**Supplementary material**

**Supplementary Figure 1.** Nucleotide sequence of the KIPyV Stockholm 60 NCCR.

1 TTTGCCTCTC TAGGCCTCTC AAATGCCTCT GCAGGCCCTC TCCTTCTTCT AGAAAAAGCT 60

61 GGGGCTTTTT TGGCCTCTGG CCTCCTGTAA TATAGAAAAA AAGGGCACAG TGATTGACAG 120

121 TTGTGTATAC AAGCATGTGT GGTATGTTTA GTGTGTAAGC CAATAAAGTT AAAGGTCACT 180

181 ACTGTGGGTG GTGACACCTG ATACCGGCGG AACTAGTTGC TACGTGCCAC ACAATAGCTT 240

241 TCACTCTTGG CGTGAAGCCA ACTTCCTGGG CCGTGAGCCA GCTTCCTGCG GCCTTGTGTT 300

301 TTTTACCAAC ACACCTGGTG AACTTCTACT GTCCTTGACA CAGGTAAGAC TGGGGACCCT 360

361 TTGTAGGCCA AAGGAGAGTG AAGGGTAACT GAAATGCTAA GACTGTAAGT TCTAATCCTA 420

421 GTATTTCAGT TCGGGGATGT TGGCGCCATC GTCTCGAACC TGGCCTGCAT ACCTTTGGAT 480

481 ATAGAGGGTC ACCAATTTTT ATTTTGTTTT TAG 513

**Supplementary Table S1.** Origin of KI NCCR variants and number of mutations compared to the Stockholm 60 NCCR.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Strain | Source | GenBank | Identity with Stockholm 60 NCCR | Country | Reference |
| Stockholm 60 | NPA | EF127906 |  | Sweden | 1 |
| Stockholm 350 | NPA | EF127907 | 4 point mutations | Sweden | 1 |
| Stockholm 380 | NPA | EF127908 | 4 point mutations | Sweden | 1 |
| Brisbane 001 | NPA | EF520287 | 10 bp insert; 3 point mutations | Australia | 2 |
| Brisbane 002 | NPA | EF520288 | identical with Stockholm 60 | Australia | 2 |
| Brisbane 003 | NPA | EF520289 | 3 point mutations | Australia | 2 |
| Brisbane 004 | NPA | FJ150412 | identical with Stockholm 60 | Australia | 3 |
| Brisbane 005 | NPA | FJ150410 | identical with Brisbane 001 | Australia | 3 |
| Brisbane 006 | BAL | FJ150411 | identical with Stockholm 60 | Australia | 3 |
| Brisbane 007 | faeces | FJ150413 | identical with Stockholm 60 | Australia | 3 |
| CU-255 | NPA | EU358766 | identical with Brisbane 001 | Thailand | 4 |
| CU-258 | NPA | EU358767 | identical with Stockholm 60 | Thailand | 4 |
| FZ52 | NPA | KM085447 | 4 point mutations | China | 5 |
| HN057 | faeces | KC571691 | 5 point mutations | China | 6 |

