

# Concept mapping One-Carbon Metabolism to model future ontologies for nutrient–gene–phenotype interactions

A. C. Joslin · R. Green · J. B. German ·  
M. C. Lange

Received: 19 October 2013 / Accepted: 10 July 2014 / Published online: 5 August 2014  
© Springer-Verlag Berlin Heidelberg 2014

**Abstract** Advances in the development of bioinformatic tools continue to improve investigators' ability to interrogate, organize, and derive knowledge from large amounts of heterogeneous information. These tools often require advanced technical skills not possessed by life scientists. User-friendly, low-barrier-to-entry methods of visualizing nutrigenomics information are yet to be developed. We utilized concept mapping software from the Institute for Human and Machine Cognition to create a conceptual model of diet and health-related data that provides a foundation for future nutrigenomics ontologies describing published nutrient–gene/polymorphism–phenotype data. In this model, maps containing phenotype, nutrient, gene product, and genetic polymorphism interactions are visualized as triples of two concepts linked together by a linking phrase. These triples, or “knowledge propositions,” contextualize aggregated data and information into easy-to-read knowledge maps. Maps of these triples enable visualization of genes spanning the One-Carbon Metabolism (OCM) pathway, their sequence variants, and multiple literature-mined associations including concepts relevant to nutrition, phenotypes, and health. The concept map development process documents the incongruity of information derived from pathway databases versus literature resources. This conceptual model highlights the importance of incorporating information about genes in upstream pathways that provide

substrates, as well as downstream pathways that utilize products of the pathway under investigation, in this case OCM. Other genes and their polymorphisms, such as TCN2 and FUT2, although not directly involved in OCM, potentially alter OCM pathway functionality. These upstream gene products regulate substrates such as B12. Constellations of polymorphisms affecting the functionality of genes along OCM, together with substrate and cofactor availability, may impact resultant phenotypes. These conceptual maps provide a foundational framework for development of nutrient–gene/polymorphism–phenotype ontologies and systems visualization.

**Keywords** One-Carbon Metabolism · Nutrigenomics · Ontology · CMAPs

## Introduction

As the field of nutrigenomics accrues more information, methods of easy organization and interrogation are required for evidence-based expert systems. Novel, easy-to-use, bioinformatic tools provide a method to interrogate and understand data generated from “-omics” experiments (Lindblom and Robinson 2011). Ontological analysis has become a commonly used bioinformatics technique to organize and characterize large amounts of information. One of the most successful ontologies is the Gene Ontology. The Gene Ontology is a knowledge resource designed to allow researchers to annotate genes with terms relating to biological processes, cellular locations, and molecular functions (Ashburner et al. 2000). When novel genes are discovered, their sequences can be compared to those previously characterized by Gene Ontology terms whereby information concerning the novel gene can be

---

A. C. Joslin · J. B. German · M. C. Lange (✉)  
Department of Food Science and Technology, University of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA  
e-mail: mclange@ucdavis.edu

R. Green  
Department of Pathology and Laboratory Medicine, School of Medicine, University of California Davis, 4400 V Street, Sacramento, CA 95817, USA

hypothesized. While several studies investigate nutrient–gene interactions and their phenotypic consequences, a comprehensive ontology for annotating these interactions remains to be developed. Taboada et al. (2012) provide a disease-based ontological model of genotype–phenotype associations that outlines a preliminary approach for such an ontology. In their model, patient data are used to create a queryable database that documents concepts related to the rare disease cerebrotendinous xanthomatosis (Taboada et al. 2012). The authors suggest that semantic web technologies are essential for translating genomic and genetic research into actionable phenotype information. These technologies allow geneticists and clinicians to ask higher-order questions of the dataset such as whether a particular genotype is associated with specific phenotypic traits.

The principles and applications envisioned in the model described by Taboada et al. (2012) can be broadly extended to ontologically enriched nutrient–gene interaction databases that provide a platform enabling users to query nutrient–gene–phenotype associations. Information generated from such databases will be useful for defining, characterizing, and ultimately personalizing nutrition. Unambiguous vocabularies describing metabolic phenotypes (“metabotypes”), their genetic and environmental precursors, and the relative strength and extent of these associations will advance the state of the science of personalized nutrition and offer the potential for delivery of personalized foods and diets.

Presented here is a model for a novel form of data integration and visualization of nutrient–gene–phenotype relationships. The model highlights physiological and metabolic consequences, as well as potential alterations in nutritional requirements, resulting from genetic variation in the One-Carbon Metabolism (OCM) pathway. The OCM pathway is a well-described central pathway of metabolism containing many characterized genetic polymorphisms. Our conceptual model of OCM acts as a preliminary model for a diet personalization knowledge base, visualizing known interactions of vitamins B12, B6, folate and riboflavin with information related to genetic polymorphisms, altered disease risks resulting from these polymorphisms, and potential nutrient interventions. There are other micronutrients such as choline, known to affect functionality of the OCM, but an in-depth discussion of these nutrients lies outside the scope of this current manuscript. Independent studies have already documented evidence that optimal vitamin intakes within a population vary due to single-nucleotide differences in genes within the OCM pathway (Bailey and Gregory 1999; Kelemen et al. 2008; Stover and Garza 2002; Yang et al. 2008; Zittan et al. 2007). In addition to altered vitamin requirements, individuals with SNPs in genes encoding for enzymes residing along this pathway may also have increased susceptibility

for comorbidity risk factors such as hyperhomocysteinemia (Garrod et al. 2010) and diseases such as Alzheimer’s type dementia (Casalheira et al. 2009) and cancer (Matakidou et al. 2007; Vlastos et al. 2009).

Concept maps (CMAPs) (<http://cmap.ihmc.us/download/>, accessed on June 2, 2012) were used to visualize gene, polymorphism, disease, and allele frequency data with links to references from scientific literature. Concept mapping is designed to allow for quick and simple visualization of knowledge. Concepts can be supplemented with annotations such as ontological terms, literature references, relevant images, and references to other CMAPs. These maps are machine readable and provide a gateway for eventual development of more formalized ontologies capable of covering the broad range of information required for personalized nutrition and disease-risk assessment.

Mapping phenotype data from literature is one method of channeling information from basic research to a clinical setting (Taboada et al. 2012). Others have presented methodologies for building nutrigenomics ontologies using CMAPs (Castro et al. 2006). The project documented by Castro et al. has transitioned into Functional Genomics Ontology (FuGO), and concomitantly lost its specificity for characterization of nutrient–gene interaction data, information, and knowledge (Whetzel et al. 2006). The project presented in this paper is part of a larger effort consisting of two distinct methodologies: a manual curation component and a computational component. This paper describes the manual curation portion of the overall project.

## Methods

### Data acquisition/curation

Manual curation of data was achieved through several distinct steps: first, genes involved in OCM were consolidated from WikiPathways, KEGG, and Gene Ontology databases to form a preliminary gene list (Table 1). KEGG “One-Carbon Pool by Folate” (site <http://www.genome.jp/kegg/pathway/map/map00670.html>, accessed on April 2, 2013) and Gene Ontology “One-Carbon Metabolic Process (with child terms)” (<http://www.geneontology.org>, April 15, 2013) were chosen as references because of the sustained development efforts in their creation, and established use in scientific studies. In addition to the reasons listed above, WikiPathways was chosen because it is easily curated and visually unique from the two other databases. The genes listed in Table 1 provided terms used to query PubMed for articles relating to One-Carbon Metabolism genes and phenotype associations. Literature mining involved a step-by-step process that included searching gene names such as “MTHFR” together with phrases such as “Vitamin B12” or

**Table 1** Linking phrase comparison

OCM CMAP Linking phrases	Relation ontology	Gene ontology	SOPHARM	New relations
Can be exacerbated by				✓
Can be mediated by				✓
Catalyzes				✓
Causes decrease in				✓
Causes increase in				✓
Characterized by				✓
Compared to group				✓
Expressed in	✓			
Has association			✓	
Has cofactor				✓
Has frequency				✓
Has function	✓			
Has gene product	✓			
In group				✓
Results in	✓			
Located in	✓			
Has parts	✓	✓	✓	
Has variant	✓		✓	
May utilize gene products from				✓
Utilizes gene products from				✓

This table depicts all linking phrases used in our CMAPs and compares them to those found in the relation ontology, the gene ontology, and SOPHARM. Those with checks are present in each ontology

“One-Carbon Metabolism” in order to find and select articles from the PubMed database. Articles were included as references in the maps if their content contained information such as OCM gene function, changes in nutrient requirements due to SNPs in the OCM pathway, or phenotype associations with OCM genetic variants. Some articles mined during these searches listed One-Carbon genes not found in either KEGG, WikiPathways, or GO resources, we termed these “literature mined genes”. These literature mined genes were combined with the preliminary list to create a master gene list. This updated list was used to query PubMed for additional articles and to link each enzyme to HumanCyc, a curated biochemical database of enzymatic reactions in humans ([www.humancyc.org](http://www.humancyc.org)) (Romero et al. 2004). Next, information such as allelic frequencies, study group data, gene–SNP–disease associations, and gene–SNP–nutrient metabolism associations were added to the CMAPs as they were uncovered. Finally, an established domain expert in the field of OCM reviewed the information visualized in each map. Expert curation is important to reduce incorrect interpretation of data. This expert also reviewed the genes found via

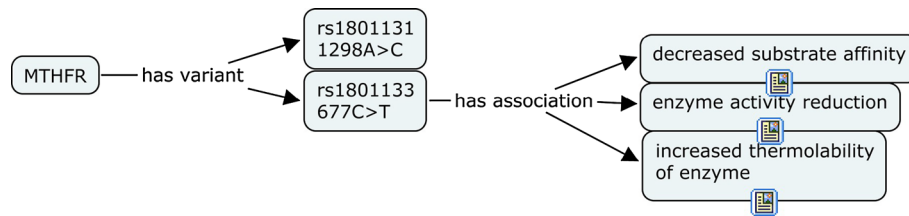
literature mining that were not present in the databases, and additionally provided an illustrative rationale for including genes that are not directly involved in a pathway but are important for pathway functionality and homeostasis.

