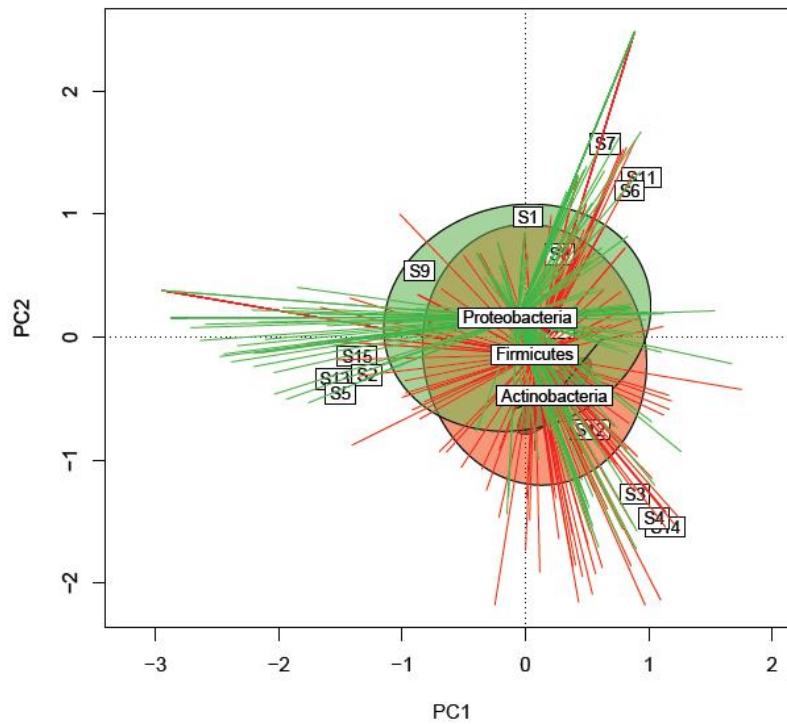


**UCC Library and UCC researchers have made this item openly available.  
Please [let us know](#) how this has helped you. Thanks!**

<b>Title</b>	A diverse range of human gut bacteria have the potential to metabolize the dietary component gallic acid
<b>Author(s)</b>	Esteban-Torres, María; Santamaría, Laura; Cabrera-Rubio, Raul; Plaza-Vinuesa, Laura; Crispie, Fiona; de las Rivas, Blanca; Cotter, Paul D.; Muñoz, Rosario
<b>Publication date</b>	2018-07-27
<b>Original citation</b>	Esteban-Torres, M., Santamaría, L., Cabrera-Rubio, R., Plaza-Vinuesa, L., Crispie, F., de las Rivas, B., Cotter, P. and Muñoz, R. (2018) 'A diverse range of human gut bacteria have the potential to metabolize the dietary component gallic acid', Applied and Environmental Microbiology, 84(19), e01558-18 (12pp). doi:10.1128/aem.01558-18
<b>Type of publication</b>	Article (peer-reviewed)
<b>Link to publisher's version</b>	<a href="https://aem.asm.org/content/aem/84/19/e01558-18.full.pdf">https://aem.asm.org/content/aem/84/19/e01558-18.full.pdf</a> <a href="http://dx.doi.org/10.1128/aem.01558-18">http://dx.doi.org/10.1128/aem.01558-18</a> Access to the full text of the published version may require a subscription.
<b>Rights</b>	© 2018, American Society for Microbiology. All rights reserved.
<b>Embargo information</b>	Access to this article is restricted until 6 months after publication by request of the publisher.
<b>Embargo lift date</b>	2019-01-27
<b>Item downloaded from</b>	<a href="http://hdl.handle.net/10468/7068">http://hdl.handle.net/10468/7068</a>

Downloaded on 2020-10-30T07:53:13Z

Figure S1



**FIG S1.** Principal component analysis results at the genus/species composition among samples from the human gut microbiota of athletes

