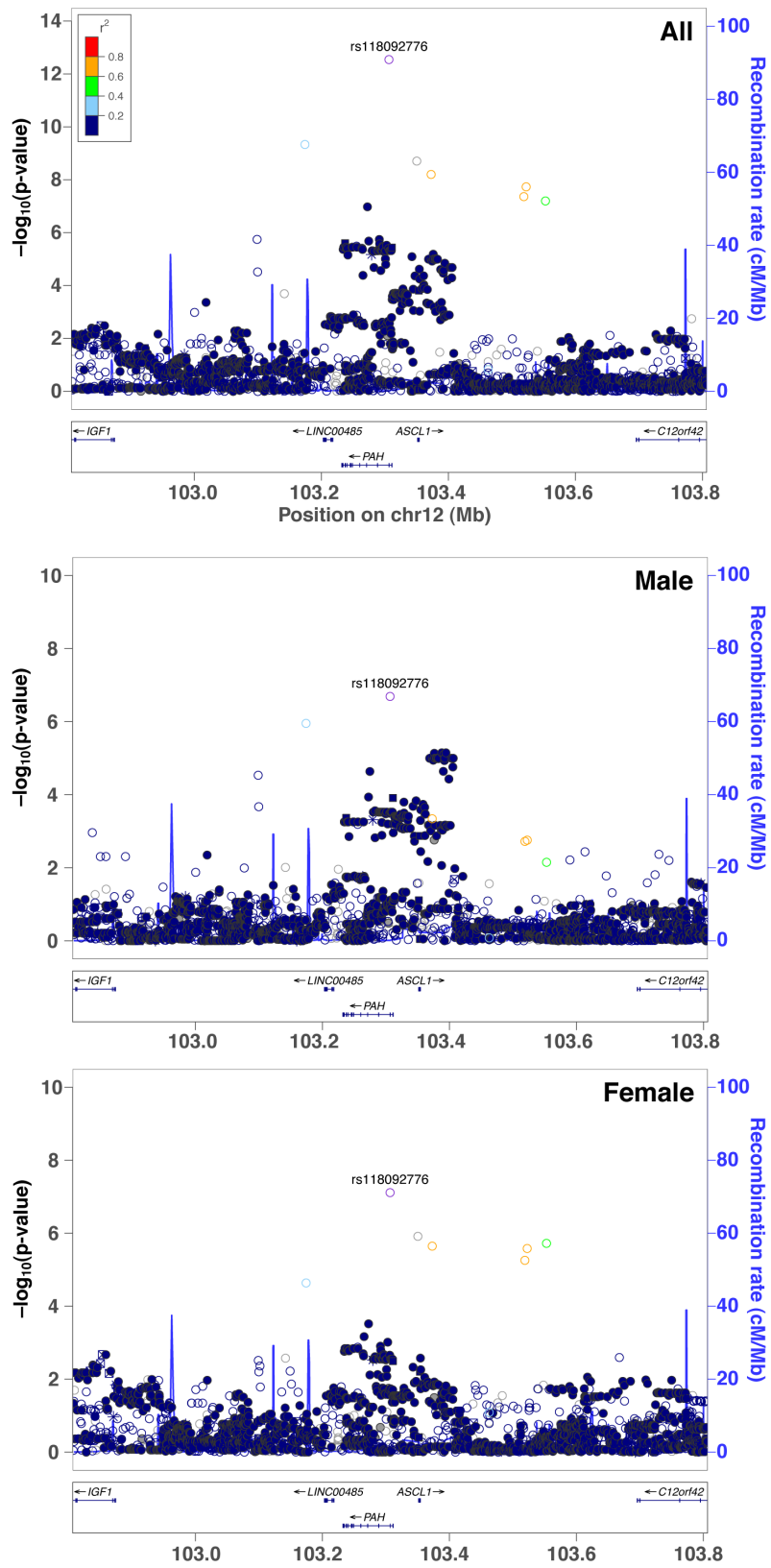
Supplementary Table S2: average values of metabolites

