Chapter 13
Functional Genomics of Selenoproteins and Se-responsive Pathways

Catherine Méplan and John Hesketh

Abstract Nutrigenomics approaches have contributed to our understanding of how selenium impacts metabolic pathways, homeostatic control and disease risk. In this chapter, we discuss the known genetic polymorphisms in genes encoding selenoproteins and components of the selenoprotein synthesis machinery. Furthermore, the consequences of these genetic variants on the synthesis and activity of individual selenoproteins within the context of the metabolic pathway in which they are involved, including the impact of these variants on the overall selenoproteome, as a result of the shared synthesis machinery between selenoproteins are discussed. The evidence for the association of these genetic variants with several chronic diseases is presented, with a specific emphasis on functional variants.

Keywords Association study • Cancer • Nutrigenomics • Selenoprotein • Selenoprotein hierarchy • Single nucleotide polymorphisms

13.1 Introduction

Selenoproteins play pivotal roles in many biochemical pathways and in particular in the response to oxidative stress and endoplasmic reticulum (ER) stress. In a nutritional context, it is important to consider the effect of selenium (Se) intake on the overall balance of selenoproteins and related pathways; similarly, from a genetic perspective, the effects of inter-individual variations need to be assessed in the overall Se pathway. Here, we discuss how genetic studies are providing novel perspectives about the mechanisms by which Se intake and metabolism are related to disease risk.

Single Nucleotide Polymorphisms (SNPs) represent 90 % of the genetic variations among individuals, and some genetic variants alter gene or protein regulation or protein activity and are thus functionally relevant. During evolution, the ability of our ancestors to survive relied on their capacity to combat metabolic stresses and

C. Méplan (✉) • J. Hesketh
School of Biomedical Sciences, Faculty of Medical Sciences, Newcastle University, The Medical School, Framlington Place, Newcastle-upon-Tyne NE2 4HH, UK

Human Nutrition Research Centre and Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle-upon-Tyne, NE2 4HH, UK
e-mail: Catherine.meplan@ncl.ac.uk; j.e.hesketh@ncl.ac.uk

D.L. Hatfield et al. (eds.), *Selenium*, DOI 10.1007/978-3-319-41283-2_13

infections and ensure their reproduction; these functions are strongly supported by various selenoproteins and therefore highly dependent on Se intake. It is possible that genetic variants which modulate selenoprotein synthesis or activity in different conditions of Se supply could have been selected as a result of the disparate geological distribution of Se and exposure to various pathogens and stresses. Thus, the pressure of selection imposed by environmental forces may partially explain the variations in allelic frequencies currently observed among different populations. Supporting this hypothesis, several studies have identified SNPs in selenoprotein genes corresponding to signatures of recent genetic adaptation, suggesting a local adaptation to low Se content in the soil. In Asian populations living in Se-deficient regions, it was reported that a selective sweep occurred recently at the *GPX1* locus [1]. Similarly, allele frequency shifts for SNPs in selenoprotein genes were observed in populations living in the severe Se-deficient regions of China [2]. Moreover, a recent study identified a strong positive natural selection in individuals of European decent, for known functional variants in *GPX1*, *SELENBP1*, *GPX3, GPX2* and *SELO* genes suggesting local adaptation to low soil Se [3].

With the increase in human lifespan, genetic variations in Se metabolism have been proposed to be associated with risks for several age-related diseases, including cancer, diabetes, immunological or neurological disorders and cardiovascular diseases. Many of these disorders share a common basis for their development such as impaired cellular maintenance mechanisms and responses to stress. These metabolic pathways involve many selenoproteins. Thus, genetic factors affecting selenoprotein activity or synthesis, have the potential to modify individual risk to chronic disease (Fig. 13.1).

Three types of studies have shown links between genetic variants in selenoprotein genes and disease susceptibility (Tables 13.1, 13.2, 13.3, and 13.4): (1) the study of a small number of candidate SNPs of known functional relevance in which it is assumed that the functional consequences of the SNP on the gene/protein regulations could contribute to the disease risk; (2) the broader study of SNPs in several selenoprotein genes or screening of genes across the Se metabolic pathways, usually using a combination of tagSNPs and functional variants; and (3) large scale genome-wide association studies (GWAS) in which selenoprotein genes were associated with a disease trait. Focussed studies of well-characterized SNPs provide mechanistic backing to epidemiological studies and also often allow measurements of biomarkers of Se status and stress, particularly important in the case of Se since intake varies across different populations [4]. Screening studies involving a wider range of SNPs regardless of known functionality and large-scale GWAS have the advantage of wide genetic coverage but are less likely to have environmental and dietary data available.

13.2 Selenium Metabolism and Transport

Selenoprotein synthesis depends upon Se intake, Se transport from the liver to other organs, Se conversion to selenocysteine (Sec) and its incorporation into selenoproteins (Chaps. 2 and 4). Genetic variants in two genes involved in Sec synthesis (Table 13.4),

