Accepted Manuscript

The micro-evolution and epidemiology of *Staphylococcus aureus* colonization during atopic eczema disease flare

Catriona P. Harkins, Kerry A. Pettigrew, Katarina Oravcová, June Gardner, R.M.Ross Hearn, Debbie Rice, Alison E. Mather, Julian Parkhill, Sara J. Brown, Charlotte M. Proby, Matthew T.G. Holden

PII: S0022-202X(17)32980-9
DOI: 10.1016/j.jid.2017.09.023
Reference: JID 1096

To appear in: *The Journal of Investigative Dermatology*

Received Date: 15 May 2017
Revised Date: 11 August 2017
Accepted Date: 12 September 2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
The micro-evolution and epidemiology of *Staphylococcus aureus* colonization during atopic eczema disease flare

Catriona P. Harkins (ORCID:0000-0002-9099-7291)¹,²,³, Kerry A. Pettigrew (ORCID:0000-0002-6027-0462)¹, Katarina Oravcová (ORCID:0000-0001-5930-6803)¹,⁴, June Gardner², R.M.Ross Hearn², Debbie Rice⁵, Alison E. Mather⁶, Julian Parkhill (ORCID:0000-0002-7069-5958)⁷, Sara J. Brown (ORCID:0000-0002-3232-5251)²,⁸, Charlotte M. Proby²,³ Matthew T.G. Holden (ORCID:0000-0002-4958-2166)¹

¹ School of Medicine, University of St Andrews, St Andrews, KY11 9TF, UK
² Department of Dermatology, Ninewells Hospital, Dundee, DD1 9SY, UK
³ School of Medicine, University of Dundee, Dundee, DD1 9SY, UK
⁴ Institute of Biodiversity, Animal Health and Comparative Medicine, College of Medical, Veterinary and Life Sciences, University of Glasgow, G12 8QQ, Glasgow, UK
⁵ Scottish Children’s Research Network, MACH 2 Building, Level 5, Ninewells Hospital, Dundee, DD1 9SY, UK
⁶ Department of Veterinary Medicine, University of Cambridge, Cambridge, CB3 0ES, UK
⁷ The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK
⁸ Skin Research Group, Division of Cancer Research, School of Medicine, University of Dundee, DD1 9SY, UK

**Study location:** Dundee, Scotland, UK

**Corresponding author:** Catriona P. Harkins - E-mail: ch246@st-andrews.ac.uk

Department of Dermatology, Ninewells Hospital and Medical School, Dundee, DD1 9SY, UK, Tel: +44 7542 250517, Fax: +44 1382 633925.

**Short Title:** Micro-epidemiology *S. aureus* colonization

**Abbreviations:** *S. aureus*, *Staphylococcus aureus*; AE, Atopic eczema; CC, clonal complex; ST sequence type; WGS, Whole Genome Sequencing; SNP, single nucleotide polymorphisms.
ABSTRACT

Staphylococcus aureus is an opportunistic pathogen and variable component of the human microbiota. In atopic eczema (AE) a characteristic of the disease is colonization by S. aureus, with exacerbations associated with an increased bacterial burden of the organism. Despite this, the origins and genetic diversity of S. aureus colonizing individual patients during AE disease flares is poorly understood. To examine the micro-evolution of S. aureus colonization we have deep-sequenced S. aureus populations from nine children with moderate to severe AE, and 18 non-atopic children asymptotically carrying S. aureus nasally. Colonization by clonal S. aureus populations was observed in both AE cases and controls, with all but one of the individuals containing colonies belonging to a single sequence type. Phylogenetic analysis revealed that disease flares were associated with the clonal expansion of the S. aureus population, occurring over a period of weeks to months. There was a significant difference in the genetic backgrounds of S. aureus colonizing AE patients versus controls (Fisher’s Exact test, p=0.03). Examination of intra-host genetic heterogeneity of the colonizing S. aureus populations identified evidence of within-host selection in the AE patients, with AE variants being potentially selectively advantageous for intracellular persistence and treatment resistance.
INTRODUCTION

Clinical studies have demonstrated a link between *Staphylococcus aureus* and the pathogenesis of atopic eczema (AE). Affected individuals are characteristically prone to colonization by this pathogen (Leyden et al. 1974; Hauser et al. 1985), with disease severity correlating to bacterial burden and number of colonized sites (Tauber et al. 2016).

