SHORT COMMUNICATION

BJC

British Journal of Cancer (2013) 108, 1163–1167 | doi: 10.1038/bjc.2013.60

Keywords: Hodgkin lymphoma; twin study; survivorship; hygiene hypothesis; human fecal microbiome

Fecal microbiota diversity in survivors of adolescent/young adult Hodgkin lymphoma: a study of twins

W Cozen[1,2,3,7], G Yu[4,7], M H Gail[4], V K Ridaura[5], B N Nathwani[6], A E Hwang[1], A S Hamilton[1,2], T M Mack[1,2,3], J I Gordon[5] and J J Goedert*[,4]

[1]Department of Preventive Medicine, USC Keck School of Medicine, Los Angeles, CA 90089-9175, USA; [2]Norris Comprehensive Cancer Center, University of Southern California Keck School of Medicine, Los Angeles, CA 90089-9175, USA; [3]Department of Pathology, University of Southern California Keck School of Medicine, Los Angeles, CA 90089-9175, USA; [4]Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, Room 7068, Rockville, MD 20852, USA; [5]Center for Genome Sciences and Systems Biology, Washington University in St Louis, St Louis, MO 63108, USA and [6]Department of Pathology, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

Background: Adolescent/young adult Hodgkin lymphoma (AYAHL) survivors report fewer exposures to infections during childhood compared with controls, and they have functional lymphocyte aberrations. The gut microbiota plays a central role in immunity.

Methods: We investigated whether fecal microbial diversity differed between 13 AYAHL survivors and their unaffected co-twin controls. Pyrosequencing of fecal bacterial 16S rRNA amplicons yielded 252 943 edited reads that were assigned to species-level operational taxonomic units (OTUs) and standardised for sequencing depth by random sampling. Microbial diversity was compared within *vs* between twin pairs and by case–control status.

Results: The number of unique OTUs was more similar within twin pairs compared with randomly paired participants ($P=0.0004$). The AYAHL cases had fewer unique OTUs compared with their co-twin controls (338 *vs* 369, $P=0.015$); this difference was not significant (169 *vs* 183, $P=0.10$) when restricted to abundant OTUs.

Conclusion: In this small study, AYAHL survivors appear to have a deficit of rare gut microbes. Further work is needed to determine if reduced microbial diversity is a consequence of the disease, its treatment, or a particularly hygienic environment.

Successful treatment of adolescent/young adult Hodgkin lymphoma (AYAHL) has enabled investigation of the pathogenesis and consequences of this disease. Adolescent/young adult Hodgkin lymphoma is associated with reduced exposures to infections, including small sibship size, less crowded living conditions, and higher socioeconomic status (Westergaard *et al*, 1997). Similarly, a previous study of AYAHL survivors suggested fewer early childhood fecal-oral exposures compared with unaffected co-twin controls (Cozen *et al*, 2009). Adolescent/young adult Hodgkin lymphoma is associated with suppressed T-helper 1 (Th1) immunity, a hyper-inflammatory Th2 response, and persistent immune defects (Fisher *et al*, 1980; Poppema, 1996; Salas *et al*, 2012). Colonisation, expansion, and maturation of gut microbial populations during infancy (Dominguez-Bello *et al*, 2010; Koenig *et al*, 2011) coincide with a switch from a fetal Th2-dominated to a mature Th1-dominated immune profile (Matricardi and Bonini, 2000; Martin *et al*, 2010). Elevated levels of IgE and Th2 cytokines in AYAHL patients, as well as deficits of cytotoxic T cells and

*Correspondence: Dr JJ Goedert; E-mail: goedertj@mail.nih.gov
[7]Co-first authors

Received 23 October 2012; revised 9 January 2013; accepted 20 January 2013; published online 26 February 2013

© 2013 Cancer Research UK. All rights reserved 0007 – 0920/13

natural killer cells in Hodgkin lymphoma tumours (Romagnani *et al*, 1980; Samoszuk, 1992; Skinnider and Mak, 2002; Poppema, 2005), suggest a failure to make this early Th2-to-Th1 switch. These observations raise the possibility that the gut microbiota, which strongly influences adaptive and innate immunity (Blumberg and Powrie, 2012; Hooper *et al*, 2012), may impact AYAHL. Thus, the current study employed next generation sequencing of fecal microbial DNA to test whether gut microbial diversity or taxonomy differed between AYAHL survivors and unaffected co-twin controls.