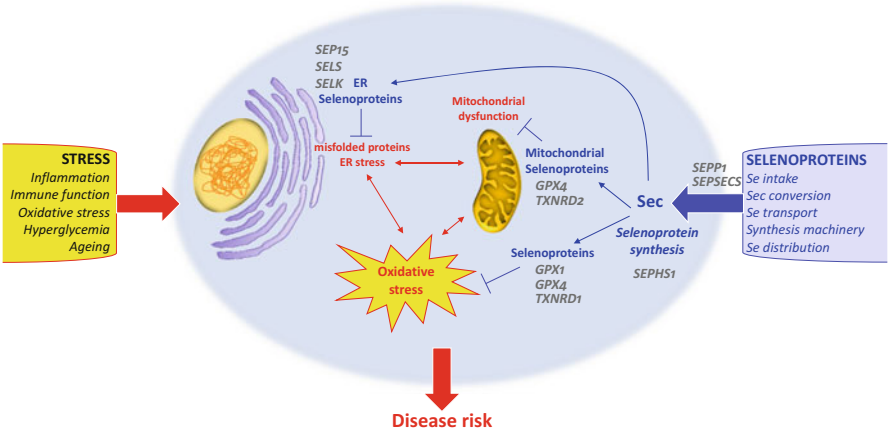

Fig. 13.1 Schematic diagram illustrating interactions of selenoprotein SNPs and Se supply on selenoprotein function and disease risk. The diagram illustrates how SNP-Se interactions could affect selenoprotein synthesis and activity and consequently downstream biochemical pathways known to be crucial in the response to oxidative and ER (endoplasmic reticulum) stress and in the maintenance of mitochondrial redox status. As disruption of these biochemical pathways is commonly observed in many chronic disorders, the existence of genetic associations linking these SNP with diseases indicate the importance of selenoproteins' function in these pathways. Genes (*dark grey*, *italic*) in which functional SNPs have been associated with disease risk are indicated at subcellular location in which the corresponding selenoproteins function is required

Sec tRNA synthase *(SEPSECS)* and selenophosphate synthetase 1 *(SEPHS1)*, have been linked to Crohn's Disease in a Caucasian population from New Zealand with suboptimal Se intake [5, 6]. Moreover, SNPs in *SEPP1*, the gene coding for selenoprotein P (SePP, SEPP1), which transports hepatic Se to other tissues, have been shown to affect Se delivery, Sec incorporation throughout the body and disease risk [7]. In particular, two genetic variants, rs3877899 and rs7579 (*SEPP1*), affect the plasma SePP isoform pattern, response to Se supplementation in healthy individuals with sub-optimal Se intake and selenoprotein synthesis [8, 9]. In support of this, expression of SePP and GPx1 was affected by both variants in patients with mild cognitive impairment supplemented with Se [10]. Moreover, rs3877899 and rs7579 were associated with risk for several cancers (Table 13.3) in different European populations [7, 11–13] and various *SEPP1* variants were linked to colorectal (CRC) [14] or breast (BC) cancer in Native Americans [15]. In Europeans, risk of prostate cancer (PCA) or advanced PCA was linked to rs7579 (*SEPP1*) [12, 13] and modulated by an interaction between rs3877899 (*SEPP1)* and rs4880, a known functional SNP in *SOD2* (manganese superoxide dismutase) [16]. In a US population, rs13168440 (*SEPP1)* was also shown to affect PCA risk [17], but other studies have failed to link SNPs in *SEPP1* to PCA [13, 18, 19] or disease recurrence [20]. In addition, both rs3877899 and rs7579 (*SEPP1*) were also shown to influence body mass index, blood pressure, peripheral arterial

Table 13.1 Genetic association of SNPs in the *GPX1* gene with disease risk[a]

SNP	Base change	Cases/controls	Target/location	Functionality	Population	Association	Reference
Breast cancer							
rs1050450	C>T	1038/1088	Pro198Leu	Enzyme activity Pro>Leu	USA	None	[62]
		1229/1629			USA	None	[61]
		79/517			USA	T allele: ↑ BC risk	[46]
		399/372			Canada	None	[63]
		2293/2278			UK	None	[86]
		4371/0			UK	No association with BC survival risk	[41]
		377/377			Denmark	T allele: ↑ BC risk	[47]
		933/959			Denmark	T allele: ↑ non-ductal BC; interaction with rs3877899(SEPP1); ↓ eGPx activity	[11]
Prostate cancer							
rs1050450	C>T	745/0	Pro198Leu	Enzyme activity pro>Leu	USA	None	[87]
		500/1391			USA	None	[88]
		247/487			Germany	T allele: ↓ PCA risk with ↑ serum Se levels	[12]
		82/123			Macedonia	T allele: ↓ PCA risk	[50]
		262/435			New Zealand	T allele: ↑PCA risk	[51]
rs1800668	C>T	951/25408	TagSNP/promoter	TagSNP/high LD with rs1050450	Netherlands	TT: ↓advanced (stage III/IV) PCA risk	[13]
rs17650792	A>G	952/25426	TagSNP/promoter	unknown	Netherlands	GG: ↑advanced (stage III/IV) PCA risk	[13]

rs1050450	C>T	656/743	Pro198Leu	Enzyme activity pro>Leu	USA	No association with advanced distal colorectal adenoma	[14]
		981/397			Norway	None	[89]
		375/779			Denmark	None	[90]
		832/705			Czech Republic	No association SNP alone, but genetic interaction with rs37413471 (*SELS*)	[7]
		827/733			Korea	None	[76]
Lung cancer							
rs1050450	C>T	237/234	Pro198Leu	Enzyme activity pro>Leu	USA	CC :↑ lung cancer in smokers>80 years of age	[52]
		315/313			Finland/men	T allele: ↑ risk	[53]
		95/176			Poland	T allele: ↓ risk	[40]
		432/798			Denmark	T allele: ↓ risk	[54]
		186/207			Germany	T allele: ↓ risk	[91]
Laryngeal cancer							
rs1050450	C>T	111/213	Pro198Leu	Enzyme activity pro>Leu	Poland	T allele: ↓ risk	[40]
Bladder cancer							
rs1050450	C>T	224/0	Pro198Leu	Enzyme activity Pro>Leu	USA	T allele: ↑ bladder cancer recurrence risk	[56]
		213/209			Japan	T allele: ↑ risk	[55]