Increasingly there has been a shift towards understanding how cutaneous dysbiosis contributes to AE aetiology (Chng et al. 2016; Kennedy et al. 2016). Metagenomic analysis has shown that changes in populations of microbial communities are significantly associated with disease activity in AE, and that increased in *S. aureus* is linked to increasing severity (Kong et al. 2012). Whilst providing a more holistic overview of microbial population structure in AE, current metagenomic approaches have thus far failed to resolve the fine-scale population dynamics of *S. aureus* in AE, or the genetic changes occurring in the colonizing population of this versatile pathogen.

Whole genome sequencing (WGS) is a high-resolution genotyping tool that can be used to study within host evolution and transmission. Deep sequencing studies of *S. aureus* populations have revealed heterogeneity arising within the host and the impact on the disease causing potential of the population (Young et al. 2012; Paterson et al. 2015; Azarian et al. 2016). By using WGS to genetically characterize bacteria it is possible to investigate how populations differentiate and adapt within a host during colonization as well as reconstructing the evolutionary events shaping populations (Didelot et al. 2016). In this study we deep sequenced *S. aureus* populations from children with AE, enabling us to investigate the micro-evolution of colonization during disease flare. From this we have uncovered evidence of selection enriching for colonization by *S. aureus* of specific genetic backgrounds, as well as
genetic diversification promoting the survival and persistence of these strains during colonization of AE patients.

RESULTS

*S. aureus colonization in cases and controls*

Nine AE cases were recruited through Ninewells Hospital in Dundee, and skin swabs obtained from 5 body sites, including a nostril, 2 areas of inflamed eczema skin, and 2 separate areas of clinically unaffected skin. All were colonized by *S. aureus* at one or more sampled sites. Eighteen community *S. aureus* nasal carriers were selected from the larger control study population, from whom swabs were taken from a single nostril. Extra-nasal skin swabs obtained from controls were in all instances negative for *S. aureus*. Control age matching was attempted but age-appropriate carriers identified had a history of atopy and were therefore excluded on this basis (Table 1).

All cases had generalized moderate-to-severe eczema, the exception being patient 5, who had locally severe disease only (Table S1). Four of the nine cases (44%) were nasally colonized in addition to eczema-affected skin, whilst seven (78%) were also colonized on clinically unaffected skin. Bacterial burden varied across colonization sites within individuals, and also between individuals (Table S2). Where available, 5 colonies from the primary isolation plates were randomly selected per swabbed body site to provide representative sampling and detect co-colonization (Votintseva et al. 2014). In cases with low bacterial burden, fewer than 5 colonies were available for WGS. Eczema site sub-sampling was included to investigate *S. aureus* population heterogeneity within disease sites. We undertook WGS of between 10 and 28 colonies per AE case dependent on recovered colony counts.
AE patients are colonized with distinct clonal populations of S. aureus

All controls, and all but one case, were colonized by clonal populations of S. aureus represented by a single sequence type (ST) defined by multilocus sequence typing (MLST) (Table S3). The exception, Patient 8, was co-colonized by two distinct STs. Comparison of the genetic backgrounds of the S. aureus colonizing populations in cases and controls revealed different ST distributions (Figure 1). Eczema cases were more frequently colonized with STs belonging to clonal complex (CC) 1; 4 of the 9 AE cases (44%) carried STs belonging to CC1 (ST1 and ST188), compared with a single control (6%; Fisher’s Exact test, p=0.03). Conversely, controls were principally colonized with CC30; 11 of the 18 controls (61%) carried STs belonging to CC30 (ST30 and ST2889), compared with one case (11%; Fisher’s Exact test, p=0.02). This suggested a subset of the S. aureus population is more frequently associated with colonization in AE. Overall, the proportions of CCs of isolates from cases and controls in revealed a significant difference in the distribution in CC of strains on the basis of disease status (Fisher’s Exact test; p=0.005).