MATERIALS AND METHODS

Detailed Materials and methods are in Supplementary information. Briefly, subjects were AYAHL-discordant twin pairs participating in the International Twin Study (Mack *et al*, 1995). Living pairs, resident in the United States or Canada, were eligible if one member had been diagnosed with histologically verified AYAHL under 50 years of age (actual age 18–44 years), if both members of the pair were alive according to a death index linkage, and if the pair was discordant for reported early childhood fecal-oral exposures as assessed by a questionnaire completed and returned by both members of the pair ($n=33$ pairs) (Cozen *et al*, 2009). At least 1 member of 15 pairs either declined or was not able to be located, and 5 pairs were excluded for gastrointestinal illness or recent use of certain medications in at least 1 member, leaving 13 pairs. All 26 participants provided written informed consent and 1 fecal specimen. The majority of the HL survivors had nodular sclerosis subtype that was negative for Epstein Barr virus (EBV). Participants were provided with a kit that fits over the toilet seat for fecal specimen collection. Specimens were immediately frozen at -20°C, shipped overnight to University of Southern California, and frozen at -80°C after replacement of identifying information with a code number. The study was approved by the responsible Institutional Review Boards.

Following microbial cell lysis, DNA was extracted from three aliquots of each fecal specimen, pooled, and used to amplify V2 region 16S rRNA genes using primers FWD: AGAGTTTG ATCCTGGCTCAG and REV: TGCTGCCTCCCGTAGGAGT (Turnbaugh *et al*, 2009). The amplicons were purified, pooled, and sequenced (454 Life Sciences FLX, standard chemistry; Roche Company, Branford, CT, USA). Low-quality and chimeric sequences were removed (Wang and Wang, 1996; Caporaso *et al*, 2010; Edgar *et al*, 2011). Remaining sequences with $\geqslant$97% identity were defined as species-level operational taxonomical units (OTUs) (Edgar, 2010) and assigned to taxa with the Ribosomal Data Project classifier version 2.2 (Wang *et al*, 2007). Each taxon's proportion was its relative abundance.

Comparisons were performed on 5520 sequences (initial analysis) or 5258 sequences ('conservative' analysis that discarded rare OTUs) randomly sampled from each individual. Difference in OTU alpha diversity (within a sample) within twin pairs, compared with randomly paired individuals, was tested by one-way analysis of variance (Howell, 2002). Cases and controls were compared on alpha diversity and body mass index by paired *t*-test (Freeman *et al*, 2007). Differences in microbial communities within *vs* between twin pairs (beta diversity) were tested with UniFrac (weighted and unweighted for relative abundance; Lozupone and Knight, 2005) and permutation tests with 1000 replications. Each time, the sum of the paired UniFrac distances across the 13 twin pairs was compared with the corresponding sum in 13 randomly paired sets of individuals. For case–control comparisons of beta diversity, the sum of UniFrac values across all pairings of the 13 cases ($13\times12/2=78$ summands) was compared with the corresponding sum in 1000 random selections of 13 individuals. Monte Carlo methods (Caporaso *et al*, 2010) provided nearly identical results (not presented). Case–control comparisons of taxa with mean relative abundance $\geqslant$0.1% (Freeman *et al*, 2007) were performed with paired *t*-tests. All statistical tests were two-sided. $P<0.05$ was deemed significant.

RESULTS

Table 1 presents demographic characteristics of the 13 twin pairs. Cases were diagnosed an average of 22.5 years (range 10–36 years) before participation in this study. After editing to remove low quality sequences and amplification artefacts (chimeras), there were 252 943 bacterial V2-16S rRNA reads from the 26 fecal samples. These were assigned to 2513 species-level OTUs. Table 2 presents within-pair differences in alpha diversity (Shannon index and number of unique OTUs). Mean differences were significantly smaller within pairs than between pairs (Shannon index, $P=0.02$; OTUs, $P=0.0004$).

In the initial analysis, all measures of alpha diversity were higher in controls than in AYAHL cases (Table 3). Controls had significantly more unique OTUs (mean difference, 31), a measure that does not weight for relative abundance. Accounting for abundance, controls still had higher diversity, but the differences were attenuated and not statistically significant (Chao1 difference = 41, $P=0.066$; PD_whole tree difference = 1.6, $P=0.051$; Shannon index = 0.2, $P=0.27$). Alpha diversity did not differ significantly by zygosity, sex, current age, age at separation, more oral exposures in early childhood, or time since AYAHL diagnosis ($P\geqslant0.19$, data not presented). Case-minus-control differences in body mass index were not significant (median 0.5 kg m^{-2}, mean -0.08 kg m^{-2}, $P=0.96$).