(continued)

Table 13.1 (continued)

SNP	Base change	Cases/controls	Target/location	Functionality	Population	Association	Reference
Cardiovascular disease							
rs1050450	C>T	184/0	Pro198Leu	Enzyme activity pro>Leu	Japan (diabetic)	T allele: ↑ CVD risk in diabetic patients and ↑ intima-media thickness	[92]
Kashin-Beck disease							
rs1050450	C>T	638/324	Pro198Leu	Enzyme activity pro>Leu	China	None	[93]

[a]The table presents results from association studies between functional and tagSNPs in the *GPX1* gene and disease risk. The disease, SNP, allele or genotype associated with disease risk or progression is indicated together with the studied population and the known functional consequences of the SNP on the protein function or expression

Table 13.2 Genetic association of SNPs in the *GPX4* gene with disease risk[a]

SNP	Base change	Cases/controls	Target/location	Functionality	Population	Association	Reference
Breast cancer							
rs713041	C>T	2182/2264	3′UTR, near SECIS	Sec-insertion efficiency C>T	UK	None	[86]
		4356			UK	T allele: ↑ risk of mortality by BC	[41]
		939/960			Denmark	T allele: ↓ eGPx activity	[11]
Prostate cancer							
rs713041	C>T	739/0	3′UTR, near SECIS	Sec-insertion efficiency C>T	USA	None	[87]
		245/490			Germany	None	[12]
		260/439			New Zealand	None	[51]
Colorectal cancer							
rs713041	C>T	745/758	3′UTR, near SECIS	Sec-insertion efficiency C>T	USA	No association with advanced distal colorectal adenoma	[14]
		252/187			UK	TT:↓ CRC risk	[36]
		832/705			Czech Republic	CT: ↑ CRC risk; interaction with rs4880 (*SOD2*), rs9605031 (*TXNRD2*) and rs3877899 (*SEPP1*)	[7]
		827/733			Korea	None	[76]

(continued)

Table 13.2 (continued)

SNP	Base change	Cases/controls	Target/location	Functionality	Population	Association	Reference
Lung cancer							
rs713041	C>T	95/176	3′UTR, near SECIS	Sec-insertion efficiency C>T	Poland	T allele: ↓ risk	[40]
Laryngeal cancer							
rs713041	C>T	325/287	3′UTR, near SECIS	Sec-insertion efficiency C>T	Poland	T allele: ↓ risk	[40]
Kashin-Beck disease							
rs713041	C>T	219/194	3′UTR, near SECIS	Sec-insertion efficiency C>T	China	None; ↓ GPX4 mRNA level in KBD patients	[39]
haplotype rs713041 -rs4807542		219/194			China Han	Haplotype A-T: ↓ KBD risk	[39]

[a]The table presents results from association studies between functional and tagSNPs in the *GPX4* gene and disease risk. The disease, SNP, allele or genotype associated with disease risk or progression is indicated together with the studied population and the known functional consequences of the SNP on the protein function or expression

Table 13.3 Genetic association of SNPs in the *SEPP1* gene with disease risk[a]

SNP	Base change	Cases/controls	Target/location	Functionality	Population	Association	Reference
Breast cancer							
rs3877899	G>A	937/957	Ala234Thr	Plasma SePP isoforms, Se bioavailability	Denmark	AA:↓ BC and ductal BC risk	[11]
rs7579	G>A	937/957	3′UTR	Plasma SePP isoforms, Se bioavailability	Denmark	None	[11]
rs230812	A>C		TagSNP		Native American	CC: ↑ BC risk in women of 28-70 % Native American ancestry Strong LD with rs230813 (associated with oxidative stress)	[15]
rs6865453	A>C		TagSNP			In LD with rs7579 AC/CC: ↓BC risk in women of 28-70 % Native American ancestry	[15]
Prostate cancer							
rs3877899	G>A	2643/1570	Ala234Thr	Plasma SePP isoforms, Se bioavailability	Sweden	None	[16]
		248/492			Germany	None	[12]

(continued)

Table 13.3 (continued)

SNP	Base change	Cases/controls	Target/location	Functionality	Population	Association	Reference
		951/25409			Netherlands	Genotype interacts with Se status: ↓advanced PCA risk	[13]
		259/436			New Zealand	None	[51]
rs7579	G>A	248/492	3′UTR	Plasma SePP isoforms, Se bioavailability	Germany	AA :↑ PCA risk; interaction with plasma SePP	[12]
		951/25408			Netherlands	A allele: ↓advanced (stage IV) PCA risk; genotype interacts with plasma Se status to ↓advanced PCA risks	[13]
rs13168440	T>C		TagSNP	Unknown	USA	C allele: interacts with plasma Se to ↓PCA risk	[17]
Colorectal cancer							
rs3877899	G>A	193/127	Ala234Thr	Plasma SePP isoforms, Se bioavailability	Germany	None	[94]
		832/705			Czech Republic	No association SNP alone, but interaction with rs5859(*SEP15)* and rs713041 (*GPX4)*	[7]
		827/733			Korea	None	[76]

rs7579	G>A	832/705	3′UTR	Plasma SePP isoforms, Se bioavailability	Czech Republic	AA:↑ CRC risk, interaction with rs5859 (*SEP15*)	[7]
		827/733			Korea	None	[76]
Promoter (−4166), Exon 5 (rs3877899, rs6413428), 3′UTR (rs12055266, rs2972994, rs3797310)		772/777			USA	Association global *SEPP1* variants with advanced distal colorectal adenoma	[14]
Type 2 diabetes							
rs28919926, rs146125471, rs16872779, rs7579		2446			Hispanics, European American, African American	Associated with fasting insulin and first phase insulin response	[25]