Clonal expansion of S. aureus populations in AE patients

The observed clonal populations in cases suggest that during disease flare there is a clonal expansion within the host, i.e. the S. aureus growing on eczema skin originates from a small progenitor population within the individual. To provide a high-resolution view of the colonizing populations we characterized the genomic diversity of the isolates and examined their phylogenetic relationships. As a measure of the relative within-host diversities of S. aureus populations we examined the core genome of the sequenced isolates to identify single nucleotide polymorphisms (SNPs). The levels of diversity we observed in both cases and controls were comparable to previous studies examining the in-host population diversity of S.
Aureus during carriage (maximum pairwise of SNP distance differentiating colonies from AE cases was 26, and in controls 17) (Golubchik et al. 2013; Tong et al. 2015).

In AE cases where nasal colonization was detected, colonies derived from skin and nasal sites were interspersed throughout the phylogeny, suggesting exchange of S. aureus between sites rather than niche-specific populations (Figures 2a, S1). Comparison of the relative genetic diversities of populations from the different body sites in patient 1, revealed a significant difference in the pairwise SNP distance separating colonies from different body sites (Kruskal-Wallis rank sum test, $p=3.9e^{-06}$; Figure 2b). Nasal isolates were distinguished by the greatest mean SNP distance per colony, suggesting that the S. aureus sampled from the patient’s nose had diversified over a longer period (Figure 2b). There was a significant difference in the pairwise SNP distance observed between nasal isolates and the eczema 1 sub-sampling site isolates (Wilcoxon-Mann-Whitney test, $p=0.015$) but not compared with eczema 2 sub-site colonies (Wilcoxon-Mann-Whitney test, $p=0.16$). This demonstrates that there is genetic and spatial diversity even within a single sampled site. Conversely, the absence of diversity in S. aureus from unaffected skin suggests very recent colonization of the site or potential variable replication rates influenced by the nutrient availability in the differing colonization environment (Figure 2a). The distribution of nasal isolates throughout the phylogeny, including basally, and intermingling with eczema isolates, suggests the nasal carriage represented a more established population, and hence a potential source of S. aureus colonizing diseased skin, therefore representing self transmission.

The S. aureus from AE cases and controls exhibited overlapping distributions of relative population diversities (cases 0.1-2.6 SNPs/colony/individual; controls 0.0-3.4 SNPs/colony/individual) suggesting similar patterns of within-host diversification between
the two groups (Table S4), even though the sampling in the controls was limited to a single niche. Using previously calculated mutation rates for *S. aureus* we estimated the age of the sampled population for each participant. This suggested periods of 0 weeks to 37 months in cases, and from less than a week to 24 months in controls, for the population to have diverged from a common ancestor (Figure 3).

**Evidence of adaption in the colonizing population**

Diseased skin represents a physiologically distinct colonization environment compared to nasal epithelium, the primary carriage site of *S. aureus* in humans. We therefore hypothesized that the genomes of *S. aureus* from AE may exhibit evidence of environmental selection and adaption compared with controls. However, comparison of the distribution of mutation types between cases and controls did not show a significant difference in frequency of non-synonymous, synonymous or intergenic SNPs (Chi squared test, p=0.74) (Figure S2a). Examining the functional distribution of non-synonymous mutations revealed several gene categories differentially represented between study populations but this was also not significant (Fisher’s Exact test, p=0.27) (Figure S2b).

In the absence of signals of selection in AE cases at the cohort level, we looked for evidence of biologically and clinically relevant adaptation occurring in *S. aureus* at the patient level. Another source of variation in the *S. aureus* genome is in its content of mobile genetic elements (MGEs), which constitute the accessory genome (Lindsay and Holden 2006). Two instances of accessory genome diversification were identified in AE cases with loss/gain events of MGEs carrying genes potentially advantageous for environmental survival. Sub-sampling within a single site in patient 5 demonstrated two distinct clades, separated by 26 SNPs, defining isolates derived from opposing borders of the site (Figure 4a). The absence of
intervening variants suggested that in the sampled skin there had either been two separate acquisitions of the clade populations from a genetically closely related external population, or a single acquisition and a selective sweep removing intermediate genetic variants, establishing spatially and genetically distinct populations. Notably, a 27 kb plasmid carrying heavy metal resistance determinants was differentially distributed between the clades. In patient 1, a plasmid carrying the β-lactamase gene, *blaZ*, was present in 21 of 25 colonies (Figure 2a), thereby rendering the colonizing population differentially sensitive to penicillin. In contrast, no such examples were seen in controls, despite similarly diverse carriage populations.