| Meabolite | SNP | gene | sex | Average | | | Fold change | |
|---------------|-------------|--------------|--------|---------|--------|-------|-------------|---------|
| | | | | wt | hetero | homo | hetero/wt | homo/wt |
| asparagine | rs8012505 | <i>ASPG</i> | all | 54.6 | 61.8 | 80.8 | 1.13 | 1.48 |
| | | | male | 56.1 | 59.8 | 84.7 | 1.07 | 1.51 |
| | | | female | 53.8 | 63.1 | 78.9 | 1.17 | 1.47 |
| phenylalanine | rs118092776 | <i>PAH</i> | all | 57.2 | 68.2 | N.A. | 1.19 | N.A. |
| | | | male | 59.9 | 70.9 | N.A. | 1.18 | N.A. |
| | | | female | 55.6 | 66.6 | N.A. | 1.2 | N.A. |
| proline | rs5747933 | <i>PRODH</i> | all | 149.5 | 179.0 | 231.1 | 1.2 | 1.55 |
| | | | male | 165.3 | 199.0 | 271.4 | 1.2 | 1.64 |
| | | | female | 139.9 | 168.2 | 210.9 | 1.2 | 1.51 |
| Glycine | rs1047891 | <i>CPSI</i> | all | 194.8 | 280.4 | 313.6 | 1.44 | 1.61 |
| | | | male | 187.8 | 217.5 | 213.5 | 1.16 | 1.14 |
| | | | female | 199.4 | 312.2 | 333.6 | 1.57 | 1.67 |
| formate | rs1801133 | <i>MTHFR</i> | all | 17.6 | 16.4 | 15.0 | 0.93 | 0.85 |
| | | | male | 18.1 | 16.6 | 15.5 | 0.92 | 0.86 |
| | | | female | 17.2 | 16.3 | 14.7 | 0.95 | 0.85 |
| creatinine | rs820336 | <i>MYLK</i> | all | 54.9 | 61.5 | N.A. | 1.12 | N.A. |
| | | | male | 66.9 | 63.2 | N.A. | 0.94 | N.A. |
| | | | female | 47.6 | 60.3 | N.A. | 1.27 | N.A. |