#### Software environment


CMAP Tools software was chosen as the preferred knowledge mapping environment due to its low-barrier-to-entry, quick learning curve, and potential for ontology creation. In CMAPS, information is viewed and entered into the software in the form of triples (propositions) consisting of a subject, linking phrase (relationship), and object (Fig. 1). The creation of triples is relatively simple due to the interchangeability of subjects and objects, meaning the object of one triplet can become the subject of another. Therefore, several triples can combine to form longer-chain conceptual phrases. The primary literature reference for each triple was added as an annotation to the object portion of the proposition. Common subjects and objects were genes, SNPs, diseases, disease states, and variations in vitamin requirements. When possible, linking phrases were chosen from the widely accepted relation ontology (RO). The RO is a collection of relations developed within the Open Biological and Biomedical Ontologies (OBO) database with the aim of standardizing ontologies (Smith et al. 2005). When linking phrases could not be used from the RO, they were taken from ontologies such as the suggested ontology for pharmacogenomics (SOPHARM), or the gene ontology. When these relations did not adequately characterize the relationship, new linking phrases were developed (Table 1). Linking phrases created during the course of this concept mapping project were to be abstract enough to convey a range of information, yet detailed enough to enable eventual computability of the relationships. For example, the linking phrase “has association” was used to represent co-occurrences between SNPs and diseases (e.g., “lung cancer”), SNPs and health states (e.g., “hyperhomocysteinemia”), SNPs and protein or enzyme activity (e.g., “increased thermolability”), and SNPs and relative vitamin requirements (e.g., “increased requirement”).

#### Results

The Gene Ontology utilizes a small subsection of the RO including “is a,” “part of,” “has part,” “regulates,” “positively regulates,” and “negatively regulates” to define GO terms (accessed 4-2-2013). The SOPHARM ontology uses a subset of linking phrases closely related to nutrigenomics, and although the fields of pharmacogenomics and nutrigenomics are similar, only a few relationships could be properly employed. Many of the linking



**Fig. 1** Examples of knowledge “triples” (or “propositions”). A triple consists of a subject, linking phrase (or “predicate”), and object. In this case, MTHFR is the subject of two triples, one with rs1801131 as an object and one with rs1801133 as an object. rs1801133 is also the subject of three triples linked with the phrase “has association” to the objects “increased thermostability of enzyme,” “enzyme activity reduction,” and “decreased substrate

affinity.” The two linking phrases, “has variant” and “has association,” were used throughout the CMAP models. Inside of the web-accessible maps, the  icon is a link to the literature reference(s) for the citation supporting the knowledge proposition that ends with the object upon which the icon sits

phrases present in the RO and other ontologies do not have the flexibility to account for conditional data. “Has association” became the most frequently used linking phrase because it conveys broad associative information while leaving room for development of more detailed conditional and/or causative relationships. As nutrient–gene–phenotype association knowledge advances, more refined linking phrases can be created as appropriate (Table 1).

A total of eighty-three genes affecting OCM were compiled. These genes were found in KEGG, WikiPathways, and Gene Ontology databases, as well as from pre-compiled literature sources such as Wernimont et al. (Table 2). WikiPathways’ OCM pathway contains seven unique genes not found in the other three databases, while the Gene Ontology’s One-Carbon Metabolic Process gene list contains seventeen unique genes. KEGG’s OCM pathway contained no unique genes. Literature mining revealed thirty-seven genes considered important to OCM, but were not listed as being either part of, or associated with, OCM in KEGG, WikiPathways, or Gene Ontology databases.

Table 3 shows the thirty-seven genes uncovered from literature mining, organized by pathway involvement, with links to the literature sources from which these associations were abstracted. The “OCM plus Upstream and Downstream Pathway Integration” CMAP highlighted in Fig. 2 displays the ways in which several groups of these genes interact with OCM. Many of these genes are involved in direct upstream effector pathways, as well as downstream pathways utilizing the products of OCM. Products from the genes in upstream pathways mediate overall levels of cofactors and substrates for OCM reactions, while variants of downstream pathway genes can influence utilization patterns of OCM products.

All genes mined from the literature can be visualized in the OCM CMAP (Fig. 3). Those found exclusively in WikiPathways are yellow concepts, those found exclusively in KEGG are purple concepts, and genes found

exclusively in GO are pink. Genes found in two or more of these pathway Web sites are blue, and the genes mined specifically from literature searches are gray. All genes found in KEGG, WikiPathways, and GO are presented with their respective enzymatic reaction from the HumanCyc database. TCN2 is not represented as part of a reaction in the HumanCyc database because TCN2 is a B12 transporter, not an enzyme. Several genes including methylenetetrahydrofolate reductase (MTHFR), methionine synthase (MTR), methionine synthase reductase (MTRR), and transcobalamin 2 (TCN2) were mapped in detail. These genes were chosen based on the relative importance for OCM and nutrient interactions, as evidenced by the sheer volume of scientific publications about each of these genes.

MTHFR, depicted in Fig. 4, is arguably the most studied gene in this pathway, with 5,032 hits when queried on GOPUBMED (<http://gopubmed.org>, accessed 5-21-2013). Variations in this gene are associated with many disorders, including vascular disease, Alzheimer disease, neural tube defects, and cancer (Cascalheira et al. 2009; Haviv et al. 2002; Matakidou et al. 2007; Shields et al. 1999). MTHFR is responsible for conversion of 5,10-methylenetetrahydrofolate into 5-methyltetrahydrofolate with a riboflavin cofactor (Ward et al. 2011). A number of SNPs within this gene are associated with variation in plasma levels of homocysteine and various B vitamins. The 677C>T mutation reduces the effectiveness of MTHFR’s enzyme activity, leading to increased homocysteine compared with the wild-type (C) allele (Yang et al. 2008). Dysregulated homocysteine levels are linked to multiple diseases, many of which are exacerbated when coupled with low vitamin levels (Cascalheira et al. 2009; Kelemen et al. 2008; Zittan et al. 2007).

MTR, depicted in Fig. 5, requires cobalamin as a cofactor during its enzymatic conversion of L-homocysteine to L-methionine. B12 supplementation has been shown to increase the activity of this enzyme twofold to 14-fold

**Table 2** Preliminary genes gathered from KEGG, WikiPathways, Gene Ontology, and text mining from journals

KEGG: "One-Carbon Pool by Folate"	WikiPathways: "One-Carbon Metabolism"	Gene Ontology: "One-Carbon Metabolic Process" curation	Literature-mined genes absent in KEGG, WikiPathway, and Gene Ontology databases
ATIC	ATIC		AMN
	AHCY	AHCY	BHMT2
		AHCYL1	CBS
		AHCYL2	CELF1
ALDH1L1 (aka FTHFD)	ALDH1L1 (aka FTHFD)	ALDH1L1 (aka FTHFD)	CTH
ALDH1L2		ALDH1L2	CUBN
AMT	AMT		DMGDH
	BHMT		DMNT2
		CA1	DPYD
		CA12	FOLR1
		CA13	FOLR2
		CA2	FOLR3
		CA3	FTH1
		CA4	FUT2
		CA5A	GCSH
		CA5B	GGH
		CA6	GLDC
		CA7	HSPA8
		CA8	LMBRD1
		CA9	MARS
DHFR	DHFR	DHFR	MBD2
DHFRL1		DHFRL1	MGMT
	DNMT1		MLH1
	DNMT3a		MRP1
	DNMT3b		MSH2
	FOLH1		NOS3
		FPGS	PCFT
FTCD	FTCD	FTCD	SARDH
GART	GART		SLC19A1
		GNMT	SLC19A2
	MAT1A	MAT1A	SLC19A3
		MAT2A	SLC25A32
	MAT2B	MAT2B	SLC46A1
MTFMT	MTFMT		TCBLR
MTHFD1	MTHFD1	MTHFD1	TCN1
MTHFD1L	MTHFD1L	MTHFD1L	UBE2I
MTHFD2	MTHFD2	MTHFD2	UBE2 N
MTHFD2L		MTHFD2L	
MTHFR	MTHFR	MTHFR	
MTHFS	MTHFS		
MTR (aka MS)	MTR (aka MS)		
	MTRR		
SHMT1	SHMT1	SHMT1	
SHMT2	SHMT2	SHMT2	
	TCN2		
TYMS	TYMS		

This is a comprehensive list of all genes gathered throughout literature and pathway mining

**Table 3** Characterization of literature mined genes absent in KEGG, WikiPathways, and GO

Function	Gene ID	OCM function	Pub Med link	References
Aminoacyl-tRNA synthetase	MARS	Methionyl-tRNA synthetase (methionine + tRNA)	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
Antigen synthesis	FUT2	Influences B12 levels	<a href="http://www.ncbi.nlm.nih.gov/pubmed/18776911">http://www.ncbi.nlm.nih.gov/pubmed/18776911</a>	Hazra et al. (2008)
B12 transport	CUBN	Intrinsic factor-cobalamin receptor	<a href="http://www.ncbi.nlm.nih.gov/pubmed/20615890">http://www.ncbi.nlm.nih.gov/pubmed/20615890</a>	Kozyraki et al. (1998), Levine et al. (2010)
	TCbIR	Transcobalamin receptor	<a href="http://www.ncbi.nlm.nih.gov/pubmed/20577008">http://www.ncbi.nlm.nih.gov/pubmed/20577008</a>	Pangilinan et al. (2010)
	LMBRDI	Lysosomal B12 transporter	<a href="http://www.ncbi.nlm.nih.gov/pubmed/20446115">http://www.ncbi.nlm.nih.gov/pubmed/20446115</a>	Rutsch et al. (2011, p. 1)
	AMN	Essential for correct CUBN expression	<a href="http://www.ncbi.nlm.nih.gov/pubmed/21750092">http://www.ncbi.nlm.nih.gov/pubmed/21750092</a>	Namour et al. (2011)
	TCN1	B12 transport protein	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
ABC transporter	MRP1	B12 export out of cells	<a href="http://www.ncbi.nlm.nih.gov/pubmed/19897579">http://www.ncbi.nlm.nih.gov/pubmed/19897579</a>	Beedholm-Ebsen et al. (2010)
Sarcosine metabolism	DMGDH	Choline metabolism	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
	SARDH	Choline metabolism	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Green et al. (2013), Wernimont et al. (2011)
			<a href="http://www.ncbi.nlm.nih.gov/pubmed/23824605">http://www.ncbi.nlm.nih.gov/pubmed/23824605</a>	
DNA synthesis and repair	MGMT	DNA methyltransferase	<a href="http://www.ncbi.nlm.nih.gov/pubmed/17119116">http://www.ncbi.nlm.nih.gov/pubmed/17119116</a>	Kelemen et al. (2008), Lim et al. (2007)
			<a href="http://www.ncbi.nlm.nih.gov/pubmed/18381459">http://www.ncbi.nlm.nih.gov/pubmed/18381459</a>	
	MLH1	DNA repair—depends on one-carbon supplies	<a href="http://www.ncbi.nlm.nih.gov/pubmed/17119116">http://www.ncbi.nlm.nih.gov/pubmed/17119116</a>	Lim et al. (2007)
	MSH2	DNA repair—depends on one-carbon supplies	<a href="http://www.ncbi.nlm.nih.gov/pubmed/17119116">http://www.ncbi.nlm.nih.gov/pubmed/17119116</a>	Lim et al. (2007)
Ferritin subunit	FTH1	Regulates folate metabolism	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011), Woeller et al. (2007)
			<a href="http://www.ncbi.nlm.nih.gov/pubmed/17702748">http://www.ncbi.nlm.nih.gov/pubmed/17702748</a>	

