ARTICLE

Received 12 Jan 2015 | Accepted 21 Sep 2015 | Published 30 Oct 2015

DOI: 10.1038/ncomms9727 OPEN

Gut mucosal microbiome across stages of colorectal carcinogenesis

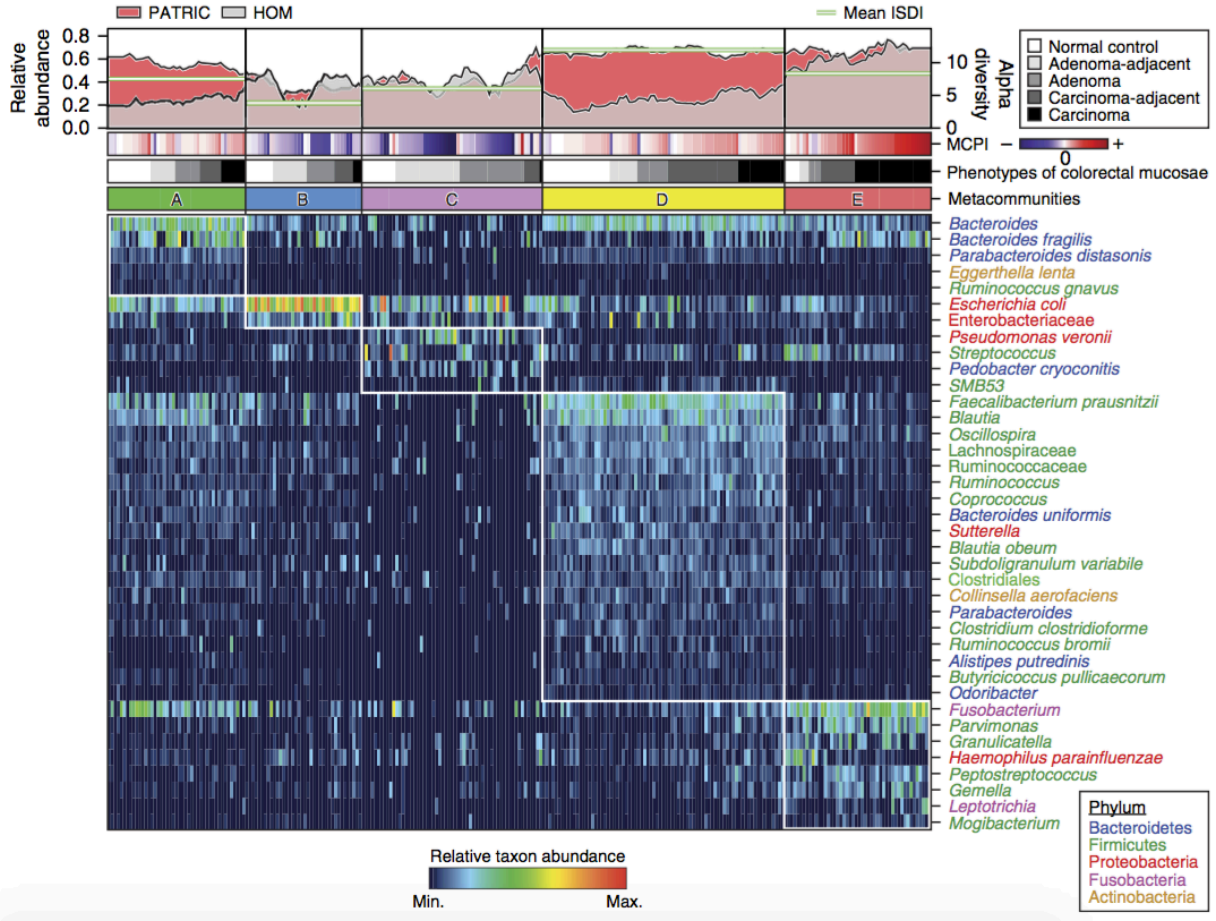
Geicho Nakatsu[1,2], Xiangchun Li[1,2,*], Haokui Zhou[3,*], Jianqiu Sheng[4,*], Sunny Hei Wong[1,2,*], William Ka Kai Wu[1,2,5,*], Siew Chien Ng[1,2], Ho Tsoi[1,2], Yujuan Dong[1,2], Ning Zhang[6], Yuqi He[4], Qian Kang[4], Lei Cao[1,2], Kunning Wang[1,2], Jingwan Zhang[1,2], Qiaoyi Liang[1,2], Jun Yu[1,2] & Joseph J.Y. Sung[1,2]

Gut microbial dysbiosis contributes to the development of colorectal cancer (CRC). Here we catalogue the microbial communities in human gut mucosae at different stages of colorectal tumorigenesis.