(a) Asparagine rs8012505

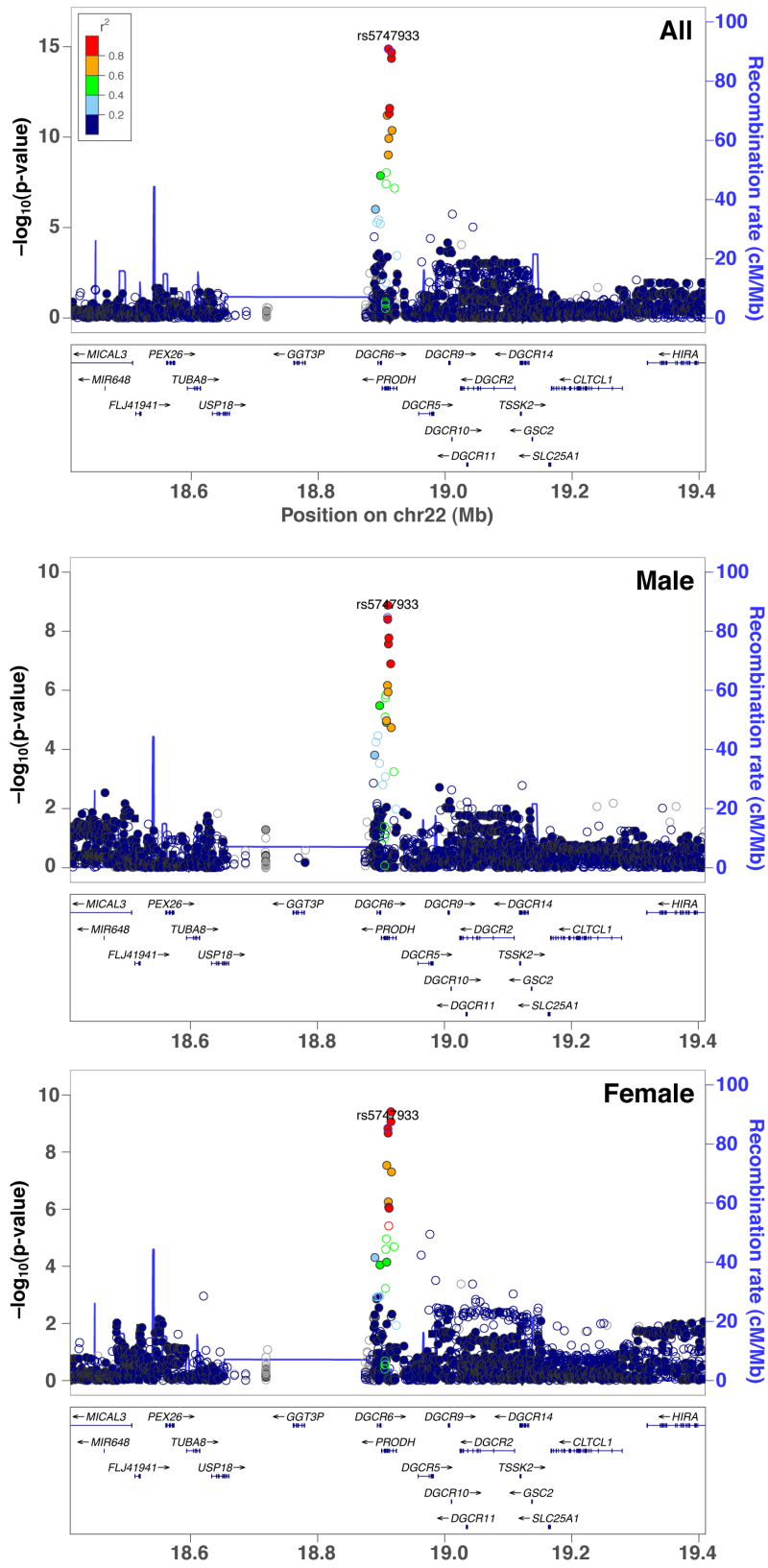


(b) Phenylalanine rs118092776

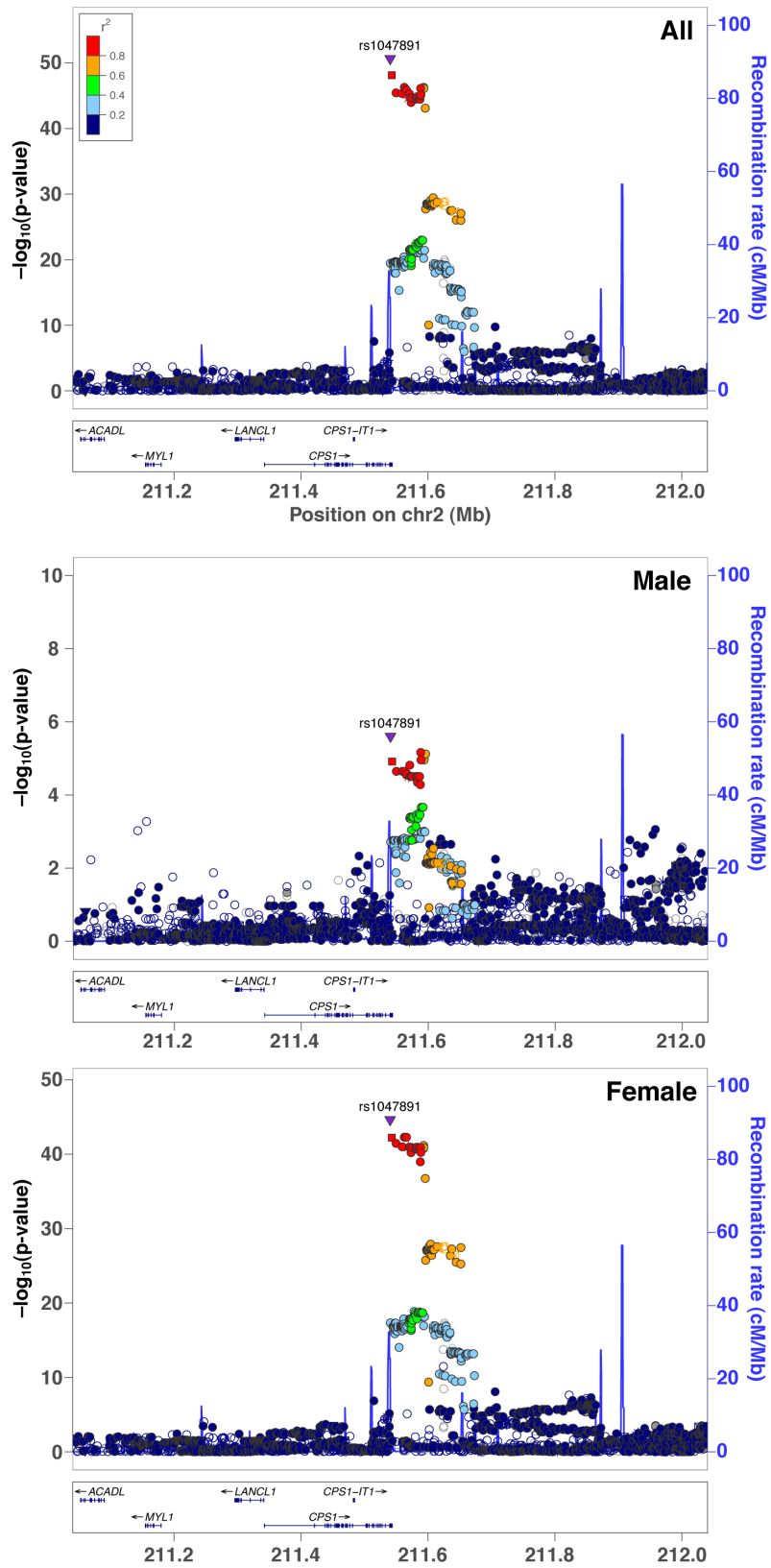


(c)