Table 3 continued

Function	Gene ID	OCM function	Pub Med link	References
Folate transport and metabolism	GGH	Determines intercellular folate availability	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Oppeneer et al. (2012), Wernimont et al. (2011)
	SLC19A1/PCFT	Folate transporter	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22018726">http://www.ncbi.nlm.nih.gov/pubmed/22018726</a>	Blanton et al. (2011)
	SLC19A2	Thiamine (B1) transporter	<a href="http://www.ncbi.nlm.nih.gov/pubmed/21254359">http://www.ncbi.nlm.nih.gov/pubmed/21254359</a>	Wernimont et al. (2011)
	SLC19A3	Thiamine (B1) transporter	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
	SLC25A32	Mitochondrial folate carrier	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
	SLC46A1	Folate transporter	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Kelemen et al. (2008)
	FOLR1	Folate receptor	<a href="http://www.ncbi.nlm.nih.gov/pubmed/18381459">http://www.ncbi.nlm.nih.gov/pubmed/18381459</a>	Blanton et al. (2011)
	FOLR2	Folate receptor	<a href="http://www.ncbi.nlm.nih.gov/pubmed/21254359">http://www.ncbi.nlm.nih.gov/pubmed/21254359</a>	Blanton et al. (2011)
	FOLR3	Folate receptor	<a href="http://www.ncbi.nlm.nih.gov/pubmed/20683905">http://www.ncbi.nlm.nih.gov/pubmed/20683905</a>	O'Byrne et al. (2010)
Glycine metabolism	GCSH	Glycine cleavage	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
	GLDC	Glycine cleavage	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
Homocysteine metabolism	CBS	Converts homocysteine to cystathionine	<a href="http://www.ncbi.nlm.nih.gov/pubmed/23913011">http://www.ncbi.nlm.nih.gov/pubmed/23913011</a>	Li et al. (2013)
	CTH	Cystathionine gamma-lyase	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011)
Methylation of aspartic acid tRNAs	DNMT2 aka TRDM1	DNA methyltransferase	<a href="http://www.ncbi.nlm.nih.gov/pubmed/17533396">http://www.ncbi.nlm.nih.gov/pubmed/17533396</a>	Matakidou et al. (2007)
Methyl-CpG complex regulation	MBD2	Binds to methylated DNA, demethylase	<a href="http://www.ncbi.nlm.nih.gov/pubmed/23913011">http://www.ncbi.nlm.nih.gov/pubmed/23913011</a>	Lim et al. (2007)
			<a href="http://www.ncbi.nlm.nih.gov/pubmed/17119116">http://www.ncbi.nlm.nih.gov/pubmed/17119116</a>	
Nitric oxide production	NOS3	Inhibits MTR	<a href="http://www.ncbi.nlm.nih.gov/pubmed/21254359">http://www.ncbi.nlm.nih.gov/pubmed/21254359</a>	Blanton et al. (2011)

Table 3 continued

Function	Gene ID	OCM function	Pub Med link	References
Betaine metabolism	BHMT2	Betaine-homocysteine methyltransferase	<a href="http://www.ncbi.nlm.nih.gov/pubmed/18457970">http://www.ncbi.nlm.nih.gov/pubmed/18457970</a> <a href="http://www.ncbi.nlm.nih.gov/pubmed/18230605/">http://www.ncbi.nlm.nih.gov/pubmed/18230605/</a> <a href="http://www.ncbi.nlm.nih.gov/pubmed/18381459">http://www.ncbi.nlm.nih.gov/pubmed/18381459</a>	Li et al. (2008), Szegedi et al. (2008) Kelemen et al. (2008), Morine et al. (2012)
Thymidylate metabolism	DPYD	Degradation of thymine	<a href="http://www.ncbi.nlm.nih.gov/pubmed/22143355">http://www.ncbi.nlm.nih.gov/pubmed/22143355</a> <a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a> <a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a> <a href="http://www.ncbi.nlm.nih.gov/pubmed/22103680">http://www.ncbi.nlm.nih.gov/pubmed/22103680</a>	Wernimont et al. (2011) Wernimont et al. (2011) Wernimont et al. (2011) Wernimont et al. (2011)
Pre-mRNA splicing	CELF1			
Heat-shock protein	HSPA8			
Ubiquitination	UBE2N			
	UBE2I			

The genes found in the literature are characterized by pathway involvement or function

(Gulati et al. 1999). B12 limitation has also been shown to reduce MTR enzyme activity (Brunaud et al. 2003). Low plasma B12 levels can lead to a buildup of L-homocysteine that may progress to hyperhomocysteinemia (Watkins et al. 2002). Certain mutations in MTR cause cblG type of methylcobalamin deficiency resulting in a combination of homocystinuria and megaloblastic anemia (Carmel et al. 1988). MTR is also responsible for the catabolism of 5-methyltetrahydrofolate to tetrahydrofolate. Low MTR activity leads to a buildup of folate in the 5-methyltetrahydrofolate form, termed the “methyl folate trap” (Herbert and Zalusky 1962).

MTRR, depicted in Fig. 6, is important for the maintenance of MTR activity. It is responsible for the reduction of catalytically inert cob(II)alamin to cob(I)alamin for use by MTR (Wolthers and Scrutton 2007). MTRR is the mutated gene in megaloblastic anemia complementation type E (cblE) which is similar to cblG but has been proven genetically unique by complementation analysis (Rosenblatt et al. 1987). Polymorphisms in this gene have been shown to cause increased homocysteine concentrations which, when combined with other polymorphisms, may further increase homocysteine levels (Vaughn et al. 2004). MTRR has also been implicated as a predictor for Down syndrome (Amorim and Lima 2012).

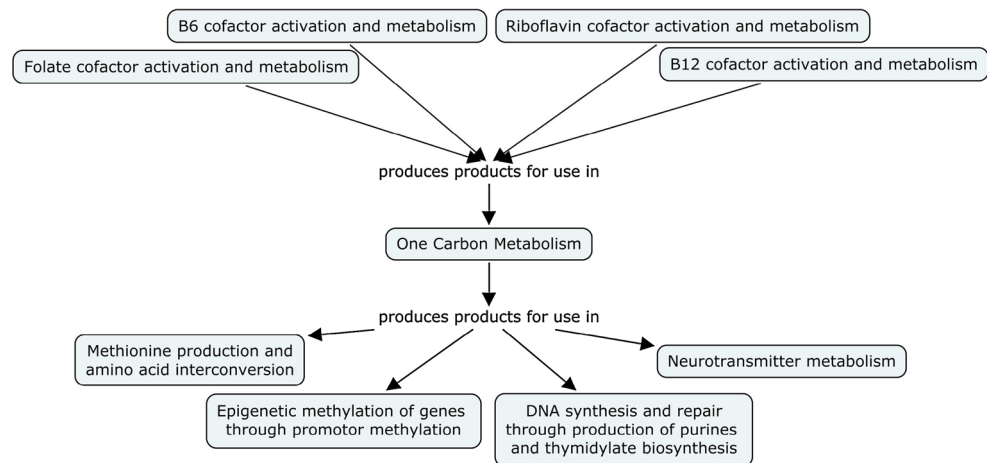
TCN2, depicted in Fig. 7, is expressed in many cells and directly mediates cellular uptake of cobalamin (Seetharam et al. 1999). The prevalence of certain alleles, such as 776C>G, in this gene varies dramatically between populations. The frequency of this allele may vary due to selective pressures such as *Plasmodium falciparum* (Guéant et al. 2007). The malaria parasite requires OCM substrates as cofactors for enzymes essential for growth (Kronenberger et al. 2014). The G allele may alter the structure of TCN2 so it has a lower affinity for B12, resulting in its more ready release (Afman et al. 2002; Garrod et al. 2010).

TCN2 efficiency directly affects the amount of vitamin B12 available for use by OCM (Garrod et al. 2010; Silla et al. 2011). This can be modeled easily within a CMAP because CMAP software is designed to flexibly incorporate all kinds of information. Surprisingly, in WikiPathways' OCM (*Homo sapiens*) pathway, TCN2 is visualized as contributing to MTR's conversion of 5-methyltetrahydrofolate to tetrahydrofolate. Although WikiPathways does not make it explicitly clear, TCN2 is not directly involved in MTR's catabolism of 5-methyltetrahydrofolate, it is instead important for the transport of B12 into cells (reference Web site: <http://www.ncbi.nlm.nih.gov/gene/6948>, accessed 6-8-2013). B12 is then used as a cofactor for MTR's enzymatic reaction.