[a]The table presents results from association studies between functional and tagSNPs in the *SEPP1* gene and disease risk. The disease, SNP, allele or genotype associated with disease risk or progression is indicated together with the studied population and the known functional consequences of the SNP on the protein function or expression

Table 13.4 Genetic association of SNPs in other selenoprotein genes with disease risk[a]

Gene symbol	SNP	Base change	Cases/ controls	Target/ location	Functionality	Population	Association	References
Prostate cancer								
SEP15	rs5859	G>A	1195/1186	3′UTR	Sec-insertion efficiency	USA	None	[84]
			248/492			Germany	AA :↓ GPX3 activity	[12]
SEP15	rs5845	G>A or C>T	259/436	3′UTR	Sec-insertion efficiency	New Zealand	AA ↑ PCA risk	[51]
SEP15	rs561104	G>A	1195/1186	TagSNP	Unknown	USA	AA: ↑risk of mortality by PCA	[84]
SELK	rs9880056	T>C	248/492	TagSNP	Unknown	Germany	C allele: interacts with serum SePP & serum Se to ↓ risk advanced and high grade PCA	[19]
TXNRD1	rs7310505	C>A	248/492	TagSNP	Unknown	Germany	CC: interacts with serum SePP & serum GPX activity to ↑ risk of advanced PCA	[19]
TXNRD2	rs9605030	C>T	248/492	TagSNP	Unknown	Germany	T allele: interacts with serum Se status to ↑ high grade PCA risk	[19]
TXNRD2	rs9605031	C>T	248/492	TagSNP	Unknown	Germany	T allele: interacts with serum Se status to ↓ high grade PCA risk	[19]
Colorectal cancer								
GPX2	rs4902347	G>A	570/762	TagSNP	Unknown	USA	GA/AA: ↓risk of rectal cancer (but not colon or adenoma)	[70]

GPX3	rs3828599	C>T	582/773	TagSNP	Unknown	USA	T allele: ↓risk of rectal cancer (but not colon or adenoma)	[70]
	rs736775	C>T	582/773	TagSNP	Unknown	USA	T allele: ↓risk of rectal cancer (but not colon or adenoma)	[70]
	rs8177447	C>T	582/773	TagSNP	Unknown	USA	T allele: ↓risk of rectal cancer (but not colon or adenoma)	[70]
SEP15	rs5859		832/705	3′UTR, SECIS	Sec-insertion efficiency	Czech Republic	no association SNP alone, but interaction with rs3877899, rs7579, rs3797310, rs12055266 in *SEPP1*	[7]
			827/733			Korea	A allele:↑ CRC risk	[76]
SEP15	rs5845		827/733	3′UTR	Sec-insertion efficiency	Korea	None	[76]
	rs35009941		772/777	C>G		USA	G allele:↓ CRC, alone and in association with rs34195484, rs4077561, rs1128446, rs5018287, rs6539137,rs10778322 and rs35776976 in *TXNRD1*	[14]

(continued)

Table 13.4 (continued)

Gene symbol	SNP	Base change	Cases/ controls	Target/ location	Functionality	Population	Association	References
TXNRD1	rs35009941	C>G	772/777			USA	G allele:↓ CRC, alone and in association with rs34195484, rs4077561, rs1128446, rs5018287, rs6539137, rs10778322 and rs35776976 in *TXNRD1*	[14]
Lung cancer								
SEP15	rs5859		325/287	3′UTR, SECIS	Sec-insertion efficiency	Poland	A allele: ↑risk in individuals with low Se status	[85]
SEP15	rs5845		325/287	3′UTR	Sec-insertion efficiency	Poland	None	[40]
Laryngeal cancer								
SEP15	rs5845		325/287	3′UTR	Sec-insertion efficiency	Poland	None	[40]
Cardiovascular disease								
SELS	rs28665122 rs4965814,rs2862845, rs7178239			TagSNPs		European Americans/ diabetic	Associated with measures of vascular calcification in European American families enriched for type 2 diabetes	[95]

Crohn's disease								
SEPHS1	rs7901303	G>T	351/853	TagSNP	Unknown	New Zealand-Caucasians	SNP-Serum Se interaction affecting Crohn's disease risk	[5]
	rs17529609	A>G	351/853	TagSNP	Unknown	New Zealand-Caucasians	SNP-Serum Se interaction affecting Crohn's disease risk	[5]
SEPSECS	rs1553153	G>A	351/853	TagSNP	Unknown	New Zealand-Caucasians	SNP-Serum Se interaction affecting Crohn's disease risk	[5]
Type 2 diabetes								
IDI2	rs225014		721	Thr92Ala		Brazil	Ala variant: less active, associated with type 2 diabetes, interaction with PPARγ2 Pro12Ala	[80–82]

[a]The table presents results from association studies between functional and tagSNPs in selenoprotein genes not presented in Tables 13.1, 13.2, and 13.3 and disease risk. The disease, SNP, allele or genotype associated with disease risk or progression is indicated together with the studied population and the known functional consequences of the SNP on the protein function or expression

disease occurrence and abdominal aortic aneurysm development in overweight and obese subjects [21].