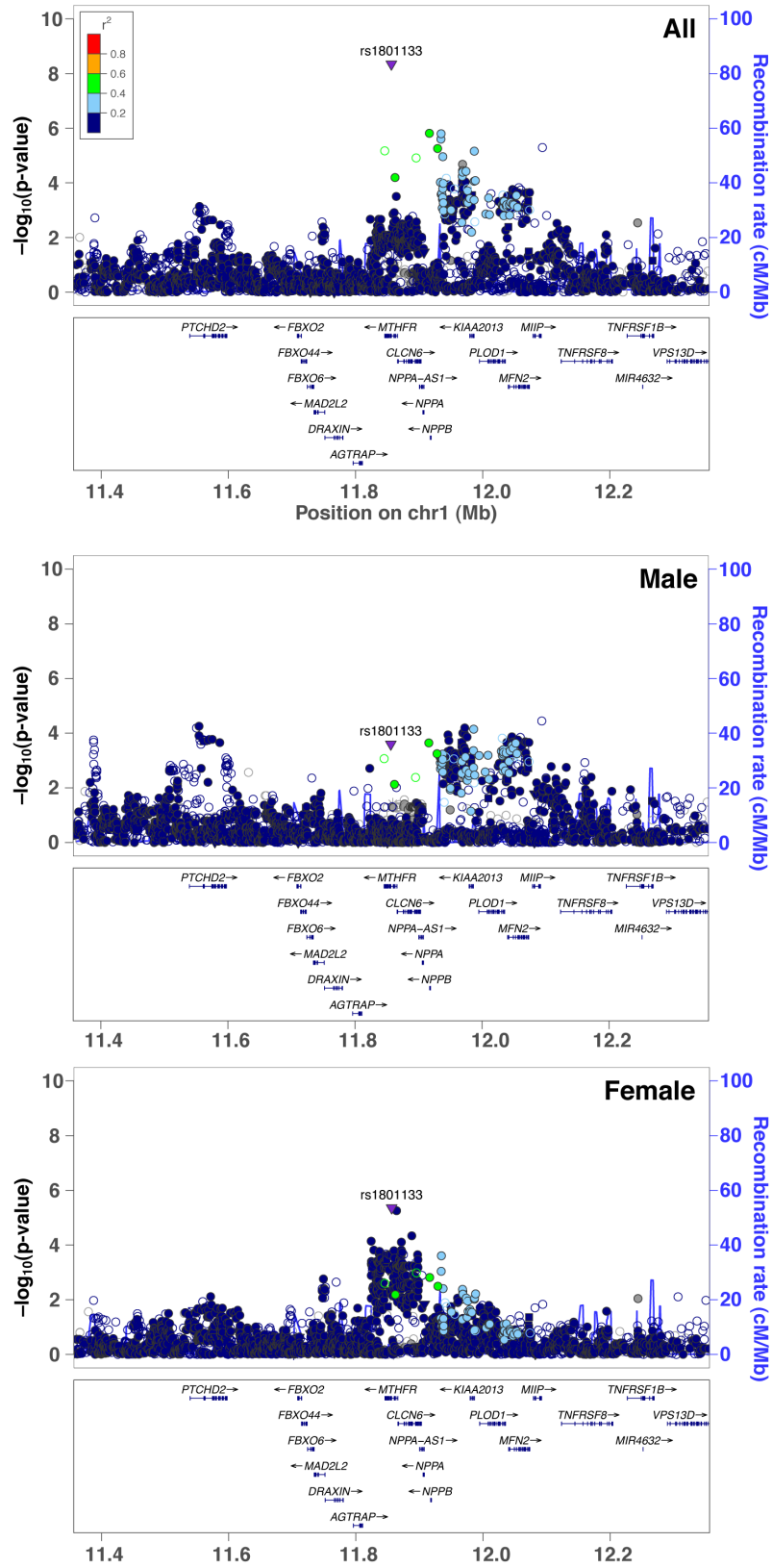
Proline rs5747933



(d) Glycine rs1047891

