Predictors of a superficial first basal cell carcinoma

J.A.C. Verkouteren

L.M. Pardo

A.G. Uitterlinden

T. Nijsten

Erasmus University Rotterdam

ABSTRACT

Background: several observational studies have suggested differences in the risk factor profile between patients with superficial basal cell carcinomas (BCCs) and non-superficial BCCs.

Objective: to test the reproducibility of previous study findings and to find new genetic and non-genetic predictors for patients with a superficial first BCC.

Methods: 14.628 participants of northwestern European descent aged 45 years or older from a prospective population-based cohort study (Rotterdam Study) were linked with the Dutch Pathology Registry (PALGA) of whom 1,528 were identified as BCC patients. After exclusion, 948 eligible BCC patients remained for further non-genetic analyses and 1,014 for genetic analyses. We included 11 phenotypic, environmental and tumor-specific characteristics, and 20 candidate single nucleotide polymorphisms (SNP) as potential predictors for patients with a superficial first BCC. We performed binary logistic multivariable regression analyses.

Results: we found that patients with a superficial first BCC were significantly younger, almost two times more often female and 12-18 times more likely to have their BCC on the trunk or extremities than patients with a non-superficial first BCC. One SNP (rs12203592), mapped to IRF4, looked promising (OR 1.83, 95% CI 1.13-2.97, p-value <0.05), but after adjustment for multiple testing, no significant differences in genetic make-up between superficial BCC and non-superficial BCC patients were found.

Conclusion: we conclude that patients with a superficial BCC differ from non-superficial BCC patients with respect to environmental factors (tumor localization as a proxy for UVR exposure) and phenotypic characteristics (age and sex), but we found no difference in genotype. As superficial BCC patients develop their first BCCs at a younger age, they could be at higher life-time risk for subsequent skin cancers and therefore be an interesting group for secondary prevention.

INTRODUCTION

Patients with basal cell carcinoma (BCC) put a strain on health care services worldwide, as a result of the high and increasing BCC incidence, especially in young white-skinned women, and the increased risk of synchronous and metachronous BCCs and other ultraviolet radiation (UVR) related skin cancers (i.e., field cancerization).[1-3] In addition, the disability adjusted life years and health care costs for BCC have risen significantly as well.[4,5]

There are different histopathological subtypes of BCC, based on the growth pattern(s) found within the tumor tissue. The nodular pattern is most frequently found histological subtype (>50%), followed by superficial (~20%) and infiltrative (~10%) and about 20% of the tumors show a mixed type.[6-11] The frequencies reported depend on the used pathological classification system and period of the study, because classification systems and subtype incidences changed over time.[12,13] BCCs mostly occur on the head and neck area (i.e., chronically sun exposed; >70%), followed by the trunk (~20%) and extremities (~10%), which are both areas intermittently exposed to UVR.[7-10,14] Several observational studies have identified associations between age, sex and anatomical site, and BCC subtypes.[6-9,13] Patients with a superficial BCC have more often their BCC on the trunk and extremities than in the head and neck region,[6-9,13] are younger[7-9,13] and more often female.[8,9] In addition, patients with an initial truncal superficial BCC developed metachronous BCCs at a faster rate than patients with other anatomical site and histology combinations.[15]

These results could indicate that different BCC subtypes, in particular superficial, have other etiologies with respect to environmental factors (e.g., UVR exposure), phenotypic characteristics (e.g., age and sex) and genetic predisposition. However, only a few studies have studied other predictors than age, sex and anatomical site, with conflicting results.[10,16,17]

The objective of this study is to test the reproducibility of these findings and to find potentially new predictors for patients with a superficial first BCC (sBCC). We hereto analyzed the data of almost 1,000 white-skinned participants with a BCC of a prospective population-based cohort study (Rotterdam Study).