Compared with the initial analysis, the conservative analysis used 5% fewer reads per sample (5258, Supplementary Table 1) and the data set had 81% fewer OTUs (488 *vs* 2513). As in the initial analysis, all measures of alpha diversity in the conservative

Table 1. Characteristics of the 13 twin pairs, discordant for Hodgkin lymphoma, whose fecal samples were used for the study

Characteristic	
Race	
White, no. of pairs (%)	12 (92)
Non-white, no. of pairs (%)	1 (8)
Zygosity	
MZ, no. of pairs (%)	5 (38)
DZ, no. of pairs (%)	8 (62)
Gender type	
MM, no. of pairs (%)	3 (23)
FF, no. of pairs (%)	6 (46)
MF, no. of pairs (%)	4 (31)
Other characteristics	
Mean age (range) at diagnosis	29 (18–44) years
Mean interval (range) between diagnosis and participation	22.5 (10–36) years
Mean no. (range) of bacterial V2-16S rRNA sequences (initial analysis)	9798 (5520–23 755)
Mean no. (range) of bacterial V2-16S rRNA sequences (conservative analysis)	9114 (5258–22 199)

Abbreviations: DZ = dizygotic; F = female; M = male; MZ = monozygotic.

Table 2. Control–case differences in Shannon index and number of unique species-level operational taxonomic units (OTUs) in the fecal samples of 13 twin pairs, discordant for Hodgkin lymphoma

Twin pairs	Zygosity	Difference in Shannon index[a]	Difference in number of unique OTUs[a]
1	DZ	0.92	71.7
2	DZ	1.68	66.3
3	DZ	−0.17	−5.2
4	DZ	−0.31	−4.1
5	DZ	0.49	37.1
6	DZ	−0.72	−51.5
7	DZ	0.93	48.5
8	DZ	−0.12	34.9
9	MZ	−0.36	34.6
10	MZ	−0.09	−10.6
11	MZ	0.08	80.7
12	MZ	0.15	67.4
13	MZ	0.25	26.8
Mean within-pair difference[b]		0.48	41.5
Mean between-pair difference[c]		0.75	77.2
P-value[d]		0.02	0.0004

Abbreviations: DZ = dizygotic; MZ = monozygotic.
[a]Control value − case value.
[b]Mean of absolute differences within twin pairs.
[c]Estimated mean of the absolute difference between values from two members chosen at random from different twin sets.
[d]Analysis of variance to test whether values are more similar within *vs* between twin pairs.

Table 3. Comparisons of alpha diversity measurements between Hodgkin lymphoma cases and co-twin controls, in 13 pairs discordant for Hodgkin lymphoma

Measurements of alpha diversity[a]	Mean (cases)	Mean (unaffected co-twins)	Mean difference (unaffected co-twin − case difference)	*P*-value[b]
Initial analysis				
No. of unique OTUs	338	369	31	0.015
Shannon index	5.6	5.8	0.2	0.27
Chao1	533	574	41	0.066
PD_whole tree	21.2	22.8	1.6	0.051
Conservative analysis				
No. of unique OTUs	183	196	13	0.10
Shannon index	5.2	5.4	0.2	0.40
Chao1	230	237	7	0.47
PD_whole tree	13.7	14.6	0.9	0.045

[a]OTUs, species-level operational taxonomic units; number of unique OTUs is also referred to as 'richness'; Shannon index, which is a conservative alpha diversity estimate that adjusts for relative abundance of each OTU, is defined as (negative) the sum over OTUs of the product of the relative abundance of the OTU times the natural logarithm of the relative abundance; Chao1 is a presence/absence of alpha diversity indicator that is bias-corrected for rare taxa; PD, phylogenetic distance; PD_whole tree is an alpha diversity estimate that reflects phylogenetic divergence among OTUs present within an individual. The initial analysis included all 16S rRNA sequence reads. The conservative analysis restricted to reads with a minimum relative abundance of 0.1%.
[b]*P*-value by paired *t*-tests.

analysis were higher in controls than in cases (Table 3). This difference was statistically significant for PD_whole tree (mean 14.6 *vs* 13.7, $P = 0.045$), marginal for unique OTUs (mean 196 *vs* 183, $P = 0.10$), and non-significant for Chao1 (mean 237 *vs* 230, $P = 0.47$) and Shannon index (mean 5.4 *vs* 5.2, $P = 0.40$).