NPA: nasopharyngeal aspirate; BAL: bronchoalveolar lavage

1. **Allander, T., Andreasson, K., Gupta, S., Bjerkner, A., Bogdanovic, G., Persson, M.A., Dalianis, T., Ramqvist, T., & Andersson, B. (2007).** Identification of a third human polyomavirus. *J Virol* **81,** 4130-4136.
2. **Bialasiewicz, S., Whiley, D.M., Lambert, S.B., Wang, D., Nissen, M.D., & Sloots, T.P. (2007).** A newly reported human polyomavirus, KI virus, is present in the respiratory tract of Australian children. *J Clin Virol* **40,** 15-18.
3. **Bialasiewicz, S., Whiley, D.M., Lambert, S.B., Nissen, M.D., & Sloots, T.P. (2009).** Detection of BK, JC, WU, or KI polyomaviruses in faecal, urine, blood, cerebrospinal fluid and respiratory samples. *J Clin Virol* **45**, 249-254.
4. **Payungporn, S., Chieochansin, T., Thongmee, C., Samransamruajkit, R., Theamboolers, A., & Poovorawan, Y. (2008).** Prevalence and molecular characterization of WU/KI polyomaviruses isolated from pediatric patients with respiratory disease in Thailand. *Virus Res* **135,** 230-236.
5. **Xiu,W., Zheng,K., Liu,G., Kang,Y., Wu,B., & Chen,W.** Sequencing and phylogenetic analysis of the complete genome sequence of KI Polyomavirus(KIPyV) in Fuzhou, China. Unpublished.
6. **Li, K., Guo, J., Zhao, R., Xue, Y., Chen, L., Yang, J., Peng, J., & Jin Q.** **(2013).** Prevalence of 10 human polyomaviruses in fecal samples from children with acute gastroenteritis: a case-control study. *J Clin Microbiol* **51,** 3107-3109.

**Supplementary Table S2.** Primers used to generate the KIPyV NCCR mutants.

|  |  |  |
| --- | --- | --- |
| Mutant | Primers | Mutation\* |
| Bld22c | F: 5’-CAGGCCCTCTCCTCCTTCTAGAAAAAGC-3’  R: 5’-GCTTTTTCTAGAAGGAGGAGAGGGCCTG-3’ | T45C |
| Bld29d | F: 5’-CTCAAATGCCTCTGCGGGCCCTCTCCTTC-3’  R: 5’-GAAGGAGAGGGCCCGCAGAGGCATTTGAG-3’ | A33G |
| Bld38c | F: 5’-GGCCTCTGGCCTCTTGTAATATAGAAAAAAAGGGC-3’  R: 5’-GCCCTTTTTTTCTATATTACAAGAGGCCAGAGGCC-3’ | C85T |