The flexibility of CMAP Tool software is further highlighted in the homocysteine CMAP (Fig. 8). This map



**Fig. 2** OCM plus upstream and downstream pathway integration. One-Carbon Metabolism relies on many closely related pathways for substrate and cofactor production. A few of the important pathways are visualized here. OCM is also important for providing cofactors and substrates for a multitude of subsequent reactions ranging from epigenetic methylation of genes to purine production



depicts SNPs found in five genes known to affect homocysteine levels. Visualizing information in this manner can aid in disease-risk assessment based on constellations of SNPs and environmental and lifestyle factors that individually, or in ensemble, affect probability of developing a disease or risk factor. Single SNPs may not confer a measurable phenotype, but individuals with genotypes containing a number of mutations known to influence homocysteine levels may be more likely to develop hyperhomocysteinemia, and increasing susceptibility for hyperhomocysteinemia-associated diseases. It is interesting to note that a single gene may contain multiple polymorphisms that alter homocysteine levels. Some of these polymorphisms may increase levels, while others in the same gene may concurrently decrease levels.

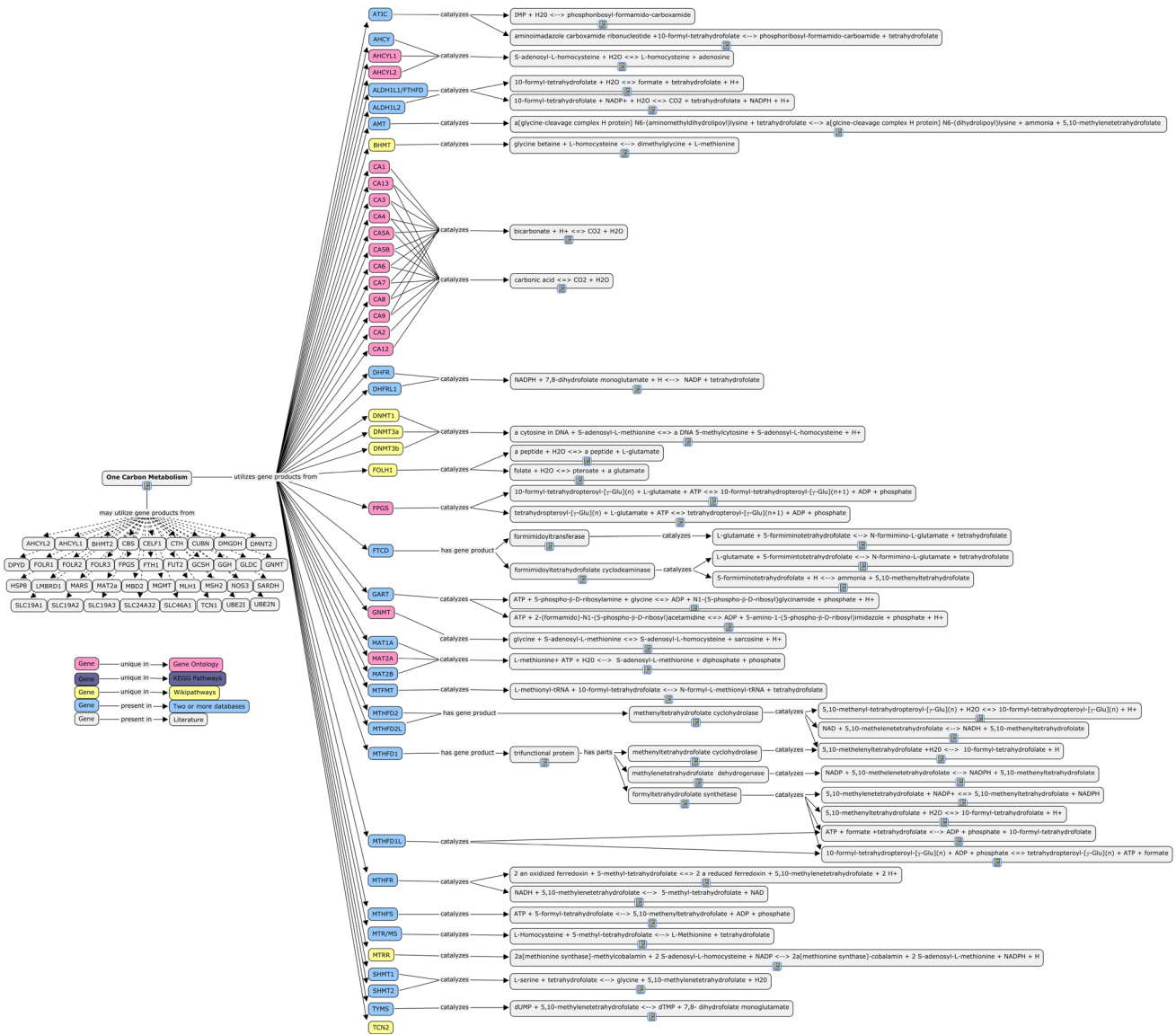
## Discussion

CMAP Tools was initially designed as educational software with a primary directive of providing a low technical barrier for users. For this study, CMAP Tools software was used as an easily accessible knowledge elicitation tool and heterogeneous information integration environment. These features—ease of use, heterogeneous information integration, and knowledge network visualization, make CMAPs a desirable system compared with many specialized bioinformatics tools. Oftentimes, these tools require technical know-how and are specific for one type of biological information (genes or pathways or proteins, etc.). It is conceivable that CMAP Tools could become an important platform for information visualization for scientists/users across disciplines. Because CMAP software lacks strict concept typing, it is possible to readily integrate nutrigenomic/nutrigenetic concepts from several sources, including literature, pathway, gene, polymorphism, and disease resources. Yet perhaps also the greatest weakness of

CMAPS, as revealed in this study, is this lack distinction between conceptual classes and individual concepts within classes. Ontological reasoners compute over classes of information, inferring meaning about the information they contain based on formally defined domain/range rules for inclusion into those classes. Practically, this means that this project prioritized ease of use and heterogeneous data integration over formalized ontology creation.

By associating genetic/genomic factors, their influence on nutrient requirements, and the potential influence of lifestyle factors including diet, physical activity, and environmental exposures, the CMAPs demonstrated in this paper represent a framework for development of formalized ontologies capable of defining metabolic and health phenotypes. Tools such as the one presented here are useful for hypothesis generation including, but not limited to, those pertaining to nutrient–gene interaction mechanisms, as well as potential nutritional consequences of, and appropriate dietary interventions for, specific genetic polymorphisms. Compiling knowledge from multiple data resources into CMAPs enables easy visualization of links between previously untested ideas. Hypotheses formed during this process can be further tested and evaluated using animal models or clinical trials. A next phase of this project includes standardization of concepts presented in this model into a formalized ontology, and alignment of this new vocabulary with existing biomedical ontologies. Future plans also include integration with other SNP annotation platforms such as SNAP (Li et al. 2007) as well as with GWAS metabotype datasets (Dharuri et al. 2013).

The gene-specific maps highlighted in this paper have strong interactions with nutrients known to be essential for the OCM pathway (Hall and Finkler 1965; Leclerc et al. 1998; Sauberlich 1980). Because these genes are integral parts of OCM, alterations in genetic information are thought to affect utilization of nutrients such as cobalamin,

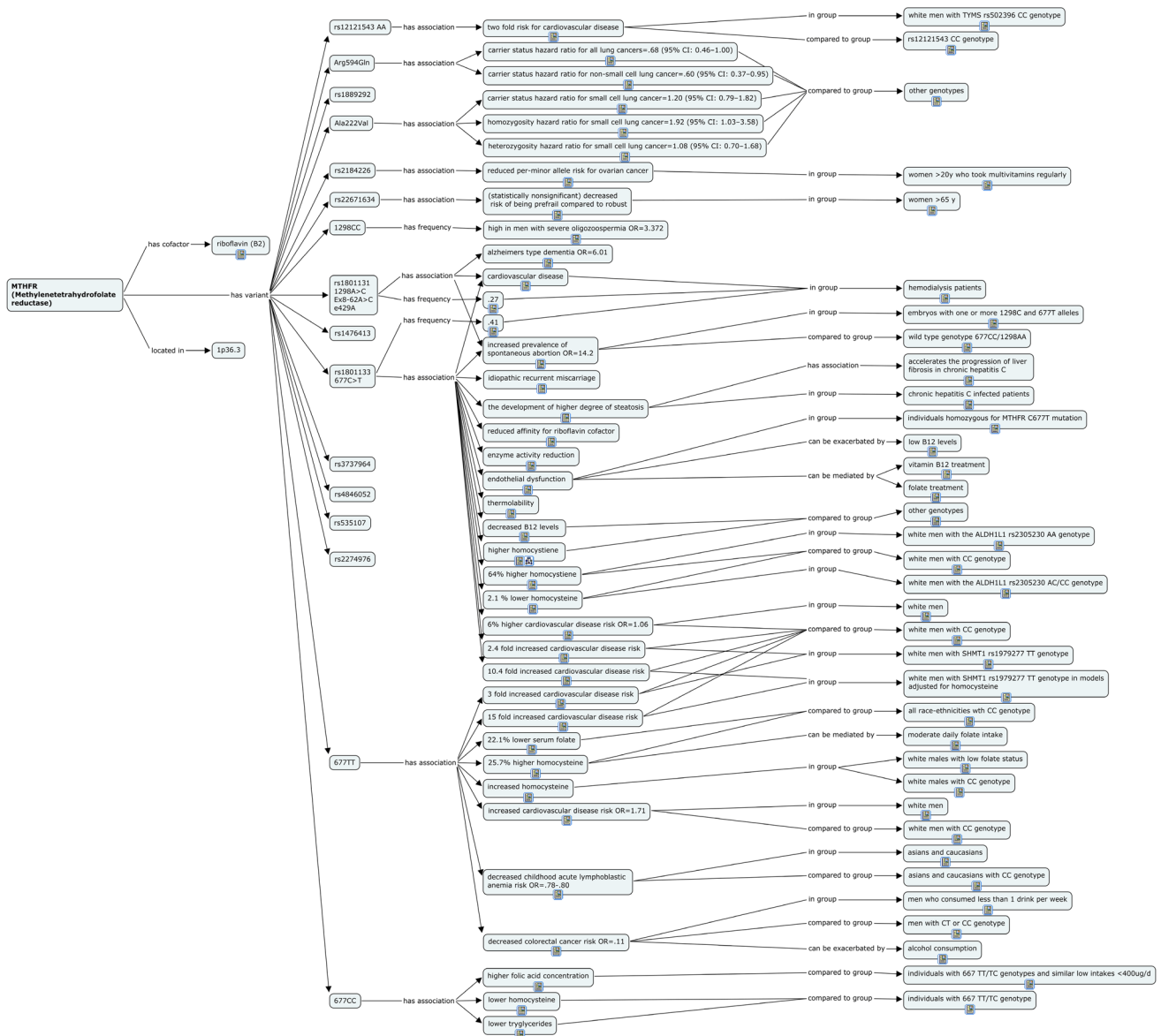


**Fig. 3** One-Carbon Metabolism CMAP. This map depicts the culmination of all genes mined from literature and pathway databases. Genes found exclusively in KEGG are colored *purple*, genes found exclusively in WikiPathways are *yellow*, and genes found exclusively in the GO are colored *pink*. Those found in two or more of these