MATERIALS AND METHODS

Study population

The Rotterdam Study is a prospective population-based cohort study of 14,926 participants (divided over three cohorts) aged 45 years or older, living in a well-defined suburb of Rotterdam, the Netherlands.[18] The cohorts predominantly consist of people

of northwestern European descent. All the participants were interviewed and examined at baseline and these examinations were repeated about every 4 years. The Rotterdam Study has been approved by the Medical Ethics Committee of the Erasmus MC and by the Ministry of Health, Welfare and Sport of the Netherlands , implementing the Wet Bevolkingsonderzoek: ERGO (Population Studies Act: Rotterdam Study) and it was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent to participate in the study and to obtain information from their treating physicians.

Phenotype / Case definition

The method by which we identified BCCs has been described in detail previously.[19] In short, the study database was linked to the Dutch nationwide network and registry of histopathology and cytopathology (PALGA) to retrieve medical history of all participants on histopathologically confirmed BCCs between 1 July 1989 and 31 December 2013.[20] Of the 14,926 RS participants, 298 did not sign informed consent for a linkage and could not be linked to PALGA. The pathology excerpts we received, contained information on date of diagnosis, anatomical location, body side, type of procedure (i.e., biopsy or excision), radicality, and diagnosis. The majority of these excerpts showed a subtyping of the BCC and these subtypes were coded based on the World Health Organization's histological classification of keratinocytic skin tumors.[21] If there was a subtype discrepancy between a biopsy and an excision or a biopsy/excision included more than one subtype, we coded it as a mixed type BCC and noted the concerned subtypes. Patients with a missing subtype were excluded and patients with a mixed type first BCC with a superficial component were excluded as well, because it was unclear to which subtype these belong (i.e., superficial or non-superficial). Metachronous BCCs that occurred within 6 months of the first BCC were counted as additional tumors at the date of the initial diagnosis, as those BCCs were most likely present at this earlier date. We randomly selected a BCC for participants with synchronous BCCs on their first diagnosis date.

Selection of non-genetic candidate predictors

A literature search up to May 2016 for English publications on phenotypic, environmental and tumor specific factors previously involved in BCC subtypes was done in PubMed. Four phenotypic factors were included, namely age at first BCC, sex, pigment status and tendency to develop sunburn.[7-10,13] The latter was a combination of eye color and hair color when young (e.g., a participant with blue eyes and red hair was scored as light). Five environmental characteristics were chosen and concerned a history of being outdoor for over 4 hours per day during more than 25 years, sun protective behavior measured by wearing sunglasses or a hat, smoking, alcohol consumption and

coffee consumption.[10,22,23] Finally, two tumor related variables were included, namely localization of the first BCC and the number of BCCs at first date of diagnosis.[6-9,13,15] All selected variables (except tumor-specific characteristics) were measured at study entry or at a study visit closest to study entry.

Selection of candidate single nucleotide polymorphisms

A literature search up to May 2016 for English genome-wide association study (GWAS) publications of loci that confer risk of BCC or non-melanoma skin cancer was done in PubMed. There was no GWAS of the histopathological subtypes of BCC. To reduce the burden of multiple testing, all selected single nucleotide polymorphisms (SNPs) had to be at least borderline genome-wide significant (p-value $< 7.0 \times 10^{-8}$) and had to be replicated in another cohort. This resulted in a list of 20 candidate SNPs located in 17 different chromosomal regions (eTable 1).[24]

Genotype

DNA was isolated from whole blood, further processed and quality checked following standard protocols.[18] The Illumina Infinium II HumanHap550 BeadChips and the Illumina Human610-Quad BeadChips were used to genotype the RS participants.

Quality control criteria included removing SNPs with Hardy-Weinberg equilibrium deviations (p-value < 0.0001), genotyping call rate $< 97\%$, gender mismatch and a high mean autosomal heterozygosity. SNPs were not included if they had a minor allele frequency of less than 1% and/or an imputation r^2 of less than 0.3.

For the candidate SNP approach we used genotypes that were estimated from the imputed 1000Genomes, GIANT Phase I version 3 dosage data[18] using the Genome-wide Complex Trait Analysis (GCTA) software with default parameters.[25] All selected candidate SNPs were included in our genetic database.