Cases and controls did not differ significantly in relative abundance of bacterial phylum-, class-, order-, or family-level taxa (data not shown). Restricting case–control comparisons with the 37 genera (of 108 identified) with a mean relative abundance of $\geqslant 0.1\%$ (Supplementary Table 2), controls had a higher relative abundance of *Actinobacteria collinsella* (control mean = 0.004, case mean = 0.002, $P = 0.03$), which was not significant after correction for multiple tests (Miller, 1991).

Consistent with the negligible differences in relative abundance, beta diversity did not differ between the 13 cases and random groupings of 13 participants (unweighted UniFrac $P = 0.07$; weighted UniFrac $P = 0.31$). In contrast, beta diversity was significantly smaller within twin pairs compared with randomly paired participants when relative abundance was not considered (unweighted UniFrac $P = 0.01$), but not when relative abundance was included (weighted UniFrac $P = 0.20$).

DISCUSSION

Our initial analysis indicated that by some measurement parameters long-term survivors of AYAHL had statistically significantly fewer unique species-level OTUs in their stool than did their unaffected co-twins. These results need to be cautiously interpreted as the absolute differences were modest and other measurements of diversity, which adjust for OTU relative abundance, were not statistically significant. In our conservative analysis, restricted to the 19% most abundant OTUs, only the PD_whole tree estimate of alpha diversity was statistically significant. Deeper sequencing could shed light on rare taxa and whether they discriminate between affected and unaffected co-twins. Based on this study, we cannot determine whether lower diversity contributed to the immune perturbations that underlie AYAHL risk, or was merely a consequence of the disease and its treatment. Two small studies suggest that various chemotherapy regimens immediately reduce and alter the composition of the gut microbiota (van Vliet *et al*, 2009; Zwielehner *et al*, 2011). However, these observations were partially confounded by antibiotic use, and long-term studies are lacking. In our study, microbial diversity was assayed many years after the case twins' diagnoses; observed differences between discordant co-twins could have been permanently induced by the disease or its treatment, or have been present before the malignancy appeared.

Twin pairs are a conservative and effective study design, not only for human genetics, but also when close matching of maternal and early life exposures is needed. Consistent with previous twin studies (Turnbaugh *et al*, 2009; Yatsunenko *et al*, 2012), fecal microbial diversity was significantly more concordant within our twin pairs than between randomly paired individuals. Likewise, concordance in diversity was similar for mono- and dizygotic pairs, suggesting that human genetic polymorphisms have little influence on the overall bacterial phylogenetic structure of the fecal microbiota (Turnbaugh *et al*, 2009; Yatsunenko *et al*, 2012). Moreover, our findings suggest that changes during adulthood related to diet, transient illnesses or treatments, are too small to modify the relatively fixed effects of twin pairing and the propensity for or consequences of obesity, inflammatory bowel disease, and, in our data, AYAHL (Turnbaugh *et al*, 2009; Willing *et al*, 2010).

Lower oral exposure to microbes in early life was associated with a 10-fold higher risk of AYAHL (Cozen *et al*, 2009) and with a pro-inflammatory Th2 profile in rodents. Specifically, invariant natural killer T cells, that can secrete Th2 cytokines, accumulate in germ-free mice affecting the pattern of colonisation with a gut microbiota from conventionally raised animals (Olszak *et al*, 2012). Neonatal rodents fail to develop a Th1 immune response on

a sterile diet, but a Th1 response can be induced by feeding them bacterial extracts (Bowman and Holt, 2001; Hrncir *et al*, 2008). In humans, genetic polymorphism and functional studies suggest that a Th2-dominated profile is a risk factor for AYAHL (Cozen *et al*, 2008; Urayama *et al*, 2012).

The current study has several limitations. First, it is very small, with only 13 pairs of AYAHL cases and co-twin controls. Thus, a statistically significant result could be due to chance. We excluded pairs in which either member had a gastrointestinal illness or had recently used a medication likely to affect the gut microbiota, thereby reducing confounding. These exclusions applied equally to the cases (HL survivors) and controls (unaffected co-twins) who were ascertained together as pairs, thus selection bias is unlikely. We noted that body mass index, which is a potential confounder (Turnbaugh *et al*, 2009), did not differ significantly between cases and controls. Second, we lacked specifics on treatment of the AYAHL, which might have helped to distinguish among the alternative hypotheses. Third, multiple fecal specimens per subject and deeper sequencing of each specimen would have been desirable to reduce misclassification of rare taxa and the potentially important fluctuations in the fecal microbiota over time (Turnbaugh *et al*, 2009; Caporaso *et al*, 2011; Wu *et al*, 2011; Consortium, 2012). Use of unaffected co-twins as controls is a strength, as it assured very tight matching on early life exposures that are related to AYAHL risk and establishment of the gut microbiota. Tight matching reduces the chance of a false-positive association.