The impact of antimicrobials in shaping the colonizing populations was also evident in Patient 1. All isolates from this patient contained two sequential point mutations in the gene encoding translation elongation factor G, *fusA*. The result being the amino acid change, L461K, which confers high level fusidic acid resistance (Chen et al. 2010). The clinical history in this individual confirmed the use of this agent 2 months prior to sampling, which had not improved symptoms relating to their eczema. The ancestor of this colonizing population is estimated to extensively predate this treatment (Figure 3), which would account for the lack of a therapeutic response, as colonization was by an already resistant population.

In addition to clinical interventions, we detected evidence of selection linked to survival in the host. In the *S. aureus* isolated from patient 4 two distinct mutations (Tables S5/6) ablating the expression of *agrA*, a global virulence regulator (Fowler et al. 2004; Cheung et al. 2011), were identified; 1 colony contained a nonsense mutation, and 18 colonies contained a frameshift mutation (Figure 4b). The phylogenetic context of these isolates revealed they arose independently, and during the predicted period of colonization. The convergent
evolution of mutations in this key regulator is evidence of a strong selective pressure favouring reduced virulence in this population.

**DISCUSSION**

Deep sequencing has allowed an unprecedented view of the micro-evolutionary changes occurring during carriage in AE with evidence of selection for specific strain backgrounds colonizing AE skin. In-depth analysis of within-host genetic variation revealed differences that impact both on the pathogenic potential of the bacterial cells and their response to common clinical interventions.

Healthy nasal carriage controls exhibited a similar clonal population structure indicating the observed clonal expansion in AE was not a unique facet of the disease, but rather a reflection of the natural colonization dynamics of *S. aureus*. In this regard, AE affected skin represents an additional opportunistic niche for colonization, distinguished by an apparent selection for certain *S. aureus* lineages. A previous study of *S. aureus* strains from Korean AE patients did not find prevailing genotypes (Kim et al. 2009). However the authors reported one third of the isolates belonged to CC1, the same lineage differentially prevailing in our cases. This clonal lineage has recently been reported for its prevalence in AE patients in association with Filagrin mutations (Clausen et al. 2017). Nasal carriage studies by definition investigate niche-site colonization and repeatedly show CC30 and 45 as the dominant lineages (Melles et al. 2004; Monecke et al. 2009). CC30 was the most prevalent lineage in our controls, in agreement with these previous studies and more contemporaneously by (Fleury et al. 2017). These findings raise the question of whether there are lineage specific features making them more adept at colonization of the vastly differing cutaneous environments of the nasal versus inflamed eczema skin. This emphasizes an area where future studies must be directed.
The nose is often cited as the primary reservoir of *S. aureus* colonization associated with AE (Hoeger et al. 1992; Lomholt et al. 2005). Our study established evidence of self-transmission, however in 5 of 9 cases nasal colonization was notably absent, suggesting the colonizing population was potentially derived from an extrinsic source. In these cases transmission could have arisen from a contact with a carrier, such as a family member or from the environment. Equally it is possible that the point prevalence sampling did not capture transitory nasal carriage, or that there was an un-sampled reservoir as the source of self-transmission. An extended study assessing longitudinal colonization of individuals and their close contacts is needed to elucidate the origins of the flare-associated *S. aureus* population.

In this study, 5 AE cases had carriage populations estimated to have diverged from a common ancestor more than 12 months previously (Figure 3). In 4 of these cases the predicted ancestral origins of the clonal colonizing population pre-dates the age of the child. This suggests that the origins of the diversity observed in these individual’s populations may have arisen during carriage within another host, with subsequent transmission of part of that population to the individual. Familial cross-transmission of *S. aureus* is likely to be highly relevant, and currently little is understood about this in the context of AE, although it has been shown that households are important reservoirs for transmission and diversification of *S. aureus* (Knox et al. 2015). In future, serially sampling AE patients and their family members would enable us to assess the impact that selective sweeps and transmission bottlenecks have on shaping the diversity we observe in AE colonization.