databases are *light blue*. The genes found during literature mining are *not colored* and have *dotted lines* leading from the One-Carbon Metabolism subject through the linking phrase and down to the genes. The HumanCyc database was used as a reference for delineating the precise involvement of each gene product in OCM reactions

folate, and riboflavin (Garrod et al. 2010; Herbert and Zalusky 1962; Silla et al. 2011; Yamada et al. 2001; Yang et al. 2008). MTR (Fig. 5) requires vitamin B12 as a cofactor for its conversion of homocysteine to methionine and conversion of 5-methyltetrahydrofolate to its active form, tetrahydrofolate. It has long been established that activity of this enzyme is partially determined by availability of its cofactor vitamin B12 (Gulati et al. 1999). MTRR (Fig. 6) reduces cob(II)alamin to cob(I)alamin for use by MTR (Wolthers and Scrutton 2007). TCN2 (Fig. 7) encodes transcobalamin, a vitamin B12 transporter critical for proper distribution of vitamin B12 to cells throughout

the body. Some SNPs in TCN2 change this transporter’s affinity for the vitamin. Altered TCN2 affinity for B12 may influence delivery of B12 to tissues, as well as levels of circulating homocysteine, thereby affecting risk of neural tube defects (Afman et al. 2002). MTHFR (Fig. 4) utilizes a riboflavin cofactor to convert 5,10-methylene-tetrahydrofolate into 5-methyltetrahydrofolate. A common polymorphism in this gene is the rs1801133 (677C>T) mutation which causes enzyme instability and thermolability. This SNP has been associated with elevated circulating homocysteine (Unfried et al. 2002; Yang et al. 2008). Homocysteine levels can be reduced in individuals

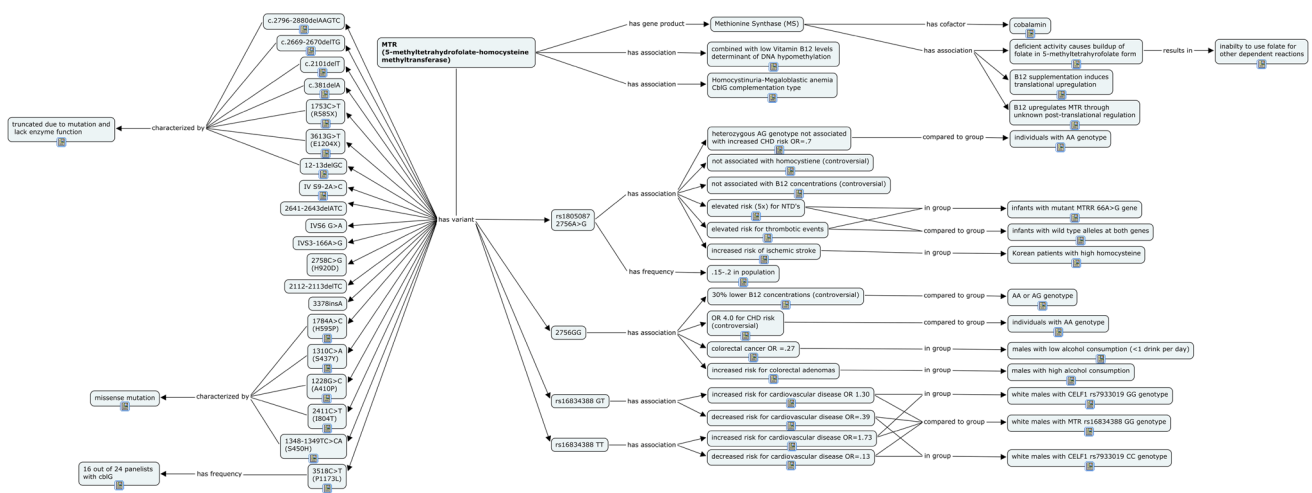


**Fig. 4** MTHFR CMAP. Conceptual map of the MTHFR gene, its polymorphisms, and literature-mined associations

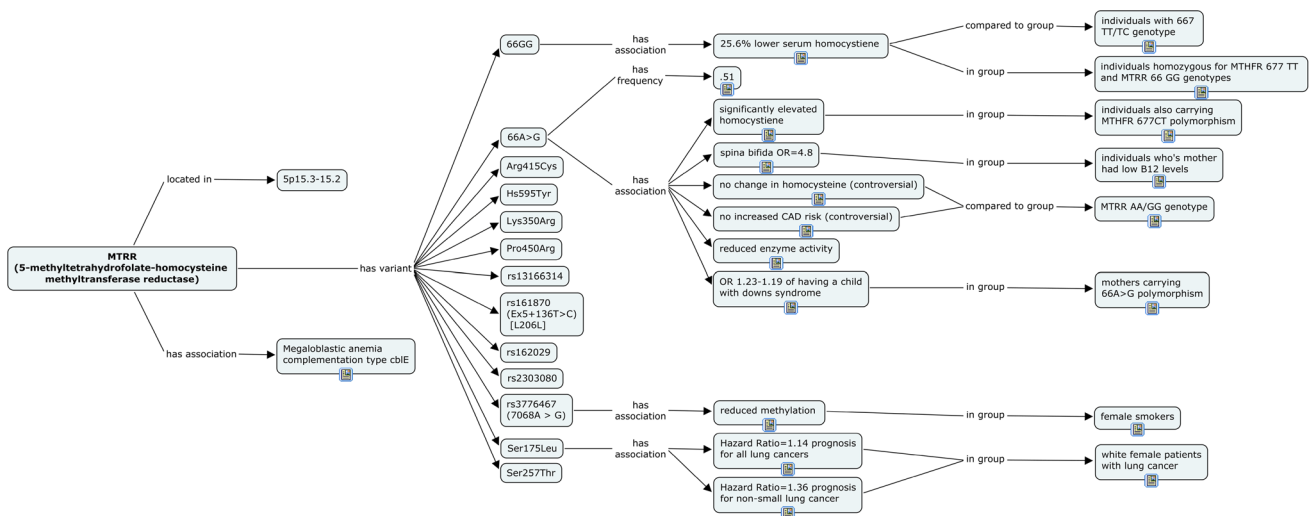
homozygous for this mutation with supplementation of riboflavin or folate (McNulty 2005; Yang et al. 2008). As resources develop, with the incorporation of genotype information into the electronic medical record, CMAPs can act as a tool to guide clinicians and scientists through evidence-based nutrient gene interactions, guiding them toward resolution of higher-order questions about genotype–phenotype associations and potential dietary interventions. Continued development and refinement of information resources such as these will provide clinicians with the ability to evaluate patient health based on current evolving evidence-based knowledge about specific SNPs or groups of SNPs, environments, and lifestyles (Taboada et al. 2012). The conceptual maps presented here highlight

the varied ways individuals possess genetic risk factors for nutrient metabolism. It is our hope that these maps will be useful in guiding development of hypotheses for testing alterations of diet and lifestyle, including dietary supplementation of vitamins for people with specific genetic predispositions. As this knowledge base grows to include more information on results from diet and lifestyle intervention studies, it is conceivable that food producers and dietitians will have the ability to leverage this expanding resource as a guide in developing personalized foods and diet plans.

The maps made for this project focus primarily on genetic contributions to phenotype. It is important to consider that in some cases, incorporation of genetic risk data