Statistical analysis

Non-genetic binary logistic regression analysis of sBCC vs non-superficial BCC (nsBCC)
All the assumptions of a binary logistic regression analysis were tested and we found no violations. There existed no strong (multi)collinearity between the selected non-genetic candidate predictors. A few outliers in the coffee consumption and alcohol consumption variables were found using the outlier labeling rule,[26] but all values were realistic. There was sufficient power to include the 11 selected candidate predictors in the multivariable binary logistic regression analysis.

We could safely assume that missing predictor values were missing at random (i.e., missing data points were not related to the missing data itself, but to the observed

data). Missing predictor values could therefore be imputed using multiple imputation (30 times) by an iterative Markov Chain Monte Carlo method. The imputation model included all candidate predictors, the outcome, the body mass index (kg/m^2), the level of education, the side of the first BCC and the Rotterdam Study cohort number. After the imputations we did both univariable and multivariable binary logistic regression analyses. No selection methods were used for the multivariable analysis.

All of the data management and the non-genetic binary logistic regression analyses were done in IBM® SPSS® Statistics for Windows version 21 (Chicago, IL).

Genetic (SNP-based) binary logistic regression analysis of sBCC vs nsBCC
All the assumptions of a binary logistic regression analysis were tested and we found one violation, namely collinearity between two selected candidate SNPs. A bivariate correlation matrix showed a Pearson correlation coefficient of 0.86 between rs12210050 and rs12202284, which means that these predictors were highly correlated. A few outliers in the age and principal component variables were found using the outlier labeling rule,[26] but all values were realistic. There was insufficient power to include the 20 selected candidate SNPs, age, sex and four principal components (PCs) in the multivariable binary logistic regression analysis. Therefore, we adjusted our analyses for multiple testing using the false discovery rate (FDR).[27] PCs were included to adjust for possible population stratification.

The SNP-based association analyses were performed on the imputed dosage data using a binary logistic regression with an additive model. The multivariable logistic regression analysis was adjusted for age at BCC diagnosis, sex and four PCs. No selection methods were used for the multivariable analysis.

The genetic data were prepared on our genetic servers and IBM® SPSS® Statistics for Windows version 21 (Chicago, IL) was used for the analyses.

Sensitivity analyses of sBCC vs. nodular BCC
We performed sensitivity analyses by doing the same non-genetic and genetic regression analyses as for sBCC versus nsBCC, but now including only patients with superficial or nodular first BCC.

RESULTS

Study population for non-genetic analyses
Of the 14,628 RS participants linked to PALGA, 1,528 had at least one BCC. After the exclusion of patients with a missing subtype (n = 71), patients with a mixed superficial first BCC (n = 58) and patients who developed at least one BCC before study entry (n =

451), 948 eligible BCC patients remained for further analyses. We randomly selected a BCC for participants with synchronous BCCs on their first diagnosis date (n = 125). Of the included patients, 137 (14%) had a superficial first BCC, 496 (52%) a nodular first BCC and the remaining 315 (33%) another subtype (infiltrative, micronodular or non-superficial mixed type; Table 1 and eTable 2).

Patients with a superficial first BCC were younger than patients with a non-superficial first BCC (median age 70.2 vs 75.5 years) and the proportion females (64%) was higher in sBCC patients than in nsBCC patients (54%; Table 1). Approximately 4 out of 5

Table 1. Non-genetic characteristics of 948 Rotterdam Study patients with a first BCC