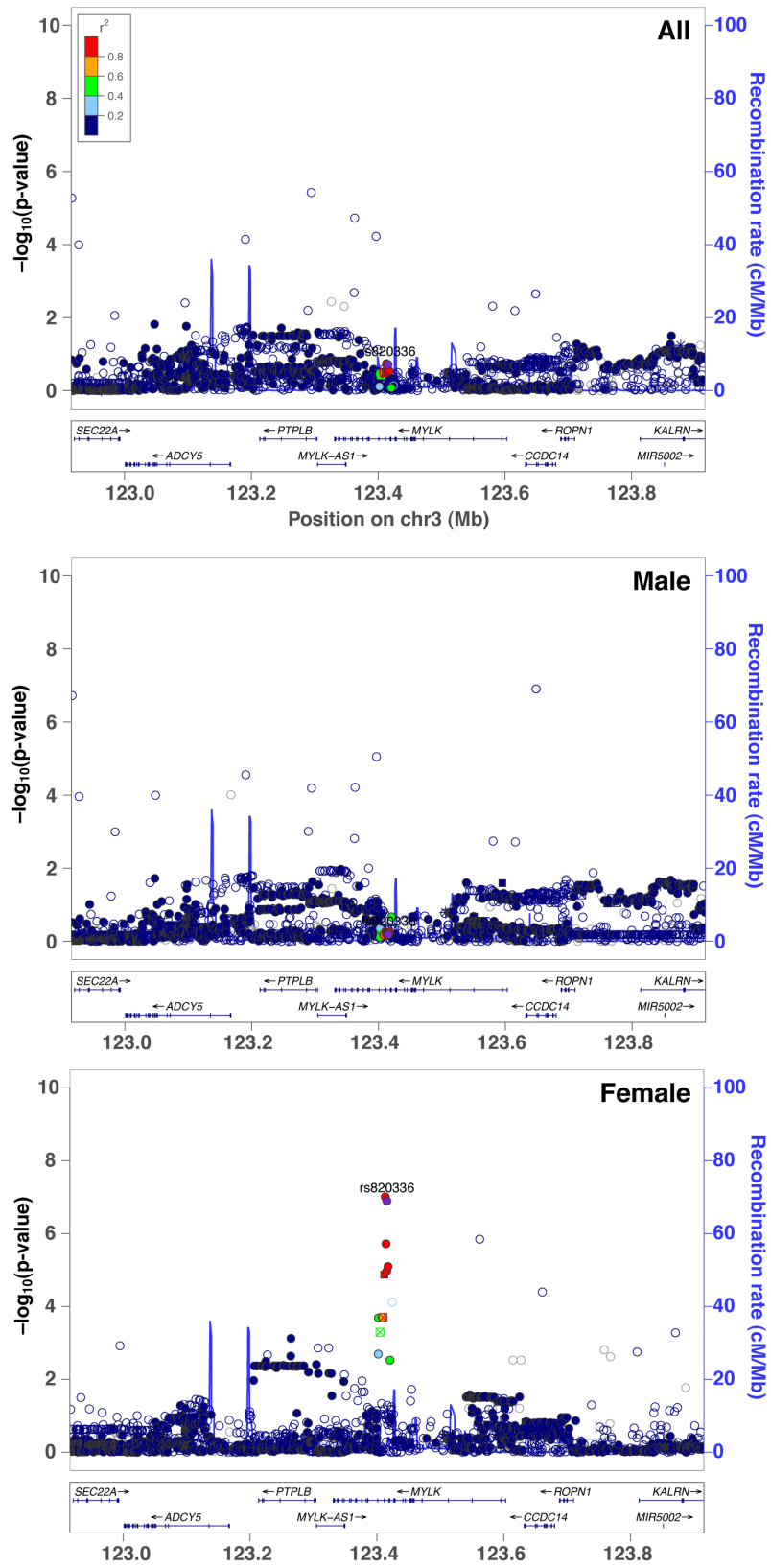


(e) Formate rs1801133



(f)