Moreover, considerable mechanistic evidence suggest a role for SePP in glucose and insulin metabolism [22–24]. In support of this, four SNPs in *SEPP1* (rs28919926, rs146125471, rs168727790 and rs7579) were shown to be associated with an altered fasting or acute insulin responses in two different Hispanic cohorts [25].

It is interesting to note that the effects of genetic variants in factors involved in Sec conversion and transport on disease risk are often influenced by Se status and/or ethnicity (hence genetic background), consistent with evolutionary adaptation of populations to geographical differences in Se distribution.

13.3 The Effects of SNPs on the Sec Incorporation Machinery

Sec incorporation requires the recoding of a UGA codon for Sec and the binding of factors from the selenoprotein synthesis machinery to the Sec insertion sequence (SECIS) stem-loop structure within the 3′untranslated region (3′UTR) of selenoprotein mRNAs (reviewed in Chap. 2). In addition, SECIS-binding protein 2 (SBP2) plays a major role in this process.

The 3′UTR regions of selenoprotein mRNAs play a key regulatory role in the so-called selenoprotein hierarchy [26], a mechanism by which the synthesis of the various selenoproteins is affected differently by limiting conditions of Se supply [27, 28]. The hierarchy reflects the fact that all selenoproteins share the same incorporation machinery and the same tRNA carrying Sec for their synthesis and therefore there is in essence a competition between the 25 selenoprotein mRNAs for the available synthesis machinery and Sec. As a result, a genetic variant in the 3′UTR of one selenoprotein has the potential to affect synthesis, not only of the selenoprotein coded by that mRNA, but also synthesis of other selenoproteins. Similarly, a SNP in a gene coding for factors involved in the selenoprotein biosynthesis machinery have the potential to affect the synthesis of all selenoproteins.

In particular, mutations in the selenoprotein N gene (*SELN*) within the gene region corresponding to the SECIS region of the 3′UTR were associated with reduced binding affinity of SBP2 for SECIS, lower expression of SelN and congenital muscular dystrophy [29]. Moreover, missense mutations in *SBP2* led to poor Sec incorporation, poor thyroid function or muscular dystrophy, low expression of all selenoproteins and increased sensitivity to oxidative stress [30, 31]. However, mutations in selenoprotein genes that directly cause genetic disease are rare. On the contrary, common SNPs may have more subtle effects on selenoprotein metabolism, but in conjunction with other factors, such as dietary Se intake, can lead to altered risk for many diseases. Three functional genetic variants, one in *GPX4* (rs713041) [32] and two in *SEP15* (rs5859 and rs5845) [33], which induce base changes in a region nearby or within the SECIS element in corresponding transcripts, have been shown

to reduce the efficiency of Sec incorporation in the corresponding protein and to affect the selenoprotein hierarchy and disease risk [27, 33–35].

Originally identified as a C/T variant in the gene region corresponding to the 3′UTR region of *GPX4* mRNA [32], rs713041 illustrates how SNPs in selenoprotein genes can be studied from a functional perspective and how a SNP affecting the 3′UTR region impacts selenoprotein hierarchy. The C variant was shown to induce reporter gene expression to a greater extent than the T counterpart [36], to have a stronger binding affinity for the selenoprotein synthesis machinery in RNA-protein binding assays [35] and to alter the pattern of selenoprotein synthesis, especially during Se-depletion [36, 37]. Moreover, in healthy individuals, rs713041 was shown to affect expression of blood selenoproteins in response to Se supplementation, consistent with an effect on the selenoprotein hierarchy in vivo [35]. Finally, human umbilical vein endothelial cells from individual donors and monocytes [38] expressing the T-variant showed an increased expression of vascular cell adhesion protein 1 and adhesion to monocytes compared with cells from the CC individuals.

13.4 Genetic Variants Affecting Redox-Active Selenoproteins

Many selenoproteins are involved in the control of cellular redox balance and anti-oxidant defense and can be divided into two main classes of redox-active selenoenzymes, the *GPX* and thioredoxin reductases (*TXNRD)*. Functional SNPs in *GPX1-4* and *TXNRD1-2* have been shown to affect anti-oxidant defense and disease risk (Tables 13.1, 13.2, 13.3, and 13.4).

13.4.1 Genetic Variants in GPX4

As mentioned above, rs713041 (*GPX4*) was shown to affect the selenoprotein hierarchy [35], as well as the sensitivity to oxidative challenge [37, 38]. Interestingly, in a Se-deficient Chinese population (Table 13.2), rs713041 together with rs4807542, another *GPX4* SNP in high linkage disequilibrium with rs713041, were found to affect Kashin-Beck disease (KBD) risk, with *GPX4* mRNA levels being reduced in KBD patients compared with controls [39]. Moreover, rs713041 has been linked to risk of CRC in Czech and Scottish populations [7, 36], to lung and laryngeal cancers in a Polish population [40] and risk of BC mortality in a British population [41]. In the Czech population, CRC risk was affected by polymorphisms in both *SEPP1* (rs7579) and *GPX4* (rs713041) [7] and was further modulated by significant genetic interactions between SNPs in *SEPP1 (rs7579* and rs3877899), known to affect Se bioavailability, and variants in *SEP15 (rs*5859) or *GPX4* (rs713041), known to affect Sec incorporation [7]. These interactions suggest that in carriers of the combined genotypes, the altered pattern of synthesis of selenoproteins affects the

individual's ability to respond to stress. Similarly, genetic interactions between rs713041 in *GPX4,* rs960531 in *TXNRD2* and rs4880 in *SOD2* mirror the interactions of these enzymes in mitochondrial redox function and suggest that the genetic interactions could affect an individual's ability to counteract oxidative stress in the mitochondria [7]. Additionally, two independent GWAS linked the *GPX4* locus to Crohn's disease [42, 43], consistent with involvement of GPx4 in inflammatory responses and NF-kB regulation [44, 45].