Patient and tumor characteristics	Coding	Overall[1]	Superficial BCC	Non-superficial BCC
Number of patients		948 (100%)	137 (100%)	811 (100%)
Age at first BCC (years)	Median (IQR)	74.6 (67.9-81.2)	70.2 (64.3-76.0)	75.5 (68.9-81.8)
Sex	Female	526 (55%)	87 (64%)	439 (54%)
Pigment status	Dark	153 (16%)	24 (18%)	129 (16%)
	Intermediate	447 (47%)	70 (51%)	377 (46%)
	Light	213 (22%)	30 (22%)	183 (23%)
	Missing	135 (14%)	13 (9%)	122 (15%)
Easily sunburned	Yes	319 (34%)	53 (39%)	266 (33%)
	Missing	65 (7%)	5 (4%)	60 (7%)
Outdoor work	Yes	124 (13%)	14 (10%)	110 (14%)
	Missing	274 (29%)	42 (31%)	232 (29%)
Sun protection	No, never or hardly ever	357 (38%)	44 (32%)	313 (39%)
	Missing	60 (6%)	4 (3%)	56 (7%)
Smoking	Current or former	623 (66%)	92 (67%)	531 (65%)
	Missing	17 (2%)	1 (1%)	16 (2%)
Alcohol consumption (glasses/day)	Median (IQR)	0.6 (0.1-1.7)	0.6 (0.1-1.4)	0.6 (0.1-1.8)
	Missing	215 (23%)	18 (13%)	197 (24%)
Coffee consumption (cups/day)	Median (IQR)	3.3 (2.0-4.5)	3.3 (1.5-4.0)	4.0 (2.0-5.0)
	Missing	215 (23%)	18 (13%)	197 (24%)
>1 BCC at initial diagnosis	Yes	125 (13%)	24 (18%)	101 (12%)
Localization of first BCC	Head and neck	630 (66%)	24 (18%)	606 (75%)
	Extremities	128 (14%)	54 (39%)	74 (9%)
	Trunk	184 (19%)	58 (42%)	126 (16%)
	Missing	6 (1%)	1 (1%)	5 (1%)

[1] Participants with a mixed-type BCC with a superficial component were excluded.
BCC, basal cell carcinoma; IQR, interquartile range.

Erasmus University Rotterdam

sBCCs were located on the extremities (39%) or trunk (42%) as opposed to 1 in 4 of the nsBCCs.

Non-genetic binary logistic regression analyses of sBCC vs. nsBCC

Of the 11 candidate predictors, 3 were significantly associated with a superficial first BCC in the univariable binary logistic regression analyses, namely a younger age at first BCC diagnosis (OR: 0.94, 95% CI: 0.92-0.96 per year), female gender (OR: 1.47, 95% CI: 1.01-2.14) and localization on the trunk (OR: 11.44, 95% CI: 6.85-19.10) or extremities (OR: 18.07, 95% CI: 10.56-30.93; Table 3).

These associations remained strongly significant after the multivariable binary logistic regression analysis and no other predictors became significant (Table 3). Female gender gave an even stronger risk increase for sBCC (OR: 1.88, 95% CI: 1.16-3.03, p-value < 0.05), but localization remained the strongest predictor (truncal OR: 12.20, 95% CI: 7.08-21.03, p-value < 0.001; extremities OR: 17.57, 95% CI: 10.06-30.70, p-value < 0.001). The 11 predictors together explained 19.7% (Cox and Snell R^2) of total variability of a superficial first BCC compared to a non-superficial first BCC.

Study population for genetic analyses

Of the 14,628 RS participants linked to PALGA, 1,257 were genotyped and had at least one BCC. After the exclusion of patients with a missing subtype (n = 181) and patients with a mixed superficial first BCC (n = 62), 1,014 eligible BCC patients remained for further analyses. We randomly selected a BCC for participants with synchronous BCCs on their first diagnosis date (n = 126). Of the included patients, 159 (16%) had a superficial first BCC, 522 (51%) a nodular first BCC and the remaining 333 (33%) another subtype (infiltrative, micronodular or non-superficial mixed type; Table 2 and eTable 3).

Patients with a superficial first BCC were younger than patients with a non-superficial first BCC (median age 68.0 vs 73.5 years) and the proportion females (65%) was higher in sBCC patients than in nsBCC patients (53%).

Table 2. Genetic characteristics of 1,014 Rotterdam Study patients with a first BCC

Patient and tumor characteristics	Coding	Overall[1]	Superficial BCC	Non-superficial BCC
Number of patients		1,014 (100%)	159 (100%)	855 (100%)
Age at first BCC (years)	Median (IQR)	72.9 (64.4-79.8)	68.0 (60.8-75.6)	73.5 (65.5-80.5)
Sex	Female	556 (55%)	103 (65%)	453 (53%)

[1] Participants with a mixed-type BCC with a superficial component were excluded.
BCC, basal cell carcinoma; IQR, interquartile range.