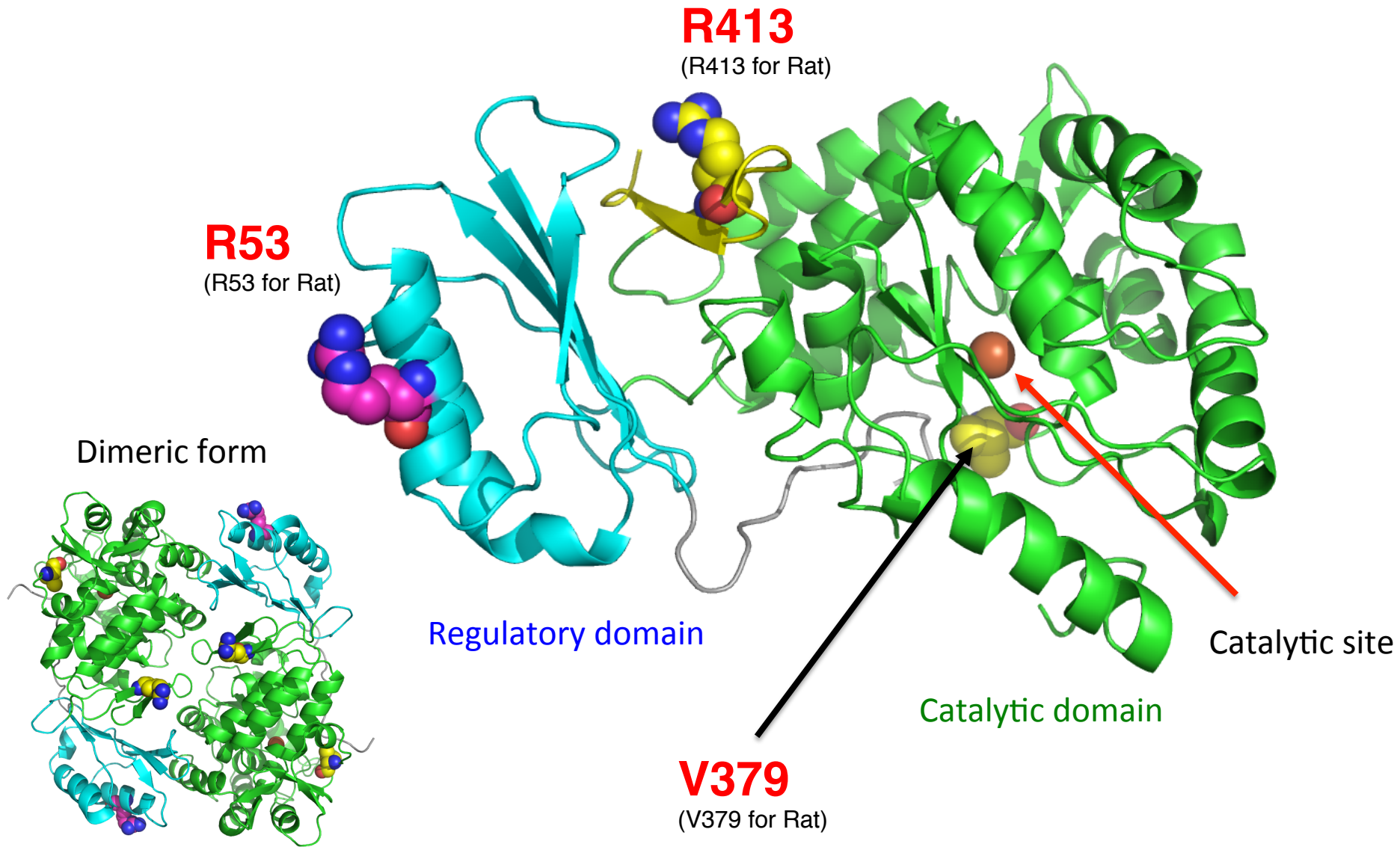
Creatinine rs820336



Supplementary Figure S1: Regional association plots for loci reported in this study. Statistical Significance of associated SNPs are plotted on the $-\log_{10}(\text{P-value})$ scale as a function of chromosomal position (NCBI 37). The identified causal SNP at each locus is shown in purple. Correlation of the causal SNP to other SNPs at each locus is shown on a scale from minimal (blue) to maximal (red). Estimate recombination rate are also shown.