13.4.2 Genetic Variants in GPX1

Cellular GPx1 is an important antioxidant enzyme in mammals. In humans, the enzyme activity is affected by rs1050450, a coding C/T SNP in the *GPX1* gene, inducing a Pro (CC) to Leu (TT) amino acid change at position 198 of the amino acid sequence. The Leu variant exhibits lower activity compared with the Pro counterpart [46, 47] and, during Se-supplementation, GPx1 activity is stimulated less in TT carriers compared with CC carriers [48, 49]. Moreover, significantly higher levels of DNA oxidation were observed in Leu carriers, probably as a result of reduced GPx1 activity [48, 49]. Surprisingly, however, an increased susceptibility to DNA strand breaks was observed in CC subjects, but not TT, during Se withdrawal [48].

Many studies have linked rs1050450 to risk for several disorders (Table 13.1), including various cancers [12, 13, 40, 46, 50–56], Alzheimer's disease [57], metabolic syndrome and obesity [58, 59], and KBD [60]. In particular, rs1050450 was linked to BC risk in some [46], but not other US populations [61–63], although a meta-analysis found an increased BC risk among women of African descent [64]. However, the association was replicated in several European populations [11, 47, 65]. The consistent differences observed between studies of European and North American populations suggest that the effect of rs1050450 on BC risk may be partially influenced by Se status. In a Danish population, the Leu variant was associated with reduced GPx1 activity, increased risk of BC and a higher grade of ductal tumors [11, 47]. Additionally, pre-diagnostic erythrocyte GPx1 activity was lower in Leu females, following hormone replacement therapy and who develop BC later in life, compared with controls [11]. Furthermore, a GCG repeat polymorphism in *GPX1,* resulting in variant protein sequences containing between 5 and 7 alanines, was linked to BC risk [63], supporting a role for GPx1 activity in protecting breast tissue from carcinogenesis.

Variants in the *GPX1* gene have also been linked to PCA, another cancer influenced by sex-hormones [13]. Moreover, rs1050450 together with rs18006688 (a *GPX1* SNP in high linkage disequilibrium) modified the association between lead exposure and glioblastoma [66], suggesting that the reduced GPx1 activity of the Leu variant could impair the protection against oxidative damage generated by lead [66]. Like *GPX4*, GWAs have linked the *GPX1* locus to inflammatory conditions such as Crohn's disease [42], inflammatory bowel disease [43] and ulcerative colitis [67].

13.4.3 Genetic Variants in TXNRDs

Mammalian cells have three isozymes of thioredoxin reductases, including cytoplasmic and nuclear TXNRD1 and mitochondrial TXNRD2. The relationship between diseases and genetic variants in *TXNRD1* and *TXNRD*2 has been investigated in studies using tagSNPs including a GWAs [68, 69], in a study investigating association of SNPs in several selenoprotein genes on PCA risk, grade and recurrence [20], and in a study investigating the association of polymorphic variants in Se metabolism with PCA risk [19]. In the latter study, carried out in a German population with low Se intake, tagSNPs in *SELK*, *TXNRD1* and *TXNRD*2 were found to interact with plasma Se or SePP status to modulate risk of advanced disease [19] (Table 13.4).

In addition, SNPs in other selenoproteins involved in the protection against oxidative damage were linked to disease. In a US population, three SNPs in *GPX3* and one in *GPX2* were significantly associated with risk of rectal cancer, but not with either colon cancer or adenoma [70] (Tables 13.3, and 13.4). In the *MSRA* gene, coding for methionine sulfoxide reductase A, rs10903323, was associated with coronary artery disease risk in a Chinese population [71].

13.5 Genetics of Endoplasmic Reticulum Selenoproteins

In eukaryotic cells, the ER is not only responsible for the synthesis, post-translational modification and correct folding of membrane and secreted proteins, but also for intracellular Ca^{2+} homeostasis and lipid biosynthesis [72]. Alteration of protein folding caused by changes in intracellular Ca^{2+} levels, redox state, nutrient status, protein synthesis rate or inflammatory stimuli, can result in ER stress and activation of the unfolded protein response to remove misfolded proteins [72]. ER dysfunction and prolonged ER stress have been implicated in diseases such as cancer, diabetes and Alzheimer's disease [72]. Seven human selenoproteins have been shown to be associated with the ER, which are the 15-kDa selenoprotein (Sep15), type 2 iodothyronine deiodinase and selenoproteins K, M, N, S, and T [73]. Selenoprotein S is a component of the ER-associated protein degradation pathway and is involved in the removal of misfolded proteins from the ER lumen [73]. Sep15 has been implicated in the formation of disulfide bonds and the quality control of protein folding [34], while SelK plays a key role in calcium signalling in immune cells [74].