Table 3. Associations between predictors and occurrence of superficial first BCC (n = 948)[1]

Patient and tumor characteristics	Coding	Univariable models[2]	Multivariable model[2,3]
Age at first BCC (years)	Continuous	0.94 (0.92-0.96)***	0.95 (0.93-0.98)***
Sex	Female	1.47 (1.01-2.14)*	1.88 (1.16-3.03)*
Pigment status	Dark	Reference	Reference
	Intermediate	1.01 (0.61-1.67)	0.91 (0.50-1.64)
	Light	0.92 (0.51-1.65)	0.80 (0.40-1.61)
Easily sunburned	Yes	1.22 (0.83-1.78)	1.13 (0.70-1.81)
Outdoor work	Yes	0.77 (0.42-1.38)	0.85 (0.43-1.69)
Sun protection	No or hardly ever	0.70 (0.48-1.04)	0.80 (0.51-1.26)
Smoking	Current or former	1.03 (0.70-1.52)	1.41 (0.85-2.33)
Alcohol consumption (glasses/day)	Continuous	0.90 (0.77-1.05)	0.84 (0.70-1.01)
Coffee consumption (cups/day)	Continuous	0.92 (0.82-1.02)	0.90 (0.79-1.02)
>1 BCC at initial diagnosis	Yes	1.49 (0.92-2.43)	1.41 (0.79-2.52)
Localization of first BCC	Head and neck	Reference	Reference
	Extremities	18.07 (10.56-30.93)***	17.57 (10.06-30.70)***
	Trunk	11.44 (6.85-19.10)***	12.20 (7.08-21.03)***

[1] Compared to nodular, micronodular, infiltrative and mixed-type BCCs; all mixed-type BCCs with a superficial component were excluded.
[2] Pooled ORs with 95% CIs between parentheses.
[3] Full model, no selection procedures used.
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001.
BCC, basal cell carcinoma.

Genetic (SNP-based) binary logistic regression analyses of sBCC vs. nsBCC

Of the 20 candidate SNPs, 2 were borderline significantly associated with a first sBCC in the univariable SNP-based binary logistic regression analyses, namely rs8015138 (OR: 0.76, 95% CI: 0.60-0.97) and rs12203592 (OR: 1.55, 95% CI: 1.01-2.37; Table 4).

Before the multivariable SNP-based binary logistic regression analyses, we excluded rs12210050 because it was highly correlated (Pearson's r: 0.86) with rs12202284 and both SNPs were also in strong linkage disequilibrium (r^2: 0.73) with each other. The multivariable analysis resulted in 1 promising SNP, namely rs12203592 (OR: 1.83, 95% CI: 1.13-2.97, p-value 0.014) mapped to pigmentation gene IRF4, but after adjustment for multiple testing (FDR) this SNP lost its significance as well. No other SNPs were significantly associated to sBCC (Table 4).

The 19 candidate SNPs together explained 1.6%, of which rs12203592 explained 0.4% (Cox and Snell R^2), of the total variability of a superficial first BCC compared to a non-superficial first BCC.

Erasmus University Rotterdam

Table 4. Associations between predictors and occurrence of superficial first BCC (n = 1,014)[1]