From the most mature colonization populations we uncovered clear evidence of selection shaping the populations, providing insights into mechanisms of bacterial persistence, and
other potential origins of the clonally expanded populations in these individuals. Data from two AE cases revealed bacterial genetic responses linked with intracellular survival. The homoplastic mutations of \textit{agrA} in patient 4 are strongly indicative of selection being exerted on the colonizing population. A recent \textit{in vitro} study demonstrated that \textit{agr} mutants are able to internalize and persist within keratinocytes subverting host clearance, with \textit{agr} mutants being recovered from keratinocytes at a frequency of over 58\% (Soong et al. 2015). Independent mutations in \textit{agrA} suggest that intracellular colonization and survival may have contributed to the persistence of \textit{S. aureus} in patient 4. It is worth noting that intracellular populations are less penetrable by antimicrobial therapy, and are therefore a potential cause of treatment failure. The \textit{fusA} mutation in the \textit{S. aureus} population in patient 1 who had previous exposure to fusidic acid, is associated with a small colony variant (SCV) phenotype (Norström et al. 2007). SCVs are phenotypic sub-populations occurring within a parent strain which are slow growing and show increased propensity to persist within host cells (Sendi and Proctor 2009). Intracellular \textit{S. aureus} has been reported in the context of both Darier’s disease and chronic rhinosinusitus as a cause of antimicrobial therapy failure (Eiff et al. 2001; Hayes et al. 2015). Thus we demonstrate two different mechanisms in separate cases that would support intracellular persistence and treatment resistance.

These findings highlight the therapeutic challenges of effectively eradicating \textit{S. aureus} colonization in AE. During colonization there are potentially unrecognized genetic adaptations rendering the population both insensitive and inaccessible to antimicrobial therapy, consequently prolonging pro-inflammatory interaction with the host. Longitudinal follow up of the colonizing population for assessment of diversity, evidence of adaptation and impact of therapies will advance our understanding of the relationship between \textit{S. aureus} and disease activity. This is of particular relevance in children prone to repeated infective flares.
where identifying and understanding genetic adaptation in the colonizing population may improve precision of treatment.

MATERIALS AND METHODS

Recruitment and sampling

Prospective AE case and community control sampling studies received ethical approval from the Nottingham 1 – East Midlands Research Ethics Committee (14/EM/1299) and East of Scotland Research Ethics service (15/ES/0153) respectively. Studies were conducted in accordance with the principles of the Declaration of Helsinki. Written informed parental consent as well as child assent was obtained before participation.

Cases were recruited at paediatric eczema clinics in Ninewells Hospital, Dundee, between February and October 2015. Community control samples were obtained from 306 school children in Tayside and North Fife between November 2015 and February 2016. Case study inclusion criteria were: age 0-8 years and dermatologist-diagnosed moderate to severe AE. Exclusionary criteria were: antimicrobials (systemic or topical) within the preceding 4 weeks, topical antiseptics within the preceding 2 weeks, or ultraviolet therapy (UV) within 3 months. Control study inclusion criteria: age 0-12 years. Samples obtained from control study participants with parent reported history of inflammatory skin disease, antibiotic therapy (<4 weeks previously), antiseptic therapy (<2 weeks), UV (<3 months) were not utilised as comparators in this study. Controls were selected from this study collection on the basis of age proximity to cases, no history of atopy or antimicrobial usage as per case criteria.
All participants were examined by an experienced dermatologist and AE disease severity scored using EASI (Eczema Area Severity Index) (Hanifin et al. 2001). Clinical history was obtained from cases during clinic review, and in controls via a parental questionnaire. Cases were swabbed (Transtube Amies swab, Medical Wire, England) from 5 sites including a single nostril, 2 areas of inflamed eczema, and 2 separate areas of unaffected skin. Eczema sites were sub-sampled where a swab was taken from the lateral and medial border of the site, 4 cm apart. Controls were sampled from a single nostril and antecubital fossa.