**Fig. 5** MTR CMAP. Conceptual map of the MTR gene, its polymorphisms, and literature-mined associations

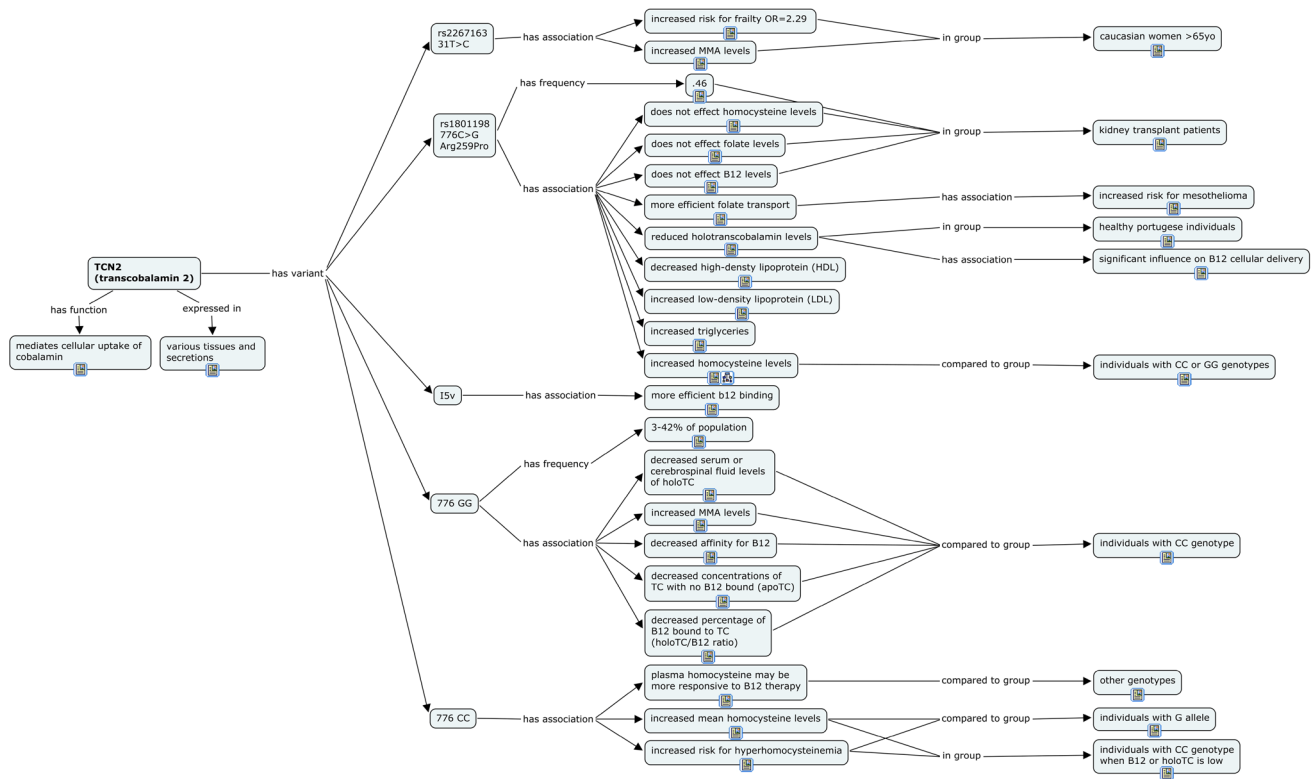


**Fig. 6** MTRR CMAP. Conceptual map of the MTRR gene, its polymorphisms, and literature-mined associations

will only moderately increase the ability to predict disease susceptibility (Meigs et al. 2008). In order to correctly determine overall risk of disease, all factors contributing to phenotype must be considered. Critical to the development of health and lifestyle personalization systems will be the development of algorithms to quantify the relative contribution to various health states of each SNP, constellations of SNPs, and environmental/behavioral exposures. As research reveals the contributions of each factor, visualization and computational tools must adapt and incorporate this knowledge while maintaining human readability. While single SNPs alone may not confer measurable phenotypic changes, the additive effect of constellations of SNPs may alter metabolism. For example, it has been shown that MTRR 66 AG or GG genotypes may additively exacerbate the effect of the MTHFR 677 TT mutation and

significantly increase homocysteine levels compared with other genotypes (Vaughn et al. 2004; Meigs et al. 2008).

As genetic tests become more affordable and are marketed directly to consumers, awareness of risk factors will help individuals manage their risks by engaging in, or avoiding certain behaviors. For example, homocysteine levels can be exacerbated by nutritionally related aberrations such as folic acid or vitamin B12 deficiency and lifestyle choices such as smoking (Garrod et al. 2010; Nygård et al. 1995). Homocysteine levels can also be mediated through nutritional therapy or physical activity (Yang et al. 2008; Garrod et al. 2010; Nygård et al. 1995; Wernimont et al. 2011). A combination of personalized nutritional therapy and lifestyle modifications could be suggested to individuals susceptible to hyperhomocysteinemia. Individuals and health information providers would



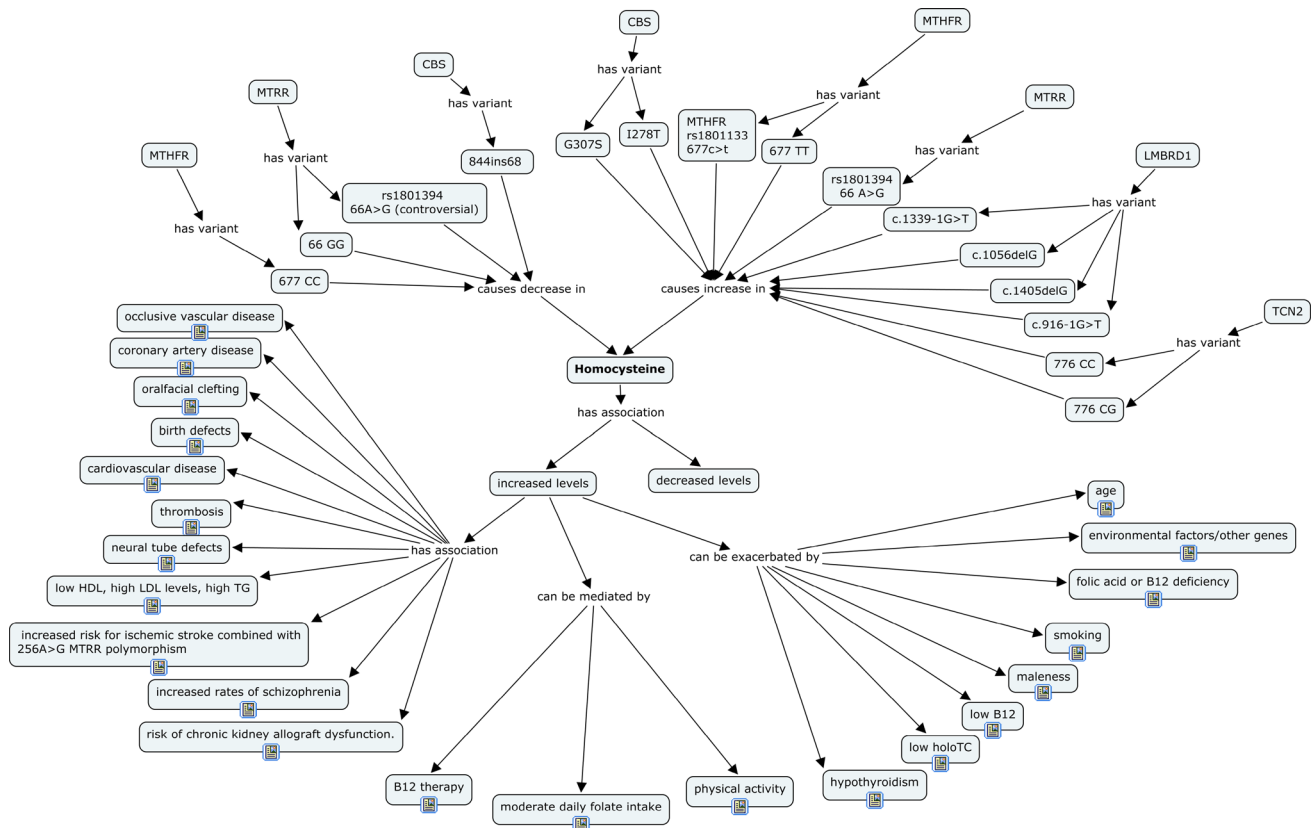
**Fig. 7** TCN2 CMAP. Conceptual map of the TCN2 gene, its polymorphisms, and literature-mined associations

benefit from a rapid, clear yet deep understanding of all factors contributing to altered homocysteine levels as well as other risk factors or diseases.

Figure 8 provides an example visualization that may guide scientists to develop hypotheses about the relative contribution of each component (genetic, environmental, and behavioral) to hyperhomocysteinemia. One genetic example from the literature-mined group of genes that has an effect on homocysteine levels is cystathionine beta synthase (CBS). The CBS gene exemplifies ways in which alterations of genes not directly involved in OCM can nevertheless modify OCM products and substrates. This particular gene codes for the enzyme that catalyzes the first step in the transsulfuration pathway of homocysteine metabolism. The enzyme requires vitamin B6 in the form of pyridoxal-5-phosphate (PLP). Mutations and SNPs in CBS, even among heterozygotes, result in elevated levels of homocysteine that varies from mild to severe. Homozygosity or compound heterozygosity for these variants is associated with classical homocystinuria and severe, often fatal, premature vascular disease. Heterozygosity may confer a risk for post-methionine load hyperhomocysteinemia with increased risk of vascular disease (Clarke et al. 1991). However, in the case of one particular insertional mutation in CBS, a protective effect toward hyperhomocysteinemia was observed in heterozygous individuals

(Yakub et al. 2012). Low levels of homocysteine are also observed in individuals with Down syndrome who have trisomy of chromosome 21. CBS is located on chromosome 21; therefore, these patients have three copies of the CBS gene (Fillon-Emery et al. 2004).

It has become clear, after mining literature and pathway databases that each differ in visual appearance, as well as breadth and depth of information coverage. Pathway databases are relatively devoid of information on important upstream and downstream regulators important for substrate availability and product metabolism. Some literature sources include genes from alternate pathways but, when this annotation process occurs, the genes were often grouped together by authors, and universally labeled as OCM genes. We attempted to combine and modify these approaches in order to include genes in pathways one degree of freedom away from the OCM pathway. This includes genes that transport cofactors, provide substrates, or utilize OCM products (such as CBS). Dysregulation in upstream pathways resulting from genetic mutations or inappropriate nutrient consumption not only changes the efficacy of the pathway under investigation (in this case OCM), it also alters downstream effects. The model presented here provides a method of displaying information not only about the pathway under investigation, but also relevant genes, substrates, and cofactors within upstream



**Fig. 8** Homocysteine CMAP. Conceptual map of genetic, environmental and lifestyle influences known to affect circulating levels of homocysteine, and several associated diseases known to be affected by hyperhomocysteinemia (at clinical and subclinical levels)

biochemical pathways and subsequent dependent pathways.

A crucial part of this process is expert curation. Like WikiPathways and other Wiki resources, data integrated into these maps should be curate-able. This way, mistakes in data aggregation can be mitigated and expert knowledge can be integrated into the pathway database for others to access, edit, and update.

### Next steps

Semi-automation along with the ability to query for specific genes, SNPs, or disease information will make information resources, such as the one proposed here, valuable to users. Users, e.g., healthcare professionals, would find such information important as an aid to answering questions concerning the probability of a given individual's genotype exhibiting any particular phenotype (Taboada et al. 2012). This would prove useful for prediction of disease susceptibility. Conversely, this type of CMAP would also facilitate the identification of possible underlying genotypes in an individual with a particular phenotype. The practical usefulness of this information lies

in attributing disease manifestations to genetic factors and for purposes of genetic counseling. Also, dietitians and other nutrition professionals can use this information to better diagnose and potentially correct metabolic abnormalities and using and personalized nutrition (Kaput 2008). Ongoing studies are developing tools that are capable of integrating automation and querying ability as well as incorporation of this information with other ontologies, such as the Reactome, to broaden the scope of information. Incorporation of alternative variants such as chromosomal rearrangements, insertions, and deletions, as well as copy number variants is necessary to paint a more comprehensive picture of how variation in human genetic information affects metabolism. Importantly, epistatic interactions and gene–environment interactions will need to be addressed to best determine relative risk. Development of methods for CMAP integration with ontology editing environments will provide the building blocks to develop a more formalized ontology for the information gathered manually and through semi-automated means. Automation will allow for near real-time updates from structured information repositories such as OMIM, KEGG, WikiPathways, and PubMed for fluid incorporation of newly discovered gene–phenotype information, into accessible, computable CMAPs.