| Bld38d | F: 5’-GGAGAGTGAAGGGCAACTGAAATGCTAAG-3’  R: 5’-CTTAGCATTTCAGTTGCCCTTCACTCTCC-3’ | T386C |
| Bld45d | F: 5’-GTGAAGGGTAACTGAGATGCTAAGACTGTAAG-3’  R: 5’-CTTACAGTCTTAGCATCTCAGTTACCCTTCAC-3’ | A393G |
| Bld60c | F: 5’-GCCTCTCAAATGCCCCTGCAGGCCCTCTCC-3’  R: 5’-GGAGAGGGCCTGCAGGGGCATTTGAGAGGC-3’ | T28C |
| Bld60d | F: 5’-CTGTGGGTGGTGACGCCTGATACCGGCGG-3’  R: 5’-CCGCCGGTATCAGGCGTCACCACCCACAG-3’ | A196G |
| Bld62c | F: 5’-GACACCTGATACCGGTGGAACTAGTTGCTAC-3’  R: 5’-GTAGCAACTAGTTCCACCGGTATCAGGTGTC-3’ | C208T |
| Bld71c | F: 5’-GCCAGCTTCCTGCGACCTTGTGTTTTTTACC-3’  R: 5’-GGTAAAAAACACAAGGTCGCAGGAAGCTGGC-3’ | G291A |
| NPA1a=NPA74d | F: 5’-CCTTTGGATATAAAGGGTCACC-3’  R: 5’-GGTGACCCTTTATATCCAAAGG-3’ | G484A |
| NPA7b | F: 5’-CCTTCTTCTAGAAAAGGCTGGGGCTTTTTTGG-3’  R: 5’-CCAAAAAAGCCCCAGCCTTTTCTAGAAGAAGG -3’ | A57G |
| NPA7d | F: 5’-GTGTAAGCCAATAGAGTTAAAGGTCACTAC-3’  R: 5’-GTAGTGACCTTTAACTCTATTGGCTTACAC-3’ | A166C |
| NPA44A | F: 5’-CCGGCGGAACTAGATGCTACGTGCCACAC-3’  R: 5’-GTGTGGCACGTAGCATCTAGTTCCGCCGG-3’ | T217A |
| NPA57a | F: 5’-CCTGTAATATAGAAAAAAAGGGCACAGTGATTG-3’  R: 5’-CAATCACTGTGCCCTTTTTTTCTATATTACAGG-3’ | A101T |
| NPA69c | F: 5’-GGCTTTTTTGGCCTCCGGCCTCCTGTAATATAG-3’  R: 5’-CTATATTACAGGAGGCCGGAGGCCAAAAAAGCC-3’ | T78C |
| NPA74c=NPA81c | F: 5’-GGTATGTTTAGTGCGTAAGCCAATAAAG-3’  R: 5’-CTTTATTGGCTTACGCACTAAACATACC-3’ | T154C |
| NPA89d | F: 5’-GTGTGTAAGCCACTAAAGTTAAAGGTCAC-3’  R: 5’-GTGACCTTTAACTTTAGTGGCTTACACAC-3’ | A163C |
| NPA94a | F: 5’-GGTCACTACTGTGGGCGGTGACACCTGATACC-3’  R: 5’-GGTATCAGGTGTCACCGCCCACAGTAGTGACC-3’ | T189C |
| Brisbane 001 ins | F: 5’-GACAGTTGTGTATACAAAGGCGCTGCGGCATGTGTGGTATG-3’  R: 5’-CATACCACACATGCCGCAGCGCCTTTGTATACACAACTGTC-3’ | insertion |
| Brisbane 001 double | F: 5’-GCCAATAAAGTTAAAGGTCACCACAGTGGGTGGTGACACC-3’  R: 5’-GGTGTCACCACCCACTGTGGTGACCTTTAACTTTATTGGC-3’ | T180C/T183A |
| Brisbane 001 C306A | F: 5’-GCGGCCTTGTGTTTTTTAACAACACACCTGGTGAAC-3’  R: 5’-GTTCACCAGGTGTGTTGTTAAAAAACACAAGGCCGC-3’ | C306A |