**Bacterial isolation**

Swabs were plated on Brilliance Staph 24 selective agar (Oxoid, UK) and incubated at 37°C for 24h. Selective enrichment was also undertaken with swabs being used to inoculate 3ml of Nutrient Broth with 7.5% NaCl (Oxoid, UK) grown statically at 37°C for 18h. 100µl of broth was plated on selective agar as described above. Colonies were then sub-cultured onto Brain Heart Infusion (BHI) agar (Sigma Aldrich, UK) and confirmed as *S. aureus* by PCR detection of species specific *femB* gene (Paterson et al. 2012).

**DNA extraction and whole genome sequencing**

Genomic DNA was extracted from overnight cultures of single colonies grown at 37°C using Masterpure Gram Positive DNA purification kit (Epicentre, UK) as per manufacturers protocol. DNA libraries were prepared using Nextera XT Library Preparation Kit (Illumina, UK) and quantified using Qubit High Sensitivity assay (LifeTechnologies, UK) and Agilent Bioanalyzer (Agilent, UK). Libraries were normalized, pooled and sequenced as paired-end reads on a MiSeq Genome Sequencer (Illumina, UK).
**Bioinformatic and statistical analyses**

Fastq files from MiSeq sequencing (Table S7) were assembled *de novo* with Velvet (Zerbino and Birney 2008). Multi-locus sequence types (MLST) were predicted from sequence reads using SRST2 (Inouye et al. 2014). To identify single nucleotide polymorphisms (SNPs) sequence reads were aligned to a reference genome of the same clonal complex (Table S8) using SMALT (http://www.sanger.ac.uk/science/tools/smalt-0). The default mapping parameters and SNP filtering were as previously described by (Hsu et al. 2015). Where an appropriate reference was unavailable a *de novo* assembly from the participants samples was used for mapping, typically this was the assembly with the lowest number of contigs derived from the highest number of reads. Accessory genome regions (Table S9) were identified in the reference chromosomes using ACT (Artemis Comparison Tool) (Carver et al. 2008) to compare pairwise BLASTN (Altschul et al. 1990) comparisons of reference genome sequences. Accessory regions were then masked from SNP alignments. The remaining core genome SNPs were individually curated by inspection of BAM files in Artemis to exclude false positives (Table S10), the subsequent SNPs were then used to construct maximum likelihood (ML) phylogenies with RAxML (Stamatakis 2006). Insertion/deletions (indels) were identified using GATK (https://software.broadinstitute.org/gatk/). Each participant’s reads were re-mapped to the *de novo* assembly, and GATK used to identify indels in comparison to the patient reference sequence. Indels were curated by manual inspection of BAM files in Artemis (Rutherford et al. 2000).

The relative diversity in each participant colony population was calculated by dividing the number of core genome SNPs per sequenced colony per individual. Temporal calculation for the age of the *S. aureus* populations was based on half the maximum pairwise core SNP distance and base substitution rates derived from analysis of the major *S. aureus* lineages as
described by (Uhlemann et al. 2014). A rate of $1.6 \times 10^{-6}$ SNPs/site/year was chosen as a median between published ranges. For each reference chromosome used an expected base substitution rate per month was calculated based on the size of their respective core genomes. Functional classification of genes was conducted on the MSSA476 reference genome (Holden et al. 2004), using the previously described classification scheme in Gram-positive organisms (Weinert et al. 2015). All comparative statistics were performed as two-tailed tests using R software version 3.3.1 (R Core Team 2016).

Data access

Short reads for all sequenced isolates have been submitted to the European Nucleotide Archive (ENA; http://www.ebi.ac.uk/ena/) under project accession PRJEB20148.