**Acknowledgments** The authors acknowledge support from the following NIH grants: R01HD059127, R01HD065122, R01HD061923, R21AT006180, R01AT007079.

**Conflict of interest** A. C. Joslin, R. Green, J. B. German, and M. C. Lange declare that they have no conflict of interest.

**Ethical standard** This article does not contain any studies with human or animal subjects performed by the any of the authors.

## References

- Afman LA, Lievers KJA, van der Put NMJ, Trijbels FJM, Blom HJ (2002) Single nucleotide polymorphisms in the transcobalamin gene: relationship with transcobalamin concentrations and risk for neural tube defects. *Eur J Hum Genet* 10:433–438. doi:10.1038/sj.ejhg.5200830
- Amorim MR, Lima MAC (2012) MTRR 66A>G polymorphism as maternal risk factor for down syndrome: a meta-analysis. *Genet Test Mol Biomarkers*. doi:10.1089/gtmb.2012.0200
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. *Nat Genet* 25:25–29. doi:10.1038/75556
- Bailey LB, Gregory JF (1999) Polymorphisms of methylenetetrahydrofolate reductase and other enzymes: metabolic significance, risks and impact on folate requirement. *J Nutr* 129:919–922
- Beedholm-Ebsen R, van de Wetering K, Hardlei T, Nexø E, Borst P, Moestrup SK (2010) Identification of multidrug resistance protein 1 (MRP1/ABCC1) as a molecular gate for cellular export of cobalamin. *Blood* 115:1632–1639. doi:10.1182/blood-2009-07-232587
- Blanton SH, Henry RR, Yuan Q, Mulliken JB, Stal S, Finnell RH, Hecht JT (2011) Folate pathway and nonsyndromic cleft lip and palate. *Birth Defects Res A Clin Mol Teratol* 91:50–60. doi:10.1002/bdra.20740
- Brunaud L, Alberto J-M, Ayav A, Gérard P, Namour F, Antunes L, Braun M, Bronowicki J-P, Bresler L, Guéant J-L (2003) Vitamin B12 is a strong determinant of low methionine synthase activity and DNA hypomethylation in gastrectomized rats. *Digestion* 68:133–140
- Carmel R, Watkins D, Goodman SI, Rosenblatt DS (1988) Hereditary defect of cobalamin metabolism (cblG mutation) presenting as a neurologic disorder in adulthood. *N Engl J Med* 318:1738–1741. doi:10.1056/NEJM198806303182607
- Cascalheira JF, João SS, Pinhanções SS, Castro R, Palmeira M, Almeida S, Faria MC, Domingues FC (2009) Serum homocysteine: interplay with other circulating and genetic factors in association to Alzheimer's type dementia. *Clin Biochem* 42:783–790. doi:10.1016/j.clinbiochem.2009.02.006
- Castro AG, Rocca-Serra P, Stevens R, Taylor C, Nashar K, Ragan MA, Sansone S-A (2006) The use of concept maps during knowledge elicitation in ontology development processes—the nutrigenomics use case. *BMC Bioinformatics* 7:267. doi:10.1186/1471-2105-7-267
- Clarke R, Daly L, Robinson K, Naughten E, Cahalane S, Fowler B, Graham I (1991) Hyperhomocysteinemia: an independent risk factor for vascular disease. *N Engl J Med* 324:1149–1155. doi:10.1056/NEJM199104253241701
- Dharuri H, Henneman P, Demirkan A, van Klinken JB, Mook-Kanamori DO, Wang-Sattler R, Gieger C, Adamski J, Hettne K, Roos M, Suhre K, Duijn CMV, van Dijk KW, 't Hoen PA (2013) Automated workflow-based exploitation of pathway databases provides new insights into genetic associations of metabolite profiles. *BMC Genom* 14:865. doi:10.1186/1471-2164-14-865
- Fillon-Emery N, Chango A, Mircher C, Barbé F, Bléhaut H, Herbeth B, Rosenblatt DS, Réthoré M-O, Lambert D, Nicolas JP (2004) Homocysteine concentrations in adults with trisomy 21: effect of B vitamins and genetic polymorphisms. *Am J Clin Nutr* 80:1551–1557
- Garrod MG, Allen LH, Haan MN, Green R, Miller JW (2010) Transcobalamin C776G genotype modifies the association between vitamin B12 and homocysteine in older Hispanics. *Eur J Clin Nutr* 64:503–509. doi:10.1038/ejcn.2010.20
- Green T, Chen X, Ryan S, Asch AS, Ruiz-Echevarría MJ (2013) TMEFF2 and SARDH cooperate to modulate one-carbon metabolism and invasion of prostate cancer cells. *Prostate* 73:1561–1575. doi:10.1002/pros.22706
- Guéant J-L, Chabi NW, Guéant-Rodríguez R-M, Mutchinick OM, Debard R, Payet C, Lu X, Villaume C, Bronowicki J-P, Quadros EV, Sanni A, Amouzou E, Xia B, Chen M, Anello G, Bosco P, Romano C, Arrieta HR, Sánchez BE, Romano A, Herbeth B, Anwar W, Namour F (2007) Environmental influence on the worldwide prevalence of a 776C->G variant in the transcobalamin gene (TCN2). *J Med Genet* 44:363–367. doi:10.1136/jmg.2006.048041
- Gulati S, Brody LC, Banerjee R (1999) Posttranscriptional regulation of mammalian methionine synthase by B12. *Biochem Biophys Res Commun* 259:436–442. doi:10.1006/bbrc.1999.0696
- Hall CA, Finkler AE (1965) The dynamics of transcobalamin II. A vitamin b12 binding substance in plasma. *J Lab Clin Med* 65:459–468
- Haviv YS, Shpichinetsky V, Goldschmidt N, Atta IA, Ben-Yehuda A, Friedman G (2002) The common mutations C677T and A1298C in the human methylenetetrahydrofolate reductase gene are associated with hyperhomocysteinemia and cardiovascular disease in hemodialysis patients. *Nephron* 92:120–126
- Hazra A, Kraft P, Selhub J, Giovannucci EL, Thomas G, Hoover RN, Chanock SJ, Hunter DJ (2008) Common variants of FUT2 are associated with plasma vitamin B12 levels. *Nat Genet* 40:1160–1162. doi:10.1038/ng.210
- Herbert V, Zalusky R (1962) Interrelations of vitamin B12 and folic acid metabolism: folic acid clearance studies. *J Clin Invest* 41:1263–1276. doi:10.1172/JCI104589
- Kaput J (2008) Nutrigenomics research for personalized nutrition and medicine. *Curr Opin Biotechnol* 19:110–120
- Kelemen LE, Sellers TA, Schildkraut JM, Cunningham JM, Vierkant RA, Pankratz VS, Fredericksen ZS, Gadre MK, Rider DN, Liebow M, Goode EL (2008) Genetic variation in the one-carbon transfer pathway and ovarian cancer risk. *Cancer Res* 68:2498–2506. doi:10.1158/0008-5472.CAN-07-5165
- Kozyraki R, Kristiansen M, Silahtaroglu A, Hansen C, Jacobsen C, Tommerup N, Verroust PJ, Moestrup SK (1998) The human intrinsic factor-vitamin B12 receptor, cubilin: molecular characterization and chromosomal mapping of the gene to 10p within the autosomal recessive megaloblastic anemia (MGA1) region. *Blood* 91:3593–3600
- Kronenberger T, Lindner J, Meissner KA, Zimbres FM, Coronado MA, Sauer FM, Schetter I, Wrenger C (2014) Vitamin B6-dependent enzymes in the human malaria parasite *Plasmodium falciparum*: a druggable target? *BioMed Res Int*. doi:10.1155/2014/108516
- Leclerc D, Wilson A, Dumas R, Gafuik C, Song D, Watkins D, Heng HH, Rommens JM, Scherer SW, Rosenblatt DS, Gravel RA (1998) Cloning and mapping of a cDNA for methionine synthase reductase, a flavoprotein defective in patients with homocystinuria. *Proc Natl Acad Sci USA* 95:3059–3064