a

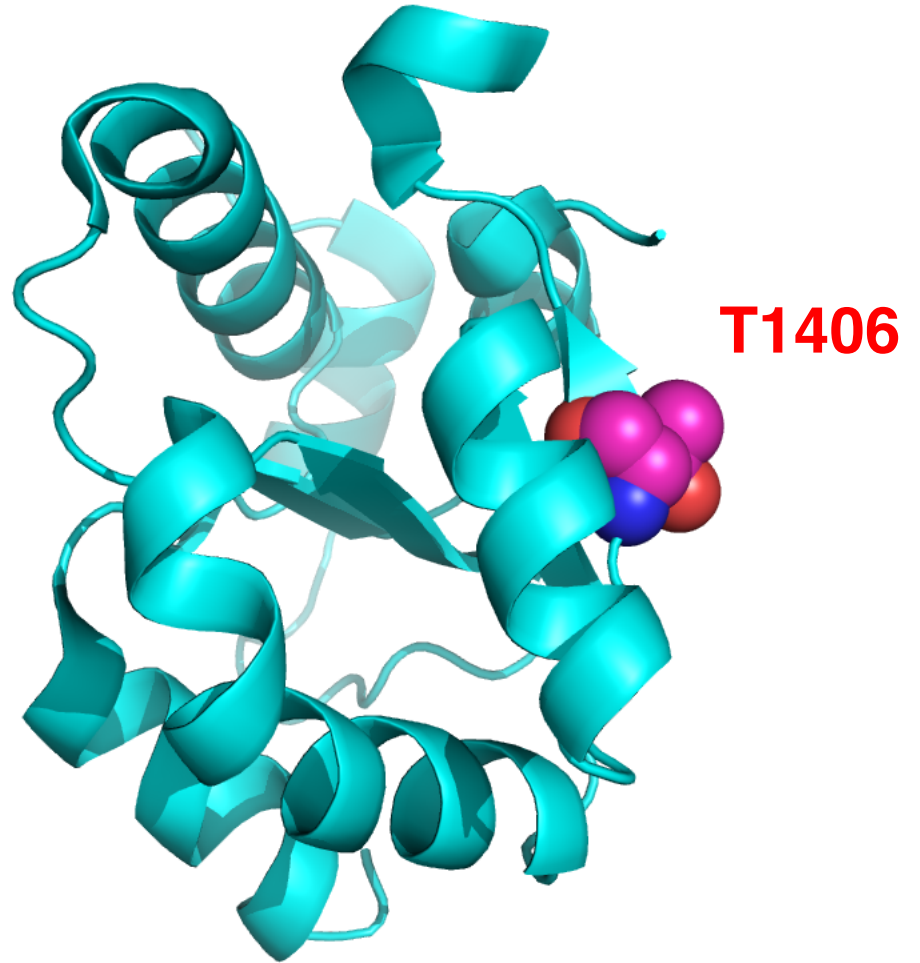
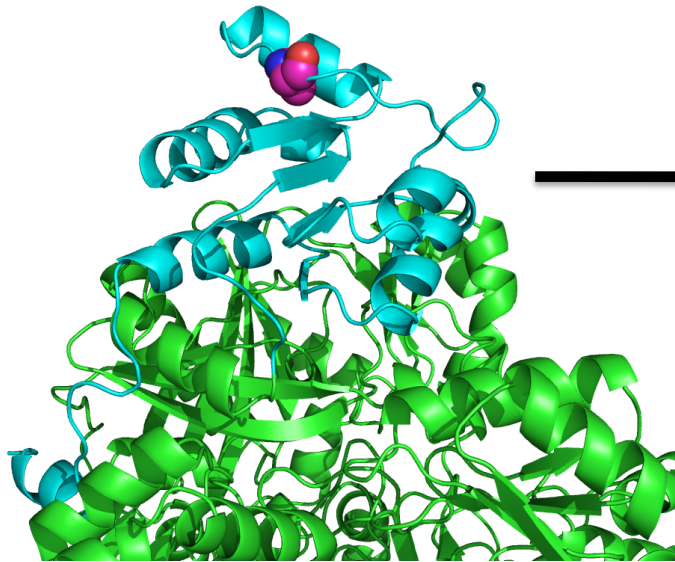
Rat phenylalanine hydroxylase (PAH) (PDB ID: 2PHM)