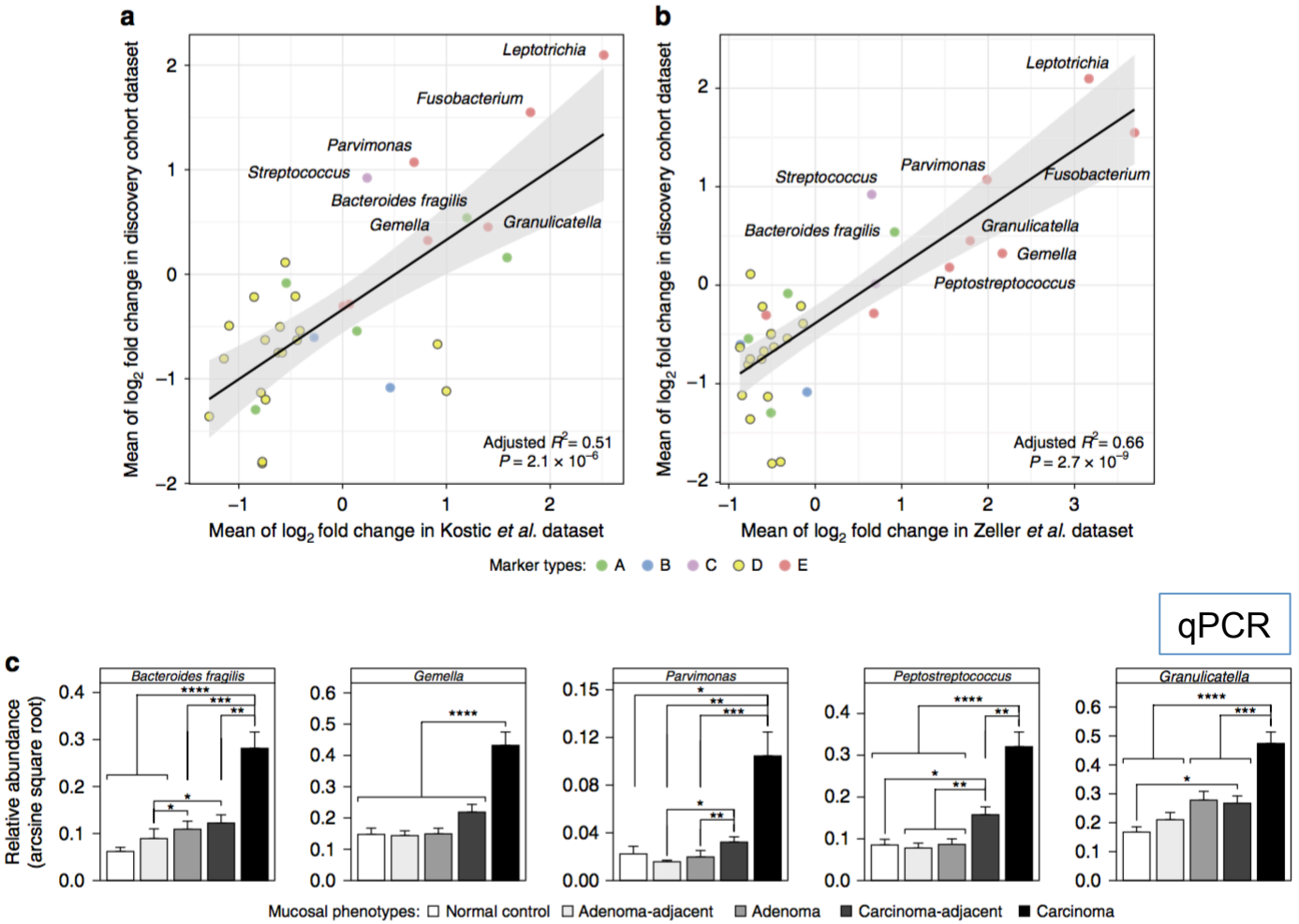
We analyse the gut mucosal microbiome of 47 paired samples of adenoma and adenoma-adjacent mucosae, 52 paired samples of carcinoma and carcinoma-adjacent mucosae and 61 healthy controls.