Figure S2

LRH -----MTASPWDLRKVLDELKQDPQQYHETDVQVDPDAELAGVYRYIGAGGTV  
LPL -----MAEQPWLRRVLDEIKDDPKNYHETDVEVDPNAELSGVYRYIGAGGTV  
PET -----MAKKDVYDLRKVLDELKKEPGQYHETNVEVDPDAELSGVYRYIGAGGTV  
ERA -----MSTPYDLRKVLAELEQEMPGEYHETNVEVDPHAELSGVYRYIGAGGTV  
SGA -----MSEQPYDLRKVLEELKEIPGQYHETDVEIDPNAEISGVYRYIGAGGTV  
ECL -----MQNPINDLRSALALLQRHPGHYIETDHPVDPNAELAGVYRHIGAGGTV  
ACH MEQP---FPAFIDFPPINDLRSALARLQVPGQLLHTDHPVDPHGELAGVYKRVGAGGTV  
KIN ----MNRTDHTANSPVIDLRSALARLEQSPDELISTDHEIDPRSELAGVYKRVGAGGTV  
ATR MEDNVMDRTDHAANSPIVLDLRSALARLEAFPGELISTDQEVDPSELAGVYKRVGAGGTV  
AGL -----MNDTIHDLRSALDYLRDMPGQLLETDTDEVDCDAEVSQVYRHHVAGGTV  
ERY -----MKEPYDLRSALQQLEAEEGQLLITDKLTNTDDELAGVYRYIGGGGTL  
COP -----MSKKKVRDLRSALALLQSIDGQLVETDVEVDPPLGELSGVYRHHVAGGTV  
BLA -----MAEKVRDLRSALALLQMPDQLIETEVEVDPMAELAGVYRHHVAGGTV  
FIR -----MADKVRDLRSALERLKTMEGQYIETDVEVDPMAELAGVYRYVAGGTV  
AHA -----MQDKVRDLRSALKRKLQMEGQYIETDVEVDPMAELAGVYRYVAGGTV  
LAC -----MQDKVRDLRSALKRKLQMEGQYIETDVEVDPMAELAGVYRYVAGGTV  
DLO -----MAAKVTDLRSALKMLEDMPGQLIETDVEVEPMAELSGVYRHHVAGGTV  
IBA -----MNKVRDLRSALALLNMPGQLIETDVEVDPMAELSGVYRHHVAGGTV  
CBU -----MSNKVYDLRSALALLKTLPGQLIETDVEVDPMAELAGVYRYVAGGTV  
HHA -----MVEDLRSAVEELKKYENQIACDTEVDSYAEVAGIYRYVAGGTV  
TSA ---MKIKERVNKMANKIHDLRSALIEFLKQHENQIVYTNSEVDCEAEISGIYRYVAGGTV  
PAG -----MSGKESKVNDLRSALALLSQYDNELIYTDPEVDPVAELSGVYRYVAGGTV  
HAL ----MANNADNKIKNP IHDLRSALIEFLKNQPGELVSTNVEVDPCAELSGVYRYVAGGTV  
SBL -----MKNKDKIPVHDLRSALALLKTLPGEYVETDTEVDPHAELSGVYRYVAGGTV  
KRA ----MSNSENKNTSGVTDLRSALIEFLKTLPGEYVETDTEVDPHAELSGVYRYVAGGTV  
KMI ----MANSNDKTPSSVHDLRSALALLKTLPGEYVETDTEVDPHAELSGVYRYVAGGTV