In summary, the current study found that, by some parameters, AYAHL survivors exhibited modest reductions in the diversity of bacteria in their fecal microbiota compared with unaffected co-twin controls. Additional data are needed to confirm this finding and increase the power to detect such differences for a variety of diversity measures. Different study designs will be needed to determine if early life microbial exposures are causally linked to a susceptible immunophenotype for AYAHL (Cozen *et al*, 2009), as well as the effects of treatment and disease on microbiota configuration (van Vliet *et al*, 2009; Zwielehner *et al*, 2011).

ACKNOWLEDGEMENTS

We would like to thank Maureen Cairns for her assistance with data collection, and the twins for participating in the research. This work was supported by the Intramural Research Program of the National Cancer Institute, as well as by grants from the National Cancer Institute (1R01CA58839 TMM, R03CA110836-01A2 WC, and HHSN261201000035C WC) and the Leukemia Lymphoma Society (6137-07 WC).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Blumberg R, Powrie F (2012) Microbiota, disease, and back to health: a metastable journey. *Sci Transl Med* **4**(137): 137rv7.

Bowman LM, Holt PG (2001) Selective enhancement of systemic Th1 immunity in immunologically immature rats with an orally administered bacterial extract. *Infect Immun* **69**(6): 3719–3727.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**(5): 335–336.

Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights D, Gajer P, Ravel J, Fierer N, Gordon JI, Knight R (2011) Moving pictures of the human microbiome. *Genome Biol* **12**(5): R50.

Consortium THMP (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486**(7402): 207–214.

Cozen W, Gill PS, Salam MT, Nieters A, Masood R, Cockburn MG, Gauderman WJ, Martinez-Maza O, Nathwani BN, Pike MC, Van Den Berg DJ, Hamilton AS, Deapen DM, Mack TM (2008) Interleukin-2, interleukin-12, and interferon-gamma levels and risk of young adult Hodgkin lymphoma. *Blood* **111**(7): 3377–3382.

Cozen W, Hamilton AS, Zhao P, Salam MT, Deapen DM, Nathwani BN, Weiss LM, Mack TM (2009) A protective role for early oral exposures in the etiology of young adult Hodgkin lymphoma. *Blood* **114**(19): 4014–4020.

Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, Knight R (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci USA* **107**(26): 11971–11975.

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**(19): 2460–2461.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**(16): 2194–2200.

Fisher RI, DeVita Jr VT, Bostick F, Vanhaelen C, Howser DM, Hubbard SM, Young RC (1980) Persistent immunologic abnormalities in long-term survivors of advanced Hodgkin's disease. *Ann Intern Med* **92**(5): 595–599.

Freeman DA, Pisani R, Purves R (2007) *Statistics*. 4th edn (W. W. Norton & Company: New York, NY, USA.

Hooper LV, Littman DR, Macpherson AJ (2012) Interactions between the microbiota and the immune system. *Science* **336**(6086): 1268–1273.

Howell D (2002) *Statistical Methods for Psychology*. Duxbury: Pacific Grove, CA, USA.

Hrncir T, Stepankova R, Kozakova H, Hudcovic T, Tlaskalova-Hogenova H (2008) Gut microbiota and lipopolysaccharide content of the diet influence development of regulatory T cells: studies in germ-free mice. *BMC Immunol* **9**: 65.

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE (2011) Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci USA* **108**(Suppl 1): 4578–4585.

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* **71**(12): 8228–8235.

Mack TM, Cozen W, Shibata DK, Weiss LM, Nathwani BN, Hernandez AM, Taylor CR, Hamilton AS, Deapen DM, Rappaport EB (1995) Concordance for Hodgkin's disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. *N Engl J Med* **332**(7): 413–418.

Martin R, Nauta AJ, Ben Amor K, Knippels LM, Knol J, Garssen J (2010) Early life: gut microbiota and immune development in infancy. *Benef Microbes* **1**(4): 367–382.

Matricardi PM, Bonini S (2000) Mimicking microbial 'education' of the immune system: a strategy to revert the epidemic trend of atopy and allergic asthma? *Respir Res* **1**(3): 129–132.