- Levine AJ, Figueiredo JC, Lee W, Conti DV, Kennedy K, Duggan DJ, Poynter JN, Campbell PT, Newcomb P, Martinez ME, Hopper JL, Le Marchand L, Baron JA, Limburg PJ, Ulrich CM, Haile RW (2010) A candidate gene study of folate-associated one carbon metabolism genes and colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 19:1812–1821. doi:10.1158/1055-9965.EPI-10-0151
- Li S, Ma L, Li H, Vang S, Hu Y, Bolund L, Wang J (2007) Snap: an integrated SNP annotation platform. *Nucl Acids Res* 35:D707–D710. doi:10.1093/nar/gkl969
- Li F, Feng Q, Lee C, Wang S, Pelleymounter LL, Moon I, Eckloff BW, Wieben ED, Schaid DJ, Yee V, Weinshilboum RM (2008) Human betaine–homocysteine methyltransferase (BHMT) and BHMT2: common gene sequence variation and functional characterization. *Mol Genet Metab* 94:326–335. doi:10.1016/j.ymgme.2008.03.013
- Li Q, Lan Q, Zhang Y, Bassig BA, Holford TR, Leaderer B, Boyle P, Zhu Y, Qin Q, Chanock S, Rothman N, Zheng T (2013) Role of one-carbon metabolizing pathway genes and gene–nutrient interaction in the risk of non-Hodgkin lymphoma. *Cancer Causes Control* 24:1875–1884. doi:10.1007/s10552-013-0264-3
- Lim U, Wang SS, Hartge P, Cozen W, Kelemen LE, Chanock S, Davis S, Blair A, Schenk M, Rothman N, Lan Q (2007) Gene–nutrient interactions among determinants of folate and one-carbon metabolism on the risk of non-Hodgkin lymphoma: NCI–SEER Case–Control Study. *Blood* 109:3050–3059. doi:10.1182/blood-2006-07-034330
- Lindblom A, Robinson PN (2011) Bioinformatics for human genetics: promises and challenges. *Hum Mutat* 32:495–500. doi:10.1002/humu.21468
- Matakidou A, El Galta R, Rudd MF, Webb EL, Bridle H, Eisen T, Houlston RS (2007) Prognostic significance of folate metabolism polymorphisms for lung cancer. *Br J Cancer* 97:247–252. doi:10.1038/sj.bjc.6603830
- McNulty H (2005) Riboflavin lowers homocysteine in individuals homozygous for the MTHFR 677C->T polymorphism. *Circulation* 113:74–80. doi:10.1161/CIRCULATIONAHA.105.580332
- Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, Manning AK, Florez JC, Wilson PWF, D’Agostino RB, Cupples LA (2008) Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 359:2208–2219. doi:10.1056/NEJMoA0804742
- Morine Y, Shimada M, Utsunomiya T, Imura S, Ikemoto T, Hanaka J, Kanamoto M, Kurita N, Miyake H (2012) Role of thymidylate synthase and dihydropyrimidine dehydrogenase mRNA in intrahepatic cholangiocarcinoma. *Surg Today* 42:135–140. doi:10.1007/s00595-011-0054-z
- Namour F, Dobrovoljski G, Chery C, Audonnet S, Feillet F, Sperl W, Gueant J-L (2011) Luminal expression of cubilin is impaired in Imerslund–Grasbeck syndrome with compound AMN mutations in intron 3 and exon 7. *Haematologica* 96:1715–1719. doi:10.3324/haematol.2011.043984
- Nygård O, Vollset SE, Refsum H, Stensvold I, Tverdal A, Nordrehaug JE, Ueland M, Kvåle G (1995) Total plasma homocysteine and cardiovascular risk profile. The Hordaland Homocysteine Study. *JAMA* 274:1526–1533
- O’Byrne MR, Au KS, Morrison AC, Lin J-I, Fletcher JM, Ostermaier KK, Tyerman GH, Doebel S, Northrup H (2010) Association of folate receptor (FOLR1, FOLR2, FOLR3) and reduced folate carrier (SLC19A1) genes with meningomyelocele. *Birth Defects Res A Clin Mol Teratol* 88:689–694. doi:10.1002/bdra.20706
- Oppeneer SJ, Ross JA, Koh W-P, Yuan J-M, Robien K (2012) Genetic variation in folypolyglutamate synthase and gamma-glutamyl hydrolase and plasma homocysteine levels in the Singapore Chinese Health Study. *Mol Genet Metab* 105:73–78. doi:10.1016/j.ymgme.2011.09.035
- Pangilinan F, Mitchell A, VanderMeer J, Molloy AM, Troendle J, Conley M, Kirke PN, Sutton M, Sequeira JM, Quadros EV, Scott JM, Mills JL, Brody LC (2010) Transcobalamin II receptor polymorphisms are associated with increased risk for neural tube defects. *J Med Genet* 47:677–685. doi:10.1136/jmg.2009.073775
- Romero P, Wagg J, Green ML, Kaiser D, Krummenacker M, Karp PD (2004) Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol* 6:R2. doi:10.1186/gb-2004-6-1-r2
- Rosenblatt DS, Thomas IT, Watkins D, Cooper BA, Erbe RW (1987) Vitamin B12 responsive homocystinuria and megaloblastic anemia: heterogeneity in methylcobalamin deficiency. *Am J Med Genet* 26:377–383. doi:10.1002/ajmg.1320260216
- Rutsch F, Gailus S, Suormala T, Fowler B (2011) LMBRD1: the gene for the cblF defect of vitamin B<sub>12</sub> metabolism. *J Inherit Metab Dis* 34:121–126. doi:10.1007/s10545-010-9083-9
- Sauberlich HE (1980) Interactions of thiamin, riboflavin, and other B-vitamins. *Ann NY Acad Sci* 355:80–97
- Seetharam B, Bose S, Li N (1999) Cellular import of cobalamin (vitamin B-12). *J Nutr* 129:1761–1764
- Shields DC, Kirke PN, Mills JL, Ramsbottom D, Molloy AM, Burke H, Weir DG, Scott JM, Whitehead AS (1999) The “thermolabile” variant of methylenetetrahydrofolate reductase and neural tube defects: an evaluation of genetic risk and the relative importance of the genotypes of the embryo and the mother. *Am J Hum Genet* 64:1045–1055
- Silla Y, Chandamouli B, Maiti S, Sengupta S (2011) A single nucleotide polymorphism in transcobalamin II (I5 V) induces structural changes in the protein as revealed by molecular modeling studies. *Biochemistry* 50:1396–1402. doi:10.1021/bi101095g
- Smith B, Ceusters W, Klagges B, Köhler J, Kumar A, Lomax J, Mungall C, Neuhaus F, Rector AL, Rosse C (2005) Relations in biomedical ontologies. *Genome Biol* 6:R46. doi:10.1186/gb-2005-6-5-r46
- Stover PJ, Garza C (2002) Bringing individuality to public health recommendations. *J Nutr* 132:2476S–2480S
- Szegedi SS, Castro CC, Koutmos M, Garrow TA (2008) Betaine–homocysteine S-methyltransferase-2 is an S-methylmethionine–homocysteine methyltransferase. *J Biol Chem* 283:8939–8945. doi:10.1074/jbc.M710449200
- Taboada M, Martínez D, Pilo B, Jiménez-Escrig A, Robinson PN, Sobrido MJ (2012) Querying phenotype–genotype relationships on patient datasets using semantic web technology: the example of cerebrotendinous xanthomatosis. *BMC Med Inform Decis Mak* 12:78. doi:10.1186/1472-6947-12-78
- Unfried G, Griesmacher A, Weismüller W, Nagele F, Huber JC, Tempfer CB (2002) The C677T polymorphism of the methylenetetrahydrofolate reductase gene and idiopathic recurrent miscarriage. *Obstet Gynecol* 99:614–619
- Vaughn JD, Bailey LB, Shelnett KP, Von-Castel Dunwoody KM, Maneval DR, Davis SR, Quinlivan EP, Gregory JF III, Theriaque DW, Kauwell GPA (2004) Methionine synthase reductase 66A->G polymorphism is associated with increased plasma homocysteine concentration when combined with the homozygous methylenetetrahydrofolate reductase 677C->T variant. *J Nutr* 134:2985–2990
- Vlastos F, Hillas G, Vidal P, Lacomme S, Galateau-Sallé F, Vollmer E, Guzman-Costabel J, Vignaud JM, Martinet N (2009) Survey and biological insights of pemetrexed-related therapeutic improvement in mesothelioma: The Nancy Centre of Biological Resources’ Mesothelioma Cohort. *J Thorac Oncol* 4:1259–1263. doi:10.1097/JTO.0b013e3181aba6bd
- Ward M, Wilson CP, Strain JJ, Horigan G, Scott JM, McNulty H (2011) B-vitamins, methylenetetrahydrofolate reductase (MTHFR) and hypertension. *Int J Vitam Nutr Res* 81:240–244. doi:10.1024/0300-9831/a000069



- Watkins D, Ru M, Hwang H-Y, Kim CD, Murray A, Philip NS, Kim W, Legakis H, Wai T, Hilton JF, Ge B, Doré C, Hosack A, Wilson A, Gravel RA, Shane B, Hudson TJ, Rosenblatt DS (2002) Hyperhomocysteinemia due to methionine synthase deficiency, cblG: structure of the MTR gene, genotype diversity, and recognition of a common mutation, P1173L. *Am J Hum Genet* 71:143–153
- Wernimont SM, Clark AG, Stover PJ, Wells MT, Litonjua AA, Weiss ST, Gaziano JM, Tucker KL, Baccarelli A, Schwartz J, Bollati V, Cassano PA (2011) Folate network genetic variation, plasma homocysteine, and global genomic methylation content: a genetic association study. *BMC Med Genet* 12:150. doi:10.1186/1471-2350-12-150
- Whetzel PL, Brinkman RR, Causton HC, Fan L, Field D, Fostel J, Fragoso G, Gray T, Heiskanen M, Hernandez-Boussard T, Morrison N, Parkinson H, Rocca-Serra P, Sansone S-A, Schober D, Smith B, Stevens R, Stoeckert CJ, Taylor C, White J, Wood A (2006) Development of FuGO: an ontology for functional genomics investigations. *OMICS* 10:199–204. doi:10.1089/omi.2006.10.199
- Woeller CF, Fox JT, Perry C, Stover PJ (2007) A ferritin-responsive internal ribosome entry site regulates folate metabolism. *J Biol Chem* 282:29927–29935. doi:10.1074/jbc.M706264200
- Wolthers KR, Scrutton NS (2007) Protein interactions in the human methionine synthase–methionine synthase reductase complex and implications for the mechanism of enzyme reactivation. *Biochemistry* 46:6696–6709. doi:10.1021/bi700339v
- Yakub M, Moti N, Parveen S, Chaudhry B, Azam I, Iqbal MP (2012) Polymorphisms in MTHFR, MS and CBS genes and homocysteine levels in a Pakistani population. *PLoS ONE* 7:e33222. doi:10.1371/journal.pone.0033222
- Yamada K, Chen Z, Rozen R, Matthews RG (2001) Effects of common polymorphisms on the properties of recombinant human methylenetetrahydrofolate reductase. *Proc Natl Acad Sci USA* 98:14853–14858. doi:10.1073/pnas.261469998
- Yang Q-H, Botto LD, Gallagher M, Friedman JM, Sanders CL, Koontz D, Nikolova S, Erickson JD, Steinberg K (2008) Prevalence and effects of gene–gene and gene–nutrient interactions on serum folate and serum total homocysteine concentrations in the United States: findings from the third National Health and Nutrition Examination Survey DNA Bank. *Am J Clin Nutr* 88:232–246
- Zittan E, Preis M, Asmir I, Cassel A, Lindenfeld N, Alroy S, Halon DA, Lewis BS, Shiran A, Schliamser JE, Flugelman MY (2007) High frequency of vitamin B12 deficiency in asymptomatic individuals homozygous to MTHFR C677T mutation is associated with endothelial dysfunction and homocysteinemia. *Am J Physiol Heart Circ Physiol* 293:H860–H865. doi:10.1152/ajpheart.01189.2006