\*\*\* .: : . \* : .\*:\*:\*: :\* . \*\*

LRH ERPTQ-EGPAMMFNNVVGFPTRVLIIGLMSRKRKRVGKMFHQDYHTLGRFLNKAVLNPIQP  
LPL QRPTQ-EGPAMMFNNVKGFPDTRVLTGLMASRRRVGKMFHHDYQTLGQYLNEAVSNPVAP  
PET QRPTQ-EGPAMMFNNVKGFPDTRVLMGLMASRKRKRVGKMFHHDYHTLGQFLNDSVENPVDP  
ERA KRPTK-EGPAMMFNCVKGFDPTRVLIGI IASRDRVGKILHHDPKHLGRLLKDSVQNPVKP  
SGA ERPTQ-EGPAMTFNNIKGFPNVRVNI GTMASRKRVGHI LHHDYKDLGHLNKAVERNPKP  
ECL KRPTR-TGPAMMFNSVKGYPGSRILVGMHASRERAALLGCVPSKLAQHVQAVKNPVAP  
ACH KRPTR-LGPAMMFNHKIGYPDSRVLVGMASRERVALLLDTPDRLAERMGEAIEKAIDP  
KIN MRPTR-TGPAMMFENVKGYPGARVLVGLMAKRERVALLLDSRPEELGRRMGEAVLNGIAP  
ATR MRPTR-TGPAMMFENVKGYPGARVLVGLMAKRERVALLLDSKPEELGQRMGEAVLNGIAP  
AGL ARPTK-EGPAMVFNNVKGFPDARVAIGLLSSRKRVAALLGMDEQYLGQLIGQKLAETIPP  
ERY QRPTQ-LGPAMLFTNIQNHGSRVLI GLLGDRQRCASLLNTTSEQLPFLMNKAYSKRLLDP  
COP MRPTQ-EGPAMIFNHVKGHPGARVAIGLLASRRRVGYLLDCEPQKLGFLKDSVNNPIKP  
BLA KRPTK-EGPAMIFNRKIGHPDARVAIGLLASRKRVAALLDTPENLGMMLCKSVENPIPP  
FIR KRPTK-EGPAMVFNNVKGHKDARVAIGVLASRKRVAALLDCKPEELGKKLYHSVDNPIPP  
AHA KRPTK-EGPAMVFNNVKGHKDARVAIGVLASRKHVAALLDCKPEELGKKLYHSVDNPIAP  
LAC KRPTK-EGPAMVFNNVKGHKDARVAIGVLASRKRVAALLDCKPEELGKKLYHSVDNPIAP  
DLO MRPTK-EGPAMIFNNVKGHPGARVAIGVLASRTRVGALLECDPKDLGKKLYHSVDNPIPP  
IBA QRPTK-EGPAMIFNNVKGHKDARVLI GLLASRKRVAALLGTEPEDLGKLLYHSVDNPIPP  
CBU QRPTK-EGPAMIFNNIKGHKDARVLI GLLASRRRVAALLDCEPENLGMMLCKSVENPIPP  
HHA KRPTK-EGPALLFRNIKGFDPKQVLI GLLASRKRVGYYLLDCPPDKLGFLLKDAANPVMP  
TSA MRPTK-IGPAMLFNHVKDYEHKSVLIGL FASRERVGLMLGCEPDRLGFLNLDALNHPVDP  
PAG KRPTR-VGPAMVFNKIKGFNDMRVLI GLLSSRQRVARLFGTSPENLAFMLKDSVKNPIPP  
HAL QRPTRKNGPAMVFNKIKGFDDISVTI GLNGSRKRKRVGHFLNCAPELGHLLKDSVKNAIQP  
SBL QRPTRKNGPAMVFHNKIGFRNTNVAI GLNGSRKRKRVGHFLNCAPEKLGFLKDSVKHAIAP  
KRA QRPTRKNGPVMVFNKIKGFQDISVAI GLNGSRKRKRVSHFLNCAPEKLGHLKDSVQNPPIPP  
KMI QRPTRKNGPVMVFNKIKGFKDISVAI GLNGSRQRVSHFLQCEPAKLGHLKDSVENAIAP

\*\*\*: \*\*.: \* : .. : \* ..\* : . : : \* : . : \*

LRH VTVE-ESAAPAHEVVAKASDPDFDIRKLVAAPTNTPQDAGPYITCGVVLGSNMAK-TMTD  
LPL ETVA-EADAPAHDVVYKATDEGFDIRKLVAAPTNTPQDAGPYITVGVVFGSSMDK-SKSD  
PET VMVD-EADAPTHEVVHKSTDAADFDIRKLVAAPTNTPRDAGPYITVGVVYGSNPK-TMSD  
ERA VKVA-KTDAPAQEVVHLATDEGFDIRKILAAPTNTEYDAGPYITMGVVFVGSNDPK-TMSD  
SGA VKVS-KDQAPAQEVVHLATDDDFDIRKILAAPTNTEYDAGPYITGLVYGSTPDK-SMSD  
ECL VVVP-ASQAPCQEQVYADDPDFDLRKLVPAPTNTPIDAGPFFCLGLVLASDPEDTSLTD  
ACH VVVP-REQAPCQEQVHRAEDPGFDLRKLVPAPTNTDEDAGPFFCLGLVLASDPEDTSLTD  
KIN VVLK-GQSAPCQEQIFRADDPHFDLRKLVPAPTNTTEEDAGPYFCLGLLLGSDDPN-GHTD  
ATR VVFK-EGSAPCQEQVFRADDPNFDLRHLLPAPTNTTEEDAGPYFCLGLLLGSDDPN-GHTD  
AGL APIAEGKHIDCQEVVHLATDPDFDLRRLVPAPTNTPEDAGPYVTMGLIAGHSPFN-SDTD  
ERY VVVA---KGICQEQIVHREDEGFDIRRLLPAIKATKEDAGPYITMGLCYASEPET-KESD  
COP VVIP-HEQAKCQEVKYFADDPDFDIRKLVPAPTNTLEDAGPYITLMGMCYASSPET-GESD  
BLA VDLE-G-DAPCQQVHKAADPDFDLYKLVPAPTNTPDDAGPYITLMGMCYATHPDT-GVHD  
FIR VEYQ-G-APPCQQVHKAEDPDFNLYDLVPAPTNTPDDAGPYITLMGMCYATHPDT-GVHD  
AHA VEYQ-G-DAPCQQVHKAEDPDFNLYDLVPAPTNTPDDAGPYITLMGMCYATHPDT-GVHD  
LAC VEYQ-G-DAPCQQVHKAEDPDFNLYDLVPAPTNTPDDAGPYITLMGMCYATHPDT-GVHD  
DLO VLTE-EA-APCQEVVHKATDPDFDLYKLVPAPTNTPVDAGPYITLMGMCYASHPDT-GVSD  
IBA VLTD--KTPLCQEVVHRATDEDFDLYKLVPAPTNTPDDAGPYITLMGMCYASHPDT-GCSD  
CBU VLTD-AKLPLCQQVHKAADPDFDLNKLVPAPTNTPDDAGPYITLMGMCYASHPDT-KFSD  
HHA EFTE-SGDVPCQEEIHYASDEGFDIRKILPAPQNTTEEDAGPYITMGMCYASDPVT-GDGD  
TSA IEIP-QAKAKCQEVVHLSTDEGFDIRKILPAPKNTLEDAGPYITMGLCYASDPDT-KAGD  
PAG IVIP-REHAVCQEVVHLATDPDFDILKILPTPTNTPEDAGPYITLMGMCYAADPET-GEHD  
HAL VTVS-ADKAVCQQVTHFAEDADFDLRKLVPAPTNTTEEDAGPYITMGLCYASDPDT-HESD  
SBL VDVS-TGNAVCQEVVHLATDPDFDLRKLVPAPTNTTEEDAGPYITMGLCYASDPDT-HESD  
KRA IL-T-KDNAVCQQVHHLASDTNFDLRKLVPAPTNTTEEDAGPYITMGLCYASDPET-HESD  
KMI VM-A-QGAAVCQQVHHLASDADFDRKLVPAPTNTTEEDAGPYITMGLCYASDPET-HESD  
:: \* \*:: :: : \* \*\*\*\*: . \*: . \*

LRH VTIHRMVLEDKDTLGIYIMPGGRHIGHFAEEYEKANKPMPVTINIGLDPAITIGATFEPP  
LPL VTIHRMVLEDKDKLGIYIMPGGRHIGHFAEEYEKANKPMPITINIGLDPAITIGATFEPP  
PET VTIHRMVLEDKDKLGIYIMPGGRHIGHKFAEEYEKMNKMPITINIGLDPAITIGATFEPP  
ERA VTIHRMVLEDEDTIGMYIMPGGRHIGHFQKQYEALNKMPITINIGLDPAISIGTTFEPP  
SGA VTIHRMVLEDKDTIGIYIMPGGRHIGHAFLESEYQKLNKMPITINIGLDPAILIGATFEPP  
ECL VTIHRLCVQERDELSMFLAAG-RHIEVFRKKAEEAGKPLPVTINMGLDPAIYIGACFEAP  
ACH VTIHRLCVQGRDRLSIFFAPD-RHIDKFRQKAEAAAGKPLPVTINMGLDPAILIGSCFEAP  
KIN VTIHRLCVQGRDELSVFFAPG-RHIDAFRAKAEERGEALPITINMGLDPAIPIGACFEAP  
ATR VTIHRLCVQGRDELSVFFAPG-RHIDAFRAKAEERGEALPITINMGLDPAIPIGACFEAP  
AGL VTIHRLCIESKDTMGMWITPGSRHLGAFFEQWKEKGEDMPVTVSIGLDPAVYMCAGFEAP  
ERY VTIHRLCLQSKDEMFMFFTPGIRHLDVFRKKAEEAKNKALPISISIGVDPALIEIACFEPP  
COP VTIHRMCFQSKDEISIFLQPGARHIGYFRELAEAKGEALPISISIGVDPALIEIACFEAP  
BLA VTIHRLCIQKDELSIFFTPGARHIGAMAERAEELGQRLPISISIGVDPALIEIACFEPP  
FIR VTIHRLCIQKDELSIFFTPGARHIGAMAERAEELGQKLPISISIGVDPALIEIACFEAP  
AHA VTIHRLCIQKDELSIFFTPGARHIGAMAERAEELGQKLPISISIGVDPALIEIACFEAP  
LAC VTIHRLCIQKDELSIFFTPGARHIGAMAERAEELGQKLPISISIGVDPALIEIACFEAP  
DLO VTIHRLCIQKDELSIFFTPGARHIGAMAERAEELGQRLPISISIGVDPALIEIACFEPP  
IBA VTIHRLCIQKDELSIFFTPGARHIGAMAERAEELGQKLPISISIGVDPALIEIACFEPP  
CBU VTIHRLCIQKDELSIFFTPGARHIGAMAERAEELGQNLPLISISIGVDPALIEIACFEPP  
HHA VTIHRLCLQSADEMFMFFTPGVRHLDAFREKAEREGVNLPLISISIGVDPALIEIACFEPP  
TSA ITIHRMLCQSKDELTMFFTPGVRHLDAFREKAEREGVNLPLISISIGVDPALIEIACFEPP  
PAG VTIHRLCVQSKDEISMYFVPG-RHLDTFRQKAEAAAGKPLPITISIGVDPALIEIACFEPP  
HAL ITIHRMLCVQSRDELSMWLTPG-RHIDAFRMAKAEAAAGKPLPISISIGVDPALIEIACFEPP  
SBL ITIHRMLCVQSKDELTMWLTPG-RHIDAFRIKAEAAAGKALPISISIGVDPALIEIACFEPP  
KRA ITIHRMLCVQSRDELSMWLTPG-RHIDAFRMAKAEAAAGKPLPISISIGVDPALIEIACFEPP  
KMI ITIHRMLCVQSRDELSMWLTPG-RHIDAFRMAKAEAAAGQPLPISISIGVDPALIEIACFEPP  
:\*\*\*\*: .: \* : :: . \*\*: : : \*:::.\*:\*\*\*\*: : : \*\* \*



LRH RAI PGKVKNVYNPPAGGGKLMTIMQIHKDNPADEGIQRQAALLAFSAFKEKLTWLV VDDD  
LPL RAI PGKVTNVYNPPAGGGKLMTIMQIHKDNEADEGIQRQAALLAFSAFKEKLTVIL VDED  
PET RAI PGKVLNVYNPPAGGGKLMTIMQIHKDNEADEGIQRQAALLAFSSFKELKTVFL VDED  
ERA KAI PGKVTNVYNPPAGGGKLMSTILQIHKESAEDEGIQRQAAILLALSAFKEKLTVIL VDDD  
SGA KAI PGKVLNVYNPPAGGGKLMTIMQIRKENPADEGIQRQAALLAFSSFKELKTVIL VDED  
ECL EAI PGFLQNVYAHTAGGGKFLGILQVKKRQPSDEGRQQAALIALATYSELKNIIL VDED  
ACH RAMPGLLNVAHTAGGGKLLAVLQVAKRRPGDEGRQQAALIALAVYRELKNVIL VDED  
KIN KALPGFVKNVYAHSAAGGGKLLAILQVCQRSAGDVGKARQAALIALAVYRELKNVIL VDDD  
ATR KALPGFVKNVYAHSAAGGGKLLAILQVCQRSAGDAGKARQAALIALAVYRELKNVIL VDDD  
AGL TAI PGRLNVAAPCGGGKFVAVLQFKKSSINDEGRQRNAAMLAFSAFSELKHFVFL VDED  
ERY KAMPGNVRNVHCASSGGGKYIAVLQFVKRMESDEGRQQAALLAFSAFSELKHFVFL VDED  
COP KAMPGLQNVYCCASSGGGKYIAVMQFKKSVPSDEGRQQAALLAFSAFAELKHFVFL VDED  
BLA KAMPGLQNVYCCASSGGGKYMAVLQFKKLSQSDEGRQQAALLAFSAFSELKNVFL VDED  
FIR KAMPGLQNVYCCASSGGGKYMAVLQFKKREASDEGRQQAALLAFSAFSELKHFVFL VDED  
AHA KAMPGLQNVYCCASSGGGKYMAVLQFKKREASDEGRQQAALLAFSAFSELKHFVFL VDED  
LAC KAMPGLQNVYCCASSGGGKYMAVLQFKKREASDEGRQQAALLAFSAFSELKHFVFL VDED  
DLO KAMPGLQNVYCCASSGGGKYMAVLQFKKLTASDEGRQQAALLAFSAFSELKHFVFL VDED  
IBA KAMPGLQNVYCCASSGGGKYMAVLQFKKSVASDEGRQQAALLAFSAFSELKNIFIL VDED  
CBU KAMPGLQNVYCCASSGGGKFMAVLQFKKTVASDEGRQQAALLAFSAFSELKNIFL VDED  
HHA KALAGRVKNVYAHSSGGGKYMAVIQFAKQVPSDEGRQQAALIALTAFPELKHVIV VDED  
TSA RAMPGRVLNVYAHSSGGGKYLAAILQFKKSI PSDEGRQQAALVAFSAFPELKHVIL VDED  
PAG RALPGFVQNVHCPSPGTGKYLAVLQVKKRFVAVDEGRQQAALLAFSAFSELKHFVFL VDED  
HAL RAMPGLLNVAHTAGGGKLLAVMVFQFKKSAPVDEGRQQAALLAFSAFSELKHFVIL VDED  
SBL RAMPGLLNVAHTAGGGKLLAVMVFQFKKSSVNDDEGRQQAALLAFSAFPELKHVIL VDED  
KRA RAMPGLLNVAHTAGGGKLLAVMVFQFKKSSPADEGRQQAALLAFSAFPELKHVIL VDED  
KMI RAMPGLLNVAHTAGGGKLLAVLQFKKSSPADEGRQQAALLAFSAFPELKHVIL VDED  
\* : \* : \* . \* \* \* : : \* . : \* \* : \* \* : \* \* : \* \* : \* \* : \* \* : \*

LRH VDIF DMNDVWWTMNTREFQGDQDIMVLPGMRNHPLDP SERPQYDPKSI RVRGMSSKTVIDG  
LPL VDIF DMNDVIWWTMNTREFQADQDIMVLSGMRNHPLDP SERPQYDPKSI RFRGMSSKLVIDG  
PET VDIF DMNDVWWTINNTREFQADQDIMVLPGMRNHPLDP SERPEYDPKSI RTRGMSSKLVIDG  
ERA VDIF DMNDVMWTLNNTREFQGDQDILVLPGMRNHPLDP SERPEYDPKSI RFRGMSSKTIIDG  
SGA VDIF DMNDVMWTINNTREFQAHKDIMSLEGMRNHPLDP SERPEYSPEHIRVRGMSSKLVLDG  
ECL VDIF DSDDILWAMTTRMQGDVSTITTLPGIRGHQLDPSQSPDYST-SIRNGISCKTIFDC  
ACH VDPF DSDDVLWAMQTRYQGDVDTIFVPGVPGHVLDP SQVPEYSP-SIAARGLTCKTIFDC  
KIN VDLE DSNDVLWAMQTRYQGNVDTMFLPGVTGHVLDPSQVPEYDP-SIPAKGV SCKTIFDC  
ATR VDLE DSNDVLWAMQTRYQGNVDTMFLPGVTGHVLDPSQVPEYDP-SIPAKGV SCKTIFDC  
AGL VDIF DMSDVMWAMTTRFQADTGLITIPGAHCHVLDPSNDPAFAP-SIRVHGIACKAIFDC  
ERY VDIY DMKDVLWAMTTRFQSDIDCISIPGVRCHPLDPSNDTTYDP-SIRDARGIACKAIFDC  
COP VDPF DMKDVMWAMTTRFQADLDLITIPGVQCHPLDPSNQPEYSA-HIRARGIACKAIFDC  
BLA VDCE DMNDVLWAMNTRFQGDVDVITIPGVRCHPLDPSNDPAFSP-SIRDHGIACKTIFDC  
FIR VDCE DMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPTCSP-SIRDHGIACKTIFDC  
AHA VDCE DMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPTCSS-SIRDHGIACKTIFDC  
LAC VDCE DMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPTCSS-SIRDHGIACKTIFDC  
DLO VDCE DMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPSCSG-SIRDHGIACKTIFDC  
IBA VDCE DMNDVLWAMNTRFQGDVDVITIPGVRCHPLDPSNDPSPFSP-SIRDHGIACKTIFDC  
CBU VDCE DMNDVLWAMNTRFQGDADIITIPGVRCHPLDPSNDPDYSP-SIKNHGIACKTIFDC  
HHA VDIF DSDDVLWALNTRYQGDVDTITIPGVRCHPLDPTEGPEYNP-MLKDRGISCKTIFDC  
TSA VDIF DSNDVLWALNTRYQGDLDTVFIPGVRCHPLDPSNNLYHP-SILSTGIACKTIYDC  
PAG VDVF DLSDVMWAMTTRYQGDVSTVFIPGVRCHPLDPSNDPAFSP-SVRDHGITCKTIFDC  
HAL VDIF DSDDVMWAMQTRYQGDVDTVFIPGVRCHPLDPSQMPDYSP-SILQEGMSCKTIFDC  
SBL VDIF DTDDVLWAMQTRYQGDIDTITIPGVRCHPLDPSQVPEYSP-FITQQGMTCKTIFDC  
KRA VDIF DSDDVLWAMQTRYQGDVDTVITIPGVRCHPLDPSQIPAYSP-SILQQGMSCKTLFDC  
KMI VDIF DSDDVLWAMQTRYQGDVDTIVIPGVRCHPLDPSQVPEYSP-SILQQGMSCKTIFDC  
\* \* : \* . \* \* \* : : \* . : \* \* : \* \* : \* \* : \* \* : \* \* : \* \* : \*

```

LRH      TVPFDMRDQFKRAAFKKVS--DWQKYLK-----
LPL      TVPFDMKDQFERAQFMKVA--DWEKYLK-----
PET      TVPFDMKDQFERAQFKEVK--DWEKYLK-----
ERA      TVPFDLKDDFIRAEFKEVP--DWEKYLK-----
SGA      TVPFDMKDQFERAKFKEVP--DWKKYLD-----
ECL      TVPWALKARFERAPFMEVDPTPWAPELFSDKK-
ACH      TAPWHLRERFERAKFRDVPDHPFAPFLFPQARS
KIN      TYPWKLKEHFVRAQFRDVPDHPFAPSIFPKAGM
ATR      TYPWKLKEHFVRAQFREVDHPFAPSIFPKAGM
AGL      TYPYDLKDEFQRCRFLDIDQDKWAAELAH----
ERY      TVPFSQKERFQRAAFQVEEASQWLK-----
COP      TVPFDQKARFERAKFMDVDPKCWLDPMLK----
BLA      TVPYDLKEDFHRARFMELDPEKWLKK-----
FIR      TVPYDQKERFKRARFMEVDPEHWL-----
AHA      TVPYDQKERFKRARFMEVDPEHWLS-----
LAC      TVPYDQKERFKRARFMEVDPEHWLS-----
DLO      TVPYDQKDRFKRAQFMDVDPEHWVK-----
IBA      TVPFEQKERFKRARFMDVDPEHWLKD IK-----
CBU      TVPFHMKERFKRAKFMVDPEHWL-----
HHA      TVPYGLKDRFQRSKFKEVNME DYEIRPLV----
TSA      TVPFIQKERFKRSSFKEVNLS DYDIKPLK----
PAG      TVPYNLKANFHRSEFLEVDVNRFI PGFNQK---
HAL      TVPFHLKAHFERSRFKEADV KRFLP DFE-----
SBL      TVPFHLKTHFERSTFKEVDV KRFLP DFE-----
KRA      TVPFHLKSHFERSKFKEVDV KRFLP DFE-----
KMI      TVPFHLKHTFQRSRFKEVDV KRFLP DFE-----
* *: : * * . * . :

```

**FIG S2.** Comparison of amino acid sequences of subunit C of putative gallate decarboxylase proteins from bacteria identified by HTS from athlete faecal microbiota. Multiple alignments were done using the programs ClustalOmega after retrieval of sequences from BLAST homology searches. The C subunits are from *Acetobacter tropicalis* (ATR) (WP\_006559787.1), *Actinomyces glycerinitolerans* (AGL) (WP\_073329259.1), *Anaerostipes hadrus* (AHA) (WP\_009204323.1), *Azotobacter chroococcum* (ACH) (WP\_052264016.1), *Blautia* sp. KLE 1732 (BLA) (WP\_021650680.1), *Clostridium butyricum* (CBU) (WP\_058371993.1), *Coprobacillus* (COP) (WP\_008787659.1), *Dorea longicatena* CAG:42 (DLO) (CDE20644.1), *Enterobacter cloacae* subsp. *cloacae* ATCC 13047 (ECL) (YP\_003612445.1), *Enterococcus* (ERA) (WP\_010743655.1), *Erysipelotrichaceae* bacterium 3\_1\_53 (ERY) (EFP59895.1), Firmicutes bacterium CAG:270 (FIR) (CDD72954), *Hafnia*

*alvei* ATCC 51873 (HAL) (WP\_004092226.1), *Hungatella hathewayi* (HHA) (WP\_006771947.1), *Intestinibacter bartlettii* (IBA) (WP\_007285641.1), *Klebsiella michiganensis* KCTC 1686 (KMI) (AEX02211.1), *Komagataeibacter intermedius* (KIN) (WP\_039733191.1), *Kosakonia* (KRA) (WP\_071921386.1), *Lachnospiraceae* bacterium 5\_163FAA (LAC) (EFV16561), *Lactobacillus rhamnosus* (LRH) (WP\_005712295.1), *Lactobacillus plantarum* ATCC 14917T (LPL) (D7VDD5), *Pantoea agglomerans* strain FDAARGOS\_160 (PAG) (AMG60167.1), *Pediococcus ethanolidurans* (PET) (WP\_057806460.1), *Shimwellia blattae* (SBL) (WP\_002441866.1), *Streptococcus* (SGA) (WP\_003065832.1), and *Turicibacter sanguinis* (TSA) (WP\_040763984.1). Residues that are identical (\*), conserved (:) or semiconserved (.) in all sequences are indicated. Dashes indicated gaps introduced to maximize similarities. Degenerate primers were designed on the conserved domains highlighted.