Patient and tumor characteristics	Coding	Univariable models[2]	Multivariable models[2,3]	Multivariable model[2,4]
Age at first BCC (years)	Continuous	0.97 (0.95-0.98)***		0.96 (0.95-0.98)***
Sex	Female	1.63 (1.15-2.32)**		1.67 (1.16-2.40)**
rs73635312	Yes	1.11 (0.74-1.67)	1.08 (0.71-1.64)	1.07 (0.70-1.63)
rs11170164	Yes	0.88 (0.57-1.35)	0.93 (0.59-1.45)	0.95 (0.61-1.49)
rs7335046	Yes	0.95 (0.67-1.36)	0.92 (0.64-1.32)	0.89 (0.61-1.29)
rs8015138	Yes	0.76 (0.60-0.97)*	0.78 (0.61-1.00)	0.79 (0.62-1.02)
rs1805007	Yes	0.96 (0.60-1.55)	0.96 (0.59-1.57)	0.95 (0.58-1.56)
rs78378222	Yes	0.89 (0.38-2.10)	0.95 (0.40-2.52)	0.98 (0.41-2.36)
rs7538876	Yes	1.06 (0.82-1.36)	1.02 (0.79-1.32)	1.02 (0.78-1.32)
rs801114	Yes	0.96 (0.75-1.23)	1.03 (0.80-1.33)	1.00 (0.77-1.30)
rs214782	Yes	0.89 (0.66-1.19)	0.91 (0.68-1.22)	0.93 (0.69-1.26)
rs13014235	Yes	1.07 (0.84-1.36)	1.07 (0.83-1.37)	1.07 (0.83-1.38)
rs57244888	Yes	1.06 (0.67-1.67)	1.11 (0.70-1.78)	1.12 (0.70-1.80)
rs401681	Yes	1.05 (0.82-1.33)	1.00 (0.78-1.28)	0.98 (0.76-1.26)
rs12203592	Yes	1.55 (1.01-2.37)*	1.55 (1.01-2.39)*	1.83 (1.13-2.97)*
rs12202284	Yes	0.92 (0.64-1.32)	0.92 (0.63-1.33)	0.72 (0.47-1.08)
rs12210050	Yes	1.00 (0.71-1.40)	1.03 (0.73-1.46)	
rs157935	Yes	0.95 (0.73-1.24)	0.94 (0.72-1.23)	0.95 (0.72-1.26)
rs28727938	Yes	0.89 (0.52-1.53)	0.78 (0.45-1.34)	0.75 (0.43-1.32)
rs7006527	Yes	0.81 (0.58-1.13)	0.79 (0.56-1.12)	0.79 (0.55-1.12)
rs2151280	Yes	1.09 (0.86-1.38)	1.13 (0.89-1.44)	1.14 (0.89-1.46)
rs59586681	Yes	1.09 (0.85-1.40)	1.13 (0.88-1.46)	1.10 (0.85-1.44)

[1] Compared to nodular, micronodular, infiltrative and mixed-type BCCs; all mixed-type BCCs with a superficial component were excluded.
[2] Odds ratios with 95% confidence intervals between parentheses.
[3] Included one SNP at a time, adjusted for age at first BCC, sex and first 4 principal components.
[4] Full model, adjusted for age at first BCC, sex and first 4 principal components. No selection procedures used.
* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001.
BCC, basal cell carcinoma.

Sensitivity analyses of sBCC vs. nodular BCC

After the non-genetic multivariable binary logistic regression analysis comparing sBCC to nodular BCC, the same predictors (age at first BCC diagnosis, sex and a localization on the trunk or extremities) were significantly associated with a superficial first BCC with similar effect sizes (eTable 4). The explained variability increased by 4.3% to 25.0% (Cox and Snell R^2).

The multivariable SNP-based binary logistic regression analysis comparing sBCC to nodular BCC resulted in 2 promising SNPs, namely rs12203592 (OR: 2.11, 95%

CI: 1.25-3.58, p-value 0.005) and rs12202284 (OR: 0.55, 95% CI: 0.35-0.88, p-value 0.012), both mapped to the IRF4 – EXOC2 region, but were not in strong linkage disequilibrium (r^2: 0.18) with each other (eTable 5). However, after adjustment for multiple testing (FDR) both SNPs lost their significance.

DISCUSSION

This prospective population-based cohort study replicates some previous non-genetic findings and shows that there are significant differences between patients with a superficial first BCC and a non-superficial first BCC. Patients who presented with a sBCC were younger, more often female and had their BCCs more frequently on the extremities and trunk than patients with nsBCCs. This study also looked into potential genetic differences. One SNP, mapped to IRF4, looked promising, but after adjustment for multiple testing, no significant differences in genetic make-up between sBCC and nsBCC patients were found.

The associations found between the non-genetic predictors and the occurrence of a superficial first BCC were in line with several other older and more recent observational studies from Europe and Australia.[6-10,13