Genetic polymorphisms in *SELS, SELK* and *SEP15* have now been linked to various cancers and inflammatory conditions, consistent with these SNPs affecting correct ER function (Table 13.4). In *SELS* gene, rs34713741 affects levels of pro-inflammatory cytokines, IL-6, IL-1β and TNF-α [75], and risk of CRC in Korean [76] and Czech [7] populations. In the Czech population, rs34713741 was associated with greater CRC risk, and in the Korean population a second variant in close proximity led to increased risk (in females only). The replication of the association

in these two diverse populations strongly indicates that, independently of lifestyle and dietary factors, SNPs in *SELS* influence CRC risk. Moreover, supporting a role of SelS in gastrointestinal function, rs34713741, was also linked to gastric cancer risk [77, 78]. No association was identified between six genetic variants in *SELS* and type 1 diabetes [79]. In contrast, rs225014 in (type 2 deiodinase gene), results in a Thr to Ala amino-acid change at position 92 in the protein sequence, with the Ala variant being less active and associated with type 2 diabetes [80–82].

In the *SEP15* gene, rs5859 and rs5845, known to reduce Sec incorporation efficiency and Sep15 synthesis, affect BC risk in a Se-dependent manner in African American women, but not in Caucasians [33]. In breast tumors, a loss of heterozygosity at the *SEP15* locus was observed [33, 83], indicating a potential tumor suppressor role of Sep15 in breast tissue. In addition, rs5859 and rs5845 were associated with increased risk of PCA in German [12] and New Zealand [51] males with low Se intake, and other SNPs in *SEP15* were also linked to PCA mortality (Table 13.4) [20, 84]. CRC risk was also affected by rs5859 in a South Korean population [76] and by genetic interactions between SNPs in *SEPP1* and rs5859 in *SEP15* [7]. Moreover, rs5859 was linked to lung cancer in Polish individuals with low Se status [85]. The influence of these SNPs on disease risk and progression provides insight into the potential role of ER stress and protein folding control in disease etiology as well as on the importance of key selenoproteins in maintaining a healthy tissue.

13.6 Perspectives

Two main lessons can be learned from the studies described above. First, there is now considerable evidence for a number of genetic variants affecting selenoprotein synthesis and, in a limited number of cases, this has been linked to effects on response to dietary Se intake. Genetic epidemiology studies suggest that a number of these variants, notably those in *SEPP1, SELS*, *GPX1*, *GPX4* and *SEP15*, modulate risk of various chronic diseases. In particular, these results suggest that variants in *SEPP1, SEP15* and *GPX1* affect PCA risk and progression, SNPs in *SELS* influence CRC risk and variants in *GPX1* BC risk. However, the evidence linking these variants with disease risk comes largely from relatively small studies that often lack accompanying measures of Se status and often require replication. In addition, observed effects are often inconsistent between study populations, a phenomenon that likely reflects differences in the characteristics of the study populations, notably Se status. Further work is necessary to be confident of the clinical relevance of selenoprotein SNPs in different populations, both in terms of Se intake and ethnicity. Larger studies, combining genetics and biomarkers of Se status should provide a clearer picture of the links between Se, selenoprotein genetics and disease risk.

Second, identifying that SNPs in selenoprotein genes affect risk for several chronic diseases is compatible with the observation that most of these multifactorial diseases share a common basis, with the disruption of biochemical pathways involved

in the response to oxidative and ER stress. As a result, these genetic associations help elucidate potential pathways affected in the etiology of disease for an individual. Selenoprotein synthesis depends on the distribution of Se between selenoproteins and the use of common synthesis machinery. Thus, the observation of genetic associations of a SNP with a disease may reflect an effect of this particular SNP on the whole selenoproteome, inviting us to combine genetic association studies with expression of selenoproteins in a tissue.

References

1. LH Foster, S Sumar 1997 *Crit Rev Food Sci Nutr* 37:211
2. L White et al 2015 *Mol Biol Evol* 32:1507
3. J Engelken et al 2016 *Mol Biol Evol* 33:738
4. GF Combs, Jr. 2001 *Br J Nutr* 85:517
5. L Gentschew et al 2012 *Nutrients* 4:1247
6. CD Thomson et al 2007 *Br J Nutr* 97:357
7. C Meplan et al 2010 *Carcinogenesis* 31:1074
8. C Meplan et al 2009 *Antioxid Redox Signal* 11:2631
9. C Meplan et al 2007 *FASEB J* 21:3063
10. BR Cardoso et al 2016 *Food Funct* 7:825
11. C Meplan et al 2013 *PLoS ONE* 8:e73316
12. A Steinbrecher et al 2010 *Cancer Epidemiol Biomarkers Prev* 19:2958
13. MS Geybels et al 2014 *J Natl Cancer Inst* 106:dju003
14. U Peters et al 2008 *Cancer Epidemiol Biomarkers Prev* 17:1144
15. AJ Pellatt et al 2013 *PLoS ONE* 8
16. ML Cooper et al 2008 *Cancer Res* 68:10171
17. KL Penney et al 2013 *Prostate* 73:700
18. MS Geybels et al 2015 *Cancer Epidemiol Biomarkers Prev* 24:178
19. C Meplan et al 2012 *PLoS ONE* 7:e48709
20. JP Gerstenberger et al 2015 *Prostate* 75:60
21. E Strauss et al 2014 *Sci Rep* 4:7061
22. H Misu et al 2010 *Cell Metab* 12:483
23. SJ Yang et al 2011 *J Clin Endocrinol Metab* 96:E1325
24. H Steinbrenner 2013 *Free Radic Biol Med* 65:1538
25. JN Hellwege et al 2014 *Gene* 534:33
26. FP Bellinger et al 2009 *Biochem J* 422:11
27. G Bermano et al 1996 *Biochem J* 320 (Pt 3):891
28. V Pagmantidis et al 2005 *FEBS Lett* 579:792
29. V Allamand et al 2006 *EMBO Rep* 7:450
30. E Schoenmakers et al 2010 *J Clin Invest* 120:4220
31. E Schoenmakers et al 2016 in *Diversity of Selenium Funcion in Health and Disease,* R Brigelius-Flohé, H Sies Eds (CRC Press, Boca Raton, FL) p 343
32. S Villette et al 2002 *Blood Cells Mol Dis* 29:174
33. YJ Hu et al 2001 *Cancer Res* 61:2307
34. KV Korotkov et al 2001 *J Biol Chem* 276:15330
35. C Meplan et al 2008 *Am J Clin Nutr* 87:1019
36. G Bermano et al 2007 *Genes Nutr* 2:225
37. H Gautrey et al 2011 *Biochim Biophys Acta* 1810:284