Probabilistic partitioning of relative abundance profiles reveals that a metacommunity predominated by members of the oral microbiome is primarily associated with CRC.

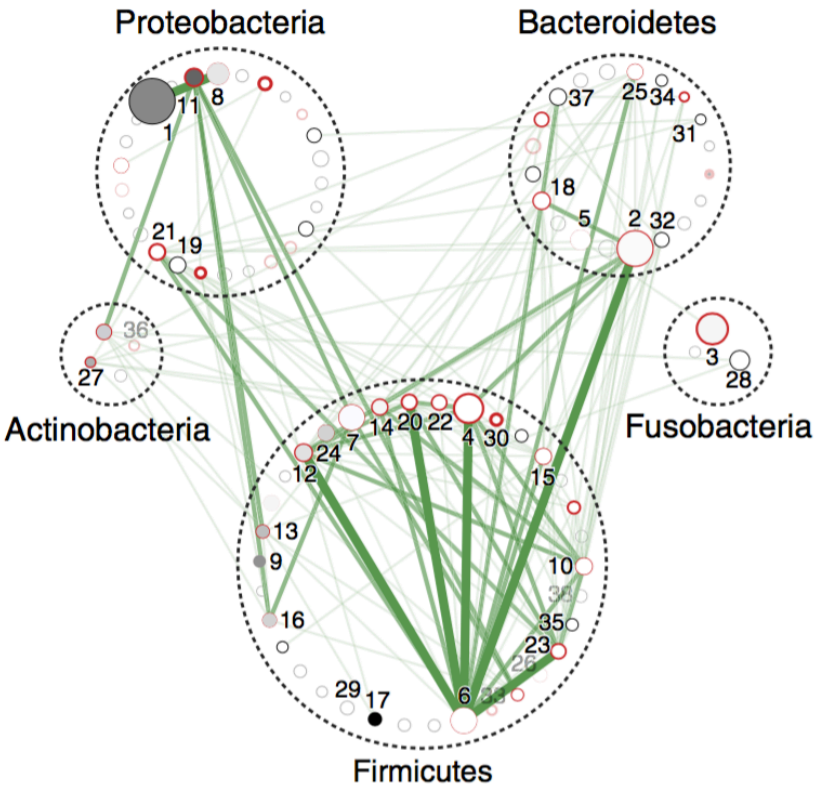
Analysis of paired samples shows differences in community configurations between lesions and the adjacent mucosae. Correlations of bacterial taxa indicate early signs of dysbiosis in adenoma, and co-exclusive relationships are subsequently more common in cancer. We validate these alterations in CRC-associated microbiome by comparison with two previously published data sets. Our results suggest that a taxonomically defined microbial consortium is implicated in the development of CRC.



Characterization of 16S rRNA gene catalogue for mucosal microbial communities in colorectal carcinogenesis. Fitting microbiome data to DMM models defined five metacommunities. Reads that are considered as being potentially originated from oral strains or known pathogenic strains in the human gut were classified against the 16S rRNA gene collections from the Human Oral Microbiome (HOM; version 13) database and PATRIC bacterial pathogen database as defined by pseudo-bootstrapped (n 1⁄4 1,000) confidence scores of 100 at species-level taxa or deeper, using the naive Bayesian classifier. The panels of metacommunity markers are ranked in the descending order of linear discriminant analysis scores from top to bottom. Columns represent microbiome profiles (arcsine square root-transformed) of 269 mucosal biopsies from individuals with or without adenoma or adenocarcinomas. (MCPIo0 for changes characteristic of adenomas; MCPI40 for changes characteristic of carcinomas).



Validations of metacommunity markers in independent cohorts. (a,b) Fold-change analyses in paired carcinoma and carcinoma-adjacent samples in two additional cohorts demonstrated significant agreement with our discovery cohort: (a) Kostic et al.7 data set (n ¼ 74) and (b) Zeller et al.20 data set (n ¼ 48). Shown are adjusted R2 and P values for goodness of fit from multiple linear regression models. (c) Real-time PCR amplifications of the most abundant sequences of representative bacterial phylotypes showed consistent enrichments in an additional Chinese cohort consisting of 207 mucosal biopsies (normal control, n ¼ 25; adenoma, n ¼ 41; adenocarcinoma, n ¼ 50). Error bars represent s.e.m. P values from Mann–Whitney U-tests are adjusted by Benjamini-Hochberg (BH) step-up procedure; *qo0.05; **qo0.01; ***qo0.001; ****qo0.0001.