CONFLICT OF INTEREST

None

ACKNOWLEDGEMENTS

The authors would like to thank the children and parents who participated in this research. We also thank Maeve McAleer, Orla Fleury, Timothy Foster, Désirée E. Bennett, Joan Geoghegan, Alan Irvine for their feedback on the project and critically reviewing the manuscript; and Ed Feil for useful discussions. C.P.H. was supported by Wellcome Trust [Grant number 104241/z/14/z]. M.T.G.H, K.A.P. and K.O. were supported by the Scottish Infection Research Network and Chief Scientist Office through the Scottish Healthcare Associated Infection Prevention Institute consortium funding [CSO Reference: SIRN10]. Bioinformatics and Computational Biology analyses were supported by the University of St Andrews Bioinformatics Unit that is funded by a Wellcome Trust ISSF award [grant
097831/Z/11/Z]. J.P and M.T.G.H were supported by Wellcome Trust grant 098051. AEM is supported by Biotechnology and Biological Sciences Research Council grant BB/M014088/1. SJB is supported by a Wellcome Trust Senior Research Fellowship in Clinical Science [106865/Z/15/Z].
### TABLES

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Atopic eczema case</th>
<th>Nasal carriage control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of participants</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Age in years, mean (range)</td>
<td>1.4 (0.25-4)</td>
<td>6.6 (5-8)</td>
</tr>
<tr>
<td>Sex</td>
<td>4Males/5Females</td>
<td>9Males/9Females</td>
</tr>
<tr>
<td>EASI score, mean (range)*</td>
<td>24.4 (12.8-37)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Atopic disease**

<table>
<thead>
<tr>
<th></th>
<th>Atopic eczema</th>
<th>Nasal carriage control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atopic eczema</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Asthma</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hayfever</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Food allergy</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Other inflammatory skin disease</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: Participant characteristics of Atopic Eczema cases and healthy nasal carriage controls. N/A not applicable. * This range excludes patient 5 who had locally severe disease only.
FIGURE LEGENDS

Figure 1. Genetic diversity of *S. aureus* strains associated with AE flares. The distribution of the clonal complex (CC) structure of *S. aureus* isolates from AE cases (n=9) and nasal carriage controls (n=18). In the single case where co-colonization (patient 8) was observed, both CCs identified in the colonizing population are represented. Skin and extra-nasal carriage was found in 4 AE cases (Table S2), of which 3 were carrying CC1 isolates at both sites, and 1 was carrying CC121 isolates (Table S3). Skin only carriage isolates in AE cases (n=5) were distributed across clonal backgrounds (Tables S2/3).

Figure 2. Clonal expansion and self-transmission in Patient 1

(A) Maximum likelihood (ML) core SNPs tree illustrating genetic relationships of colonies across 3 body sites. Body diagram shows sampling site, branch label coloring corresponds to site colony was obtained from. Perforated boxes indicate transmission between body sites. * Indicates where beta-lactamase carrying plasmid is absent. Branch labels: *E*- eczema,(1/ 2 lateral/ medial), *U*- unaffected, *N*- nose, *C*- colony number from site, *ic*-18 (pre- or post-enrichment colony). Tree rooted using ST188 reference. Scale bar is approximately 1 SNP (not applicable to root branches with strikethrough). (B) Box plot comparing pairwise SNP difference between sequenced colonies from different body sites in a single patient.

Figure 3. Estimated age of the colonizing populations

Time point to the most recent common ancestor of the population is presented in months and is calculated from half of the maximum pairwise SNP distance observed in the sampled populations. Both colony populations within patient 8 are included.
Figure 4. Evidence of selection and genetic adaptation within the host

(a) Evidence of plasmid dynamics and an evolutionary bottleneck. ML core SNPs tree illustrating genetic relationship of colonies from two spatially distinct positions in single eczema site. Branch label marked with a red arrow indicates point in phylogeny where a plasmid (backbone similar to SAP0194) carrying multiple metal resistance genes was gained in the *S. aureus* population. Tree rooted using MRSA252 reference. Branch labels: E1/3-lateral border, E2/4-medial border of single eczema site. SNP bar is indicated for scale (not applicable to root branches with strikethrough). (b) Convergent evolution of *agrA* mutants. ML core SNP tree of colonies from 5 body sites. Branch label marked with red arrows indicate the point in phylogeny where homoplastic mutations in *agrA* are predicted to have occurred. Tree rooted using MSSA476 reference.
REFERENCES


Patient 5
EASI score: locally severe only
Number of colonies: 10
Colonizing strain: ST30

Patient 4
EASI score: 25
Number of colonies: 28
Colonizing strain: ST1

agrA nonsense mutation
agrA truncation