38. LK Crosley et al 2013 *Mol Nutr Food Res* 57:2185
39. XH Du et al 2012 *Br J Nutr* 107:164
40. K Jaworska et al 2013 *PLoS ONE* 8:e59051
41. M Udler et al 2007 *J Clin Oncol* 25:3015
42. A Franke et al 2010 *Nature Genetics* 42:1118
43. L Jostins et al 2012 *Nature* 491:119
44. H Imai, Y Nakagawa 2003 *Free Radic Biol Med* 34:145
45. G Gong et al 2012 *Genes Nutr* 7:167
46. YJ Hu, AM Diamond 2003 *Cancer Res* 63:3347
47. G Ravn-Haren et al 2006 *Carcinogenesis* 27:820
48. E Jablonska et al 2015 *Eur J Nutr* doi:10.1007/s00394-015-1118-4
49. JC Miller et al 2012 *American Journal of Clinical Nutrition* 96:923
50. Z Arsova-Sarafinovska et al 2009 *Int Urol Nephrol* 41:63
51. N Karunasinghe et al 2012 *J Nutrigenet Nutrigenomics* 5:339
52. P Yang et al 2004 *Carcinogenesis* 25:1935
53. D Ratnasinghe et al 2000 *Cancer Res* 60:6381
54. O Raaschou-Nielsen et al 2007 *Cancer Lett* 247:293
55. Y Ichimura et al 2004 *J Urol* 172:728
56. H Zhao et al 2005 *Urology* 66:769
57. C Paz-y-Mino et al 2010 *Am J Med Sci* 340:373
58. M Kuzuya et al 2008 *Am J Clin Nutr* 87:1939
59. C Cominetti et al 2011 *Nutrition* 27:891
60. YM Xiong et al 2010 *Osteoarthritis Cartilage* 18:817
61. DG Cox et al 2004 *Cancer Epidemiol Biomarkers Prev* 13:1821
62. J Ahn et al 2005 *Cancer Epidemiol Biomarkers Prev* 14:2459
63. JA Knight et al 2004 *Cancer Epidemiol Biomarkers Prev* 13:146
64. J Hu et al 2010 *Breast Cancer Res Treat* 124:425
65. E Jablonska et al 2015 *BMC Cancer* 15:657
66. P Bhatti et al 2009 *Cancer Epidemiol Biomarkers Prev* 18:1841
67. A Julia et al 2014 *Hum Mol Genet* 23:6927
68. Z Ye et al 2014 *Front Genet* 5:125
69. EN Vithana et al 2012 *Nat Genet* 44:1142
70. U Haug et al 2012 *Genes Chromosomes Cancer* 51:598
71. H Gu et al 2013 *Clin Biochem* 46:1668
72. M Wang, RJ Kaufman 2014 *Nat Rev Cancer* 14:581
73. VA Shchedrina et al 2010 *Antioxid Redox Signal* 12:839
74. Z Huang et al 2011 *J Biol Chem* 286:34830
75. JE Curran et al 2005 *Nat Genet* 37:1234
76. A Sutherland et al 2010 *Genes Nutr* 5:215
77. T Shibata et al 2009 *BMC Gastroenterol* 9:2
78. H Mao et al 2015 *Int J Clin Exp Med* 8:10993
79. A Martinez et al 2008 *BMC Genomics* 9:329
80. LH Canani et al 2005 *J Clin Endocrinol Metab* 90:3472
81. JM Dora et al 2010 *Eur J Endocrinol* 163:427
82. AA Estivalet et al 2011 *Obesity (Silver Spring)* 19:825
83. MA Nasr et al 2003 *Cancer Therapy* 1:293
84. KL Penney et al 2010 *Cancer Prev Res (Phila)* 3:604
85. E Jablonska et al 2008 *Eur J Nutr* 47:47
86. A Cebrian et al 2006 *Cancer Res* 66:1225
87. M Abe et al 2011 *BJU Int* 107:126
88. JY Choi et al 2007 *Cancer Epidemiol Biomarkers Prev* 16:1115

89. R Hansen et al 2005 *Cancer Lett* 229:85
90. RD Hansen et al 2009 *Mutat Res* 664:13
91. A Rosenberger et al 2008 *BMC Cancer* 8:60
92. T Hamanishi et al 2004 *Diabetes* 53:2455
93. L Huang et al 2013 *PLoS ONE* 8:e71411
94. OH Al-Taie et al 2004 *Nutr Cancer* 48:6
95. AJ Cox et al 2013 *Acta Diabetol* 50:391