1. *Escherichia coli*
2. *Bacteroides*
3. *Fusobacterium*
4. *Faecalibacterium prausnitzii*
5. *Bacteroides fragilis*
6. *Blautia*
7. *Streptococcus*
8. Enterobacteriaceae
9. *Parvimonas*
10. *Oscillospira*
11. *Haemophilus parainfluenzae*
12. Lachnospiraceae
13. *Granulicatella*
14. Ruminococcaceae
15. *Ruminococcus*
16. *Gemella*
17. *Peptostreptococcus*
18. *Bacteroides uniformis*
19. *Pseudomonas veronii*
20. *Coprococcus*
21. *Sutterella*
22. *Subdoligranulum variabile*
23. *Blautia obeum*
24. Clostridiales
25. *Parabacteroides distasonis*
26. *SMB53*
27. *Collinsella aerofaciens*
28. *Leptotrichia*
29. *Ruminococcus bromii*
30. *Clostridium clostridioforme*
31. *Pedobacter cryoconitis*
32. *Parabacteroides*
33. *Mogibacterium*
34. *Alistipes putredinis*
35. *Butyricicoccus pullicaecorum*
36. *Eggerthella lenta*
37. *Odoribacter*
38. *Ruminococcus gnavus*

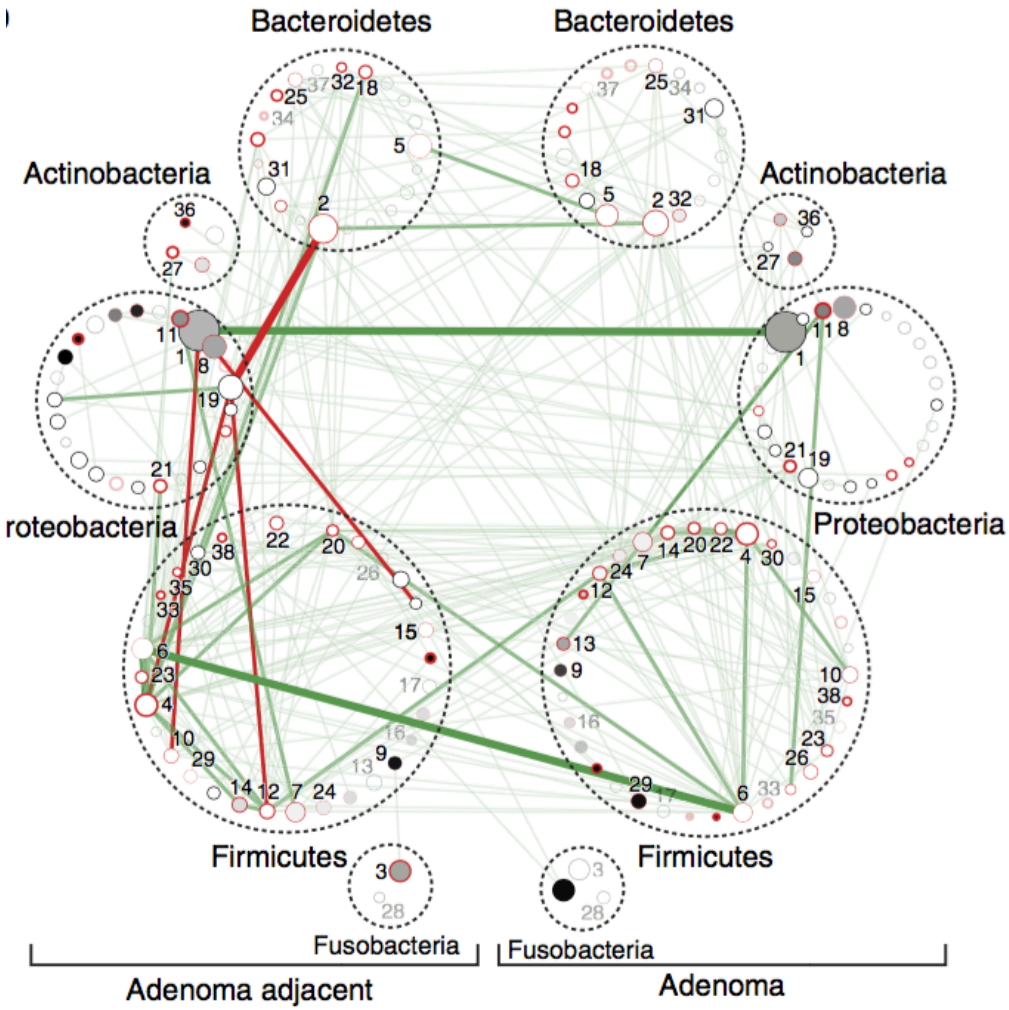
Min. Max. PATRIC

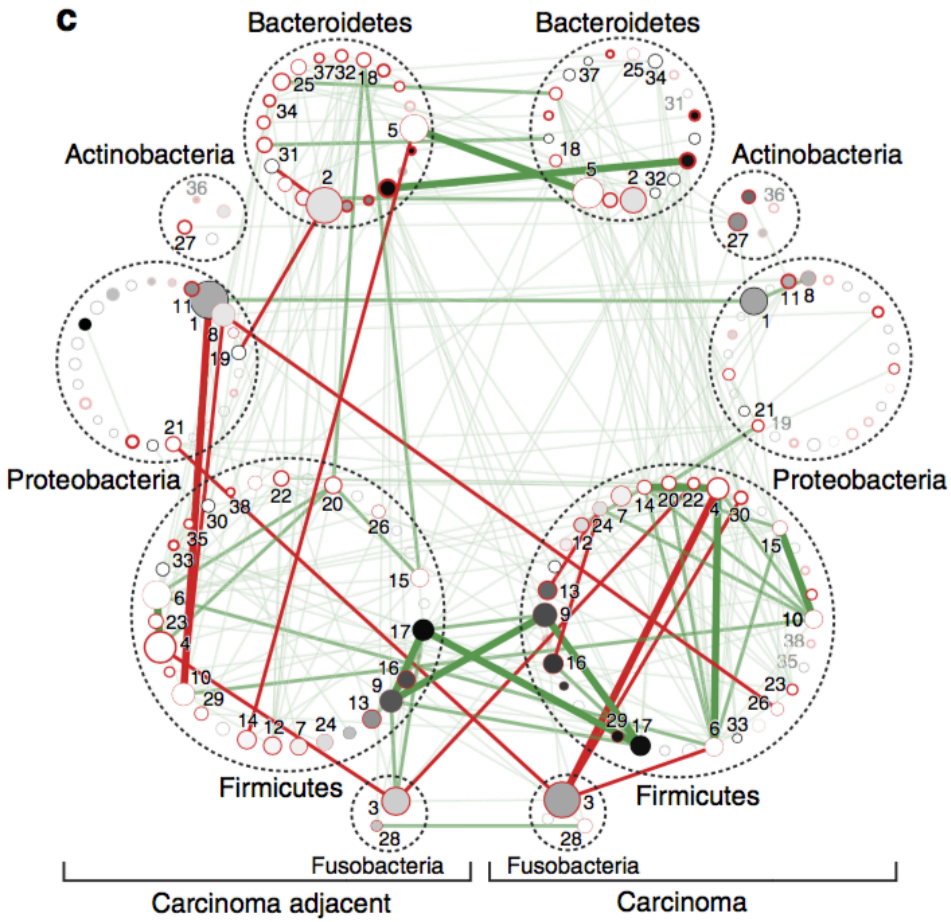
Min. Max. HOM

Neither being classified to PATRIC nor HOM

Co-occuring
- $r > 0.6$
- $r > 0.5$
- $r > 0.3$

Co-excluding
- $r < -0.4$
- $r < -0.3$


Bacteroidetes
Bacteroidetes
Actinobacteria
Actinobacteria
roteobacteria
Proteobacteria
Firmicutes
Firmicutes
Fusobacteria
Fusobacteria
Adenoma adjacent
Adenoma


c
Bacteroidetes
Bacteroidetes
Actinobacteria
Actinobacteria
Proteobacteria
Proteobacteria
Firmicutes
Firmicutes
Fusobacteria
Fusobacteria
Carcinoma adjacent
Carcinoma

In this regard, the rediscovery of CRC-specific enrichment of ***Fusobacterium*** and ***B. fragilis*** and the identification of novel CRC-associated candidates, such as ***Gemella***, ***Peptostreptococcus*** and ***Parvimonas***, expands the current scope of bacterial involvement in CRC development. In particular, *Gemella*, *Peptostreptococcus* and *Parvimonas* along with other microbes of oral origin formed a strong symbiotic network, which characterized the CRC-associated metacommunity E