Miller RGJ (1991) *Simultaneous Statistical Inference*. Springer-Verlag: New York.

Olszak T, An D, Zeissig S, Vera MP, Richter J, Franke A, Glickman JN, Siebert R, Baron RM, Kasper DL, Blumberg RS (2012) Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**(6080): 489–493.

Poppema S (1996) Immunology of Hodgkin's disease. *Bailliere's Clin Haematol* **9**(3): 447–457.

Poppema S (2005) Immunobiology and pathophysiology of Hodgkin lymphomas. *Hematology Am Soc Hematol Educ Prog* 231–238.

Romagnani S, Biagiotti R, Amadori A, Maggi E, Biti G, Bellesi G, Ricci M (1980) Hyperproduction of IgE and T-cell dysfunction in Hodgkin's disease. *Int Arch Allergy Appl Immunol* **63**(1): 64–72.

Salas C, Niembro A, Lozano V, Gallardo E, Molina B, Sanchez S, Ramos S, Carnevale A, Perez-Vera P, Rivera Luna R, Frias S (2012) Persistent genomic instability in peripheral blood lymphocytes from Hodgkin lymphoma survivors. *Environ Mol Mutagen* **53**(4): 271–280.

Samoszuk M (1992) IgE in Reed-Sternberg cells of Hodgkin's disease with eosinophilia. *Blood* **79**(6): 1518–1522.

Skinnider BF, Mak TW (2002) The role of cytokines in classical Hodgkin lymphoma. *Blood* **99**(12): 4283–4297.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI (2009) A core gut microbiome in obese and lean twins. *Nature* **457**(7228): 480–484.

Urayama KY, Jarrett RF, Hjalgrim H, Diepstra A, Kamatani Y, Chabrier A, Gaborieau V, Boland A, Nieters A, Becker N, Foretova L, Benavente Y, Maynadie M, Staines A, Shield L, Lake A, Montgomery D, Taylor M, Smedby KE, Amini RM, Adami HO, Glimelius B, Feenstra B, Nolte IM, Visser L, van Imhoff GW, Lightfoot T, Cocco P, Kiemeney L, Vermeulen SH, Holcatova I, Vatten L, Macfarlane GJ, Thomson P, Conway DI, Benhamou S, Agudo A, Healy CM, Overvad K, Tjonneland A, Melin B, Canzian F, Khaw KT, Travis RC, Peeters PH, Gonzalez CA, Quiros JR, Sanchez MJ, Huerta JM, Ardanaz E, Dorronsoro M, Clavel-Chapelon F, Bueno-de-Mesquita HB, Riboli E, Roman E, Boffetta P, de Sanjose S, Zelenika D, Melbye M, van den Berg A, Lathrop M, Brennan P, McKay JD (2012) Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. *J Natl Cancer Inst* **104**(3): 240–253.

van Vliet MJ, Tissing WJ, Dun CA, Meessen NE, Kamps WA, de Bont ES, Harmsen HJ (2009) Chemotherapy treatment in pediatric patients with acute myeloid leukemia receiving antimicrobial prophylaxis leads to a relative increase of colonization with potentially pathogenic bacteria in the gut. *Clin Infect Dis* **49**(2): 262–270.

Wang GC, Wang Y (1996) The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. *Microbiology* **142**(Part 5): 1107–1114.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**(16): 5261–5267.

Westergaard T, Melbye M, Pedersen JB, Frisch M, Olsen JH, Andersen PK (1997) Birth order, sibship size and risk of Hodgkin's disease in children and young adults: a population-based study of 31 million person-years. *Int J Cancer* **72**(6): 977–981.

Willing BP, Dicksved J, Halfvarson J, Andersson AF, Lucio M, Zheng Z, Jarnerot G, Tysk C, Jansson JK, Engstrand L (2010) A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology* **139**(6): 1844–1854. e1.

Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, Sinha R, Gilroy E, Gupta K, Baldassano R, Nessel L, Li H, Bushman FD, Lewis JD (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**(6052): 105–108.

Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI (2012) Human gut microbiome viewed across age and geography. *Nature* **486**(7402): 222–227.

Zwielehner J, Lassl C, Hippe B, Pointner A, Switzeny OJ, Remely M, Kitzweger E, Ruckser R, Haslberger AG (2011) Changes in human fecal microbiota due to chemotherapy analyzed by TaqMan-PCR, 454 sequencing and PCR-DGGE fingerprinting. *PLoS ONE* **6**(12): e28654.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)