

Supplementary Table 1: List of sulfur-metabolizing bacteria¹⁴

| | |
|---|--|
| <i>Acidaminococcus unclassified</i> | <i>Eubacterium rectale</i> |
| <i>Adlercreutzia equolifaciens</i> | <i>Gordonibacter pamelaee</i> |
| <i>Alistipes finegoldii</i> | <i>Lachnospiraceae bacterium 5 1 63FAA</i> |
| <i>Alistipes putredinis</i> | <i>Odoribacter splanchnicus</i> |
| <i>Anaerotruncus colihominis</i> | <i>Oxalobacter formigenes</i> |
| <i>Bacteroides clarus</i> | <i>Parabacteroides distasonis</i> |
| <i>Bacteroides intestinalis</i> | <i>Parabacteroides goldsteinii</i> |
| <i>Bacteroides ovatus</i> | <i>Parabacteroides johnsonii</i> |
| <i>Bacteroides plebeius</i> | <i>Parabacteroides merdae</i> |
| <i>Bacteroides stercoris</i> | <i>Parabacteroides unclassified</i> |
| <i>Bacteroides uniformis</i> | <i>Paraprevotella clara</i> |
| <i>Bacteroides vulgatus</i> | <i>Paraprevotella xylaniphila</i> |
| <i>Bilophila unclassified</i> | <i>Parasutterella excrementihominis</i> |
| <i>Bilophila wadsworthia</i> | <i>Roseburia intestinalis</i> |
| <i>Burkholderiales bacterium 1 1 47</i> | <i>Ruminococcus bromii</i> |
| <i>Clostridiales bacterium 1 7 47FAA</i> | <i>Streptococcus australis</i> |
| <i>Coprococcus catus</i> | <i>Streptococcus vestibularis</i> |
| <i>Desulfovibrio desulfuricans</i> | <i>Sutterella wadsworthensis</i> |
| <i>Eggerthella lenta</i> | <i>Veillonella atypica</i> |
| <i>Erysipelotrichaceae bacterium 21 3</i> | <i>Veillonella parvula</i> |
| <i>Erysipelotrichaceae bacterium 6 1 45</i> | <i>Veillonella unclassified</i> |
| <i>Eubacterium ramulus</i> | |

Species-level identification of microbes suspected to be involved in dietary sulfur metabolism based on prior experimental evidence, as well as the presence of genes encoding for at least two sulfur-metabolizing enzymes.

Supplementary Table 2: Components of the sulfur microbial diet pattern (Health Professionals Follow-up Study)¹⁴

| Sulfur microbial diet pattern components with the top factor loadings | Representative foods | β-coefficients (SE) |
|--|--|---|
| <i>Positive associations</i> | | |
| Processed meats | Processed meats, bacon, hot dogs | 0.64 (0.25) |
| Liquor | Vodka, gin | 0.31 (0.11) |
| Low-calorie drinks | Low-calorie cola, other low-energy carbonated beverages | 0.38 (0.11) |
| <i>Negative associations</i> | | |
| Beer | Beer | -0.54 (0.17) |
| Fruit juice | Apple juice or cider, orange juice, grapefruit juice, other fruit juice | -0.21 (0.12) |
| Legumes | String beans, peas or lima beans, beans or lentils, tofu or soybeans, alfalfa sprouts | -0.64 (0.19) |
| Other vegetables | Celery, mushrooms, green pepper, corn, mixed vegetables, eggplant, summer squash | -0.30 (0.10) |
| Sweets & desserts | Chocolate bars or pieces, candy bars, cookies, brownies, doughnuts, cake, pie, sweet roll, coffee cake, pastries | -0.23 (0.10) |

Abbreviations: SE (standard error)

Supplementary Table 3. Participant characteristics at study midpoint in the Nurses' Health Study II (NHSII), 2001

| | Sulfur microbial diet score (quartile) | | | |
|---|--|---------------|---------------|-------------------------|
| | 1 (lowest) (n=1395) | 2 (n=1396) | 3 (n=1396) | 4 (highest) (n=1396) |
| Age, years* | 46 (3.3) | 46.1 (3.2) | 46.1 (3.2) | 45.7 (3.4) |
| Height, cm | 164.9 (6.6) | 164.7 (6.5) | 165 (6.6) | 165 (7) |
| BMI, kg/m ² | 24.3 (4.8) | 24.7 (5) | 25.6 (5.3) | 27.3 (6.1) |
| Family history of colorectal cancer, % | 18.7 | 19.1 | 19.9 | 18.4 |
| Number of previous endoscopies | 1.6 (0.9) | 1.6 (1) | 1.6 (0.9) | 1.6 (1) |
| Time since most recent endoscopy, years | 3.7 (2.8) | 3.6 (2.7) | 3.5 (2.6) | 3.4 (2.6) |
| Reasons for endoscopy | | | | |
| Screening, % | 48.7 | 48.4 | 48.1 | 44.2 |
| Symptoms, % | 49.0 | 48.8 | 50.1 | 53.3 |
| Missing, % | 2.3 | 2.8 | 1.9 | 2.5 |
| Postmenopausal hormone use | | | | |
| Pre-menopause, % | 81.4 | 82.6 | 79.6 | 80.1 |
| No prior use, % | 4.3 | 3.6 | 3.5 | 3.3 |
| Current use, % | 11.9 | 11.7 | 14.5 | 13.5 |
| Past use, % | 2.5 | 2.2 | 2.4 | 3.0 |
| History of diabetes, % | 1.6 | 1.8 | 2.8 | 4.1 |
| Ever smokers, % | 32.2 | 31.9 | 32.6 | 34.5 |
| Pack-years among prior smokers | 11.1 (8.7) | 12.2 (9.7) | 13.1 (10) | 14 (10.7) |
| Alcohol intake, g/day | 4.4 (6.8) | 3.3 (4.7) | 2.9 (4.7) | 3.3 (5.7) |
| Physical activity, MET-hours/week | 24.7 (23.4) | 20.5 (20.4) | 19.6 (18.8) | 18.9 (18.4) |
| Regular aspirin use, % | 11.4 | 11.5 | 12.0 | 12.4 |
| Regular non-aspirin NSAID use, % | 32.0 | 30.2 | 34.0 | 36.9 |
| Multivitamin use, % | 63.2 | 56.7 | 55.0 | 52.1 |
| Total calorie intake, kcal/day | 2145 (445) | 1826 (431) | 1690 (422) | 1671 (454) |
| Dietary intake, servings/week | | | | |
| Processed meat | 1.3 (1.3) | 1.5 (1.3) | 1.6 (1.4) | 2 (2.1) |
| Liquor | 0.2 (0.6) | 0.2 (0.6) | 0.2 (0.7) | 0.5 (1.6) |
| Low-calorie drinks | 2.2 (3.4) | 2.9 (3.8) | 5.3 (5.2) | 16.6 (11.2) |
| Beer | 1.2 (2.8) | 0.7 (1.3) | 0.5 (1.2) | 0.5 (1.2) |
| Fruit juice | 7.8 (6.1) | 4.9 (4) | 3.4 (2.9) | 2.9 (3) |
| Legumes | 4.4 (3) | 2.7 (1.5) | 2.2 (1.4) | 1.9 (1.4) |
| Other vegetables | 10.4 (6.4) | 6.5 (3.4) | 5.5 (3.3) | 5.1 (3.5) |
| Sweets & desserts | 10 (9) | 7.3 (5.2) | 5.9 (4.1) | 6.1 (4.7) |

| | | | | |
|---|------------|------------|------------|------------|
| Alternate Healthy Eating Index 2010 (without alcohol) | 47.2 (9.6) | 44.3 (9.1) | 43.7 (8.7) | 43.3 (8.3) |
|---|------------|------------|------------|------------|

Values are means (SD) for continuous variables. Percentages or Ns or both for categorical variables. Variables standardized to the age distribution of the study population.

Values of polytomous variables may not sum to 100% due to rounding

*Not age adjusted


Abbreviations: g (gram), kcal (kilocalories), kg (kilogram), m (meters), MET (metabolic equivalent of task), NSAID (non-steroidal anti-inflammatory drug)

Supplementary Table 4. Sulfur microbial diet and risk of early-onset (age <50 years) serrated lesion by polyp location and size, NHSII, 1991-2015

| | Sulfur microbial diet score (quartile) | | | | <i>P</i> _{trend} § |
|----------------------------|--|---------------------|---------------------|---------------------|-----------------------------|
| | 1 (lowest) | 2 | 3 | 4 (highest) | |
| Overall | 420 | 420 | 414 | 415 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 0.99 (0.86 to 1.14) | 0.98 (0.84 to 1.13) | 1.01 (0.87 to 1.17) | 0.86 |
| Multivariable OR (95% CI)‡ | 1 [Ref] | 0.96 (0.84 to 1.11) | 0.93 (0.80 to 1.08) | 0.90 (0.77 to 1.05) | 0.08 |
| Size | | | | | |
| <u>Large (≥1 cm)</u> | 27 | 40 | 40 | 31 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.42 (0.87 to 2.32) | 1.42 (0.86 to 2.37) | 1.13 (0.65 to 1.98) | 0.73 |
| Multivariable OR (95% CI)‡ | 1 [Ref] | 1.44 (0.88 to 2.35) | 1.40 (0.83 to 2.36) | 1.02 (0.58 to 1.76) | 0.94 |
| <u>Small (<1 cm)</u> | 369 | 352 | 358 | 371 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 0.94 (0.81 to 1.10) | 0.97 (0.83 to 1.13) | 1.04 (0.89 to 1.21) | 0.89 |
| Multivariable OR (95% CI)‡ | 1 [Ref] | 0.92 (0.79 to 1.08) | 0.93 (0.79 to 1.09) | 0.93 (0.79 to 1.09) | 0.22 |
| Location | | | | | |
| <u>Proximal</u> | 149 | 145 | 147 | 143 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 0.95 (0.75 to 1.21) | 0.96 (0.75 to 1.23) | 0.97 (0.75 to 1.24) | 0.66 |
| Multivariable OR (95% CI)‡ | 1 [Ref] | 0.95 (0.75 to 1.21) | 0.96 (0.75 to 1.23) | 0.93 (0.72 to 1.20) | 0.48 |
| <u>Distal</u> | 179 | 182 | 179 | 193 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 0.98 (0.80 to 1.22) | 0.97 (0.77 to 1.21) | 1.08 (0.87 to 1.34) | 0.59 |
| Multivariable OR (95% CI)‡ | 1 [Ref] | 0.95 (0.77 to 1.17) | 0.90 (0.71 to 1.12) | 0.91 (0.73 to 1.14) | 0.34 |
| <u>Rectal</u> | 143 | 150 | 152 | 138 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.08 (0.85 to 1.37) | 1.14 (0.89 to 1.45) | 1.07 (0.83 to 1.37) | 0.66 |
| Multivariable OR (95% CI)‡ | 1 [Ref] | 1.05 (0.82 to 1.33) | 1.06 (0.83 to 1.36) | 0.90 (0.69 to 1.17) | 0.35 |

†Adjusted for age (continuous), total caloric intake (in quintiles), time period of endoscopy (in 2-year intervals), number of reported endoscopies (continuous), time in years since the most recent endoscopy (continuous), and reason for the current endoscopy (screening, symptoms, missing).

‡Additionally adjusted for height (continuous), body mass index (in quintiles), family history of colorectal cancer (yes, no), menopausal status (premenopausal, postmenopausal), menopausal hormone use (never, past, current use of menopausal hormones), personal history of type 2 diabetes (yes, no), pack-years of smoking (never, 1-4.9, 5-19.9, 20-39.9, ≥40 pack-years), physical activity (in metabolic equivalent of task-hours/week, quintiles), current use of multivitamin (yes, no), regular use of aspirin (yes, no), and regular use of non-steroidal anti-inflammatory



drugs (yes, no).

§Calculated using the median of each quartile as a continuous variable.

Abbreviations: CI (confidence interval), OR (odds ratio)

Supplementary Table 5. Sulfur microbial diet in high school and risk of early-onset (age <50 years) conventional adenoma by malignancy risk, NHSII, 1998-2015

| | Sulfur microbial diet score (median) | | P_{trend} |
|------------------------------------|--------------------------------------|---------------------|--------------------|
| | Below | Above | |
| Histology | | | |
| <u>Tubulovillous/villous</u> | 48 | 56 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.21 (0.77 to 1.91) | 0.27 |
| Multivariable model 1 OR (95% CI)‡ | 1 [Ref] | 1.26 (0.80 to 2.00) | 0.17 |
| Multivariable model 2 OR (95% CI)§ | 1 [Ref] | 1.27 (0.79 to 2.04) | 0.17 |
| <u>Tubular</u> | 254 | 286 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.13 (0.94 to 1.37) | 0.04 |
| Multivariable model 1 OR (95% CI)‡ | 1 [Ref] | 1.10 (0.91 to 1.33) | 0.11 |
| Multivariable model 2 OR (95% CI)§ | 1 [Ref] | 1.11 (0.92 to 1.34) | 0.09 |
| Size | | | |
| <u>Large (≥1 cm)</u> | 58 | 70 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.25 (0.84 to 1.85) | 0.16 |
| Multivariable model 1 OR (95% CI)‡ | 1 [Ref] | 1.23 (0.82 to 1.85) | 0.19 |
| Multivariable model 2 OR (95% CI)§ | 1 [Ref] | 1.28 (0.85 to 1.92) | 0.14 |
| <u>Small (<1 cm)</u> | 208 | 219 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.05 (0.85 to 1.29) | 0.31 |
| Multivariable model 1 OR (95% CI)‡ | 1 [Ref] | 1.02 (0.83 to 1.26) | 0.54 |
| Multivariable model 2 OR (95% CI)§ | 1 [Ref] | 1.03 (0.83 to 1.27) | 0.51 |
| Location | | | |
| <u>Proximal</u> | 92 | 112 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.27 (0.94 to 1.72) | 0.38 |
| Multivariable model 1 OR (95% CI)‡ | 1 [Ref] | 1.24 (0.92 to 1.68) | 0.52 |
| Multivariable model 2 OR (95% CI)§ | 1 [Ref] | 1.26 (0.93 to 1.71) | 0.47 |
| <u>Distal</u> | 124 | 121 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.01 (0.77 to 1.33) | 0.06 |
| Multivariable model 1 OR (95% CI)‡ | 1 [Ref] | 0.97 (0.74 to 1.28) | 0.11 |
| Multivariable model 2 OR (95% CI)§ | 1 [Ref] | 0.97 (0.73 to 1.28) | 0.10 |
| <u>Rectal</u> | 40 | 45 | |
| Age-adjusted OR (95% CI)† | 1 [Ref] | 1.12 (0.68 to 1.87) | 0.18 |
| Multivariable model 1 OR (95% CI)‡ | 1 [Ref] | 1.14 (0.68 to 1.91) | 0.14 |

Multivariable model 2 OR (95% CI)§

1 [Ref]

1.16 (0.69 to 1.95)

0.13

†Adjusted for age (continuous), total caloric intake at age 13-18 (continuous), time period of endoscopy (in 2-year intervals), number of reported endoscopies (continuous), time in years since the most recent endoscopy (continuous), and reason for the current endoscopy (screening, symptoms, missing).

‡Additionally adjusted for height (continuous), body mass index at age 18 (continuous), pack-years of smoking before age 20 (continuous), physical activity at grade 9-12 (continuous), and multivitamin use at age 13-18 (yes, no).

§Additionally adjusted for putative colorectal cancer risk factors in mid-adulthood: body mass index (in quintiles), family history of colorectal cancer (yes, no), menopausal status (premenopausal, postmenopausal), menopausal hormone use (never, past, current use of menopausal hormones), personal history of type 2 diabetes (yes, no), pack-years of smoking (never, 0.1-19.9, ≥20 pack-years), physical activity (in metabolic equivalent of task-hours/week, quintiles), current use of multivitamin (yes, no), regular use of aspirin (yes, no), and regular use of non-steroidal anti-inflammatory drugs (yes, no).

|| Calculated using the median of each quantile as a continuous variable

Abbreviations: CI (confidence interval), cm (centimeter), OR (odds ratio)