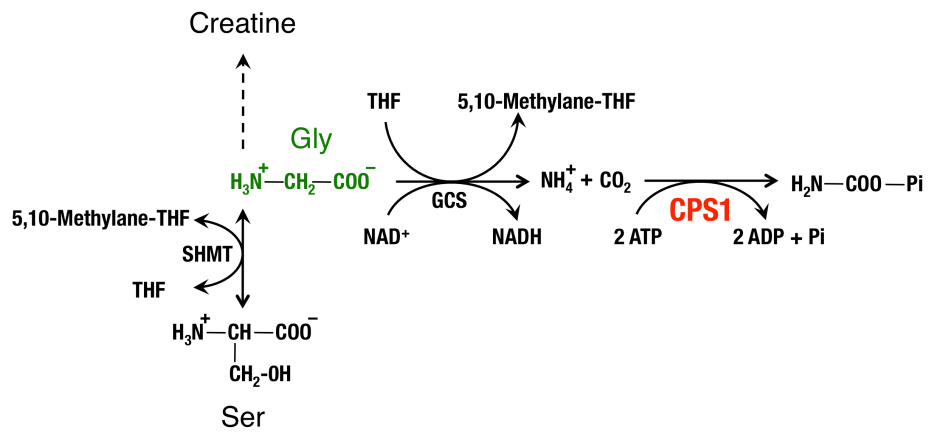
b

Human CPS1 C-terminal regulatory domain
(PDB ID:2YVQ)

E. Coli CPS (PDB ID: 1BXR)

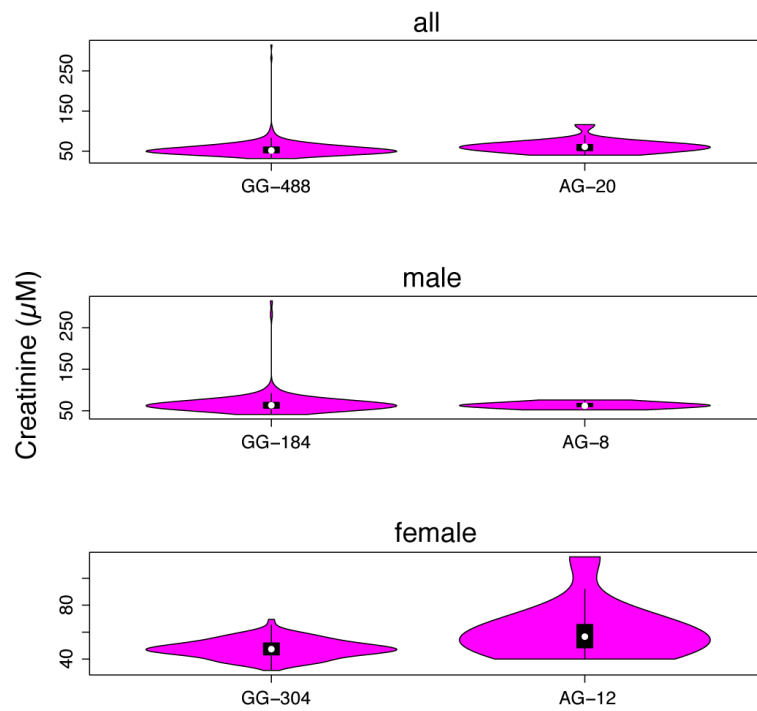


Supplementary Figure S2: (a) Structural characteristics of the non-synonymous variants in PAH. (b) Orientation of the regulatory domain relative to the catalytic domain in *E coli* CPS structure (PDB ID: 2BXR). All figures were made using the PyMOL Software Package (<https://www.pymol.org/>).



Supplementary Figure S3: Schematic representation of metabolite pathway involving glycine. CPS1 was shown in red. Glycine was shown in green.

rs820336 (MYLK)



Supplementary Figure S4: Distribution of the plasma creatinine concentration across the genotype of a locus summarized in Supplementary Table S3b. The distribution was shown using a violin plot. This figure was made using R package.