\*The numbers refer to the nucleotide number in the NCCR.

**Supplementary Table S3.** NCCR mutations compared to Stockholm 60 in different isolates and effect on putative transcription factor binding sites. The prediction of transcription factor binding sites is based on the ALGGEN PROMO algorithm [1,2].

|  |  |  |
| --- | --- | --- |
| Sample | Mutations | Putative transcription factor sites affected  Removed Added |
| Bld22a  Bld22b  Bld22c  Bld22d | T45C  T19C,C74T | none p300,BTEB3,Elk-1,Nkx2-1  POU2F2,NF1,Nkx2-1 EIIaE-A,p300,R2,NF1,Elk-1,NHP1 |
| Bld29a  Bld29b  Bld29c  Bld29d | T28C  A33G | AP-2A,NKx2.1,NF-X3,GATA-2 POU3F2,p53,EIIaE-A,ETF,Sp3  E2F-1,TFIIB AP-2A |
| Bld37a  Bld37b  Bld37c | A162T | NF1,ENKTF-1,CULT1,NF-Y AP-2A,TMF |
| Bld38a  Bld38b  Bld38c  Bld38d | C85T  T386C | p300,Elk-1 SRY  VDR C/EBP,C/EBP,LF-A1,p53,Myf-3,MyoD,RFX1 |
| Bld45a  Bld45b  Bld45c  Bld45d | A393G | POU2F2,POU3F2,Nrf2 TFIIB |
| Bld60a  Bld60b  Bld60c  Bld60d | T28C  A196G | AP-2A,NKx2.1,NF-X3 POU3F2,p53,EIIaE-A,ETF,Sp3  NHP-1,Myf-3,MyoD,AREB6,Max,TGIF AP-2A,Pax5 |
| Bld61a  Bld61b  Bld61c  Bld61d | T93C,A173G,A381T,  A415G,C455T  T41C,T111C,T411C | TMF,Pax6,AP-3,HNF-1B FOXO3a,VDR,Hlf  c-Myb MZF-1,p300,R2,Elk-1 |
| Bld62b  Bld62c  Bld62d | G401A | E2F none |
| Bld69a  Bld69b  Bld69c  Bld69d | T121A,T294C | SRY,c-Myb Pbx1b,Hlf |
| Bld71a  Bld71b  Bld71c  Bld71d | G291A | AP2A none |
| NPA1a  NPA1b  NPA1c | G484A  C77T,G188A | GATA-2 none  NKx2-1,NF-X3,E47,WT1 I,VDR NF1,TCF-1A,TCF-4E,LEF-1,p300,R2,Elk-1,YY1,PEA3 |
| NPA7a  NPA7b  NPA7c  NPA7d | T28C,C46T  A57G  A166G | AP2,Nkx2,NF-X3,R2 p53,EIIaE-A,ETF,Sp3  Pu box binding factor  c-Myb |
| NPA12a  NPA12b  NPA12c  NPA12d  NPA12e  NPA12f  NPA12g | T248A,A371G | NF1,ENKTF-1,TCF-4E,TCF-1A,LEF1,E2F AP2aA,BTEB3,E47,Nkx2.1,NF-X3 |
| NPA19a  NPA19b  NPA19c  NPA19d | G188A,T247C,G484A  G224A | WT1 I,VDR,GATA-2 p300,R2,Elk-1,PEA3,YY1  none none |
| NPA44a  NPA44b  NPA44c  NPA44d | T217A  C36T  T297C,T300C | C/EBP,C/EBP,c-Myb Oct2.1  POU3F2,p53,EIIaE-A NHP-1,AREB6  SRY,PR-A,PR-B,GR,GR C/EBP,C/EBP,Pax5 |
| NPA57a  NPA57b  NPA57c  NPA57d | A101T  delG352,T416C | none Oct2.1,POU3F2,HNF3,FOXP3  MZF1,E2F:DP1 TFIIB,NF1,PPAR:RXR |
| NPA69a  NPA69b  NPA69c  NPA69d | A234G,G484A  T78C | CUTL1,SRY,GATA-2 none  AP2A,NF-X3 p300, R2, Elk-1 |
| NPA74a  NPA74b  NPA74c  NPA74d  NPA74e  NPA74f  NPA74g | T154C  G484A  A305G,G484A | none Pax5  GATA-2 none  HBNF-3,GATA-2 C/EBP,C/EBP,NF1,LF-A1,ENKTF-1 |
| NPA76a |  |  |
| NPA81a  NPA81b  NPA81c  NPA81d | T297C,T336C,G484A  A173G,C444T  T154C  T60C,G116A | SRY,PR-A,PR-B,GR,GR,Pax2,USF2b,c-JUN,C-FOS C/EBP,C/EBP,Pax5,Nkx2-1  NHP-1, TGIF,GATA-2  ENKTF-1,E2F FOXO3a,VDR,LF-A1  none Pax5  E47,Pu-box binding factor,Pax2, PR-A,PR-B,HNF-1A,HNF-1B,GR  PPAR:RXR,USF2b,CULT1,TGIF, c-JUN,c-FOS,NHP-1,  PKNOX1 |
| NPA89a  NPA89b  NPA89c  NPA89d | A231G,G290A  A99G,T235A,C418T  A163C | SRY,WT1 I,E2F-1 C/EBP,C/EBP,Pax5,Arnt,AhR  CUTL1,p300,Elk-1 TGIF  NF1,CUTL1,NF-Y none |
| NPA94a  NPA94b | T189C  C8A,C10T,T30A | VDR p53,EIIaE-A,Sp1,E2F  Nkx2-1,NF-X3 POU3F2,HNF-3,HNF-3,AP3,TFIIB |

1. **Messeguer, X., Escudero, R., Farre, D., Nunez, O., Martinez, J., & Alba, M. M.** **(2002).** PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics,* **18**, 333-334.

2. **Farre, D., Roset, R., Huerta, M., Adsuara, J. E., Rosello, L., Alba, M. M., & Messeguer, X.** **(2003).** Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Res,* **31**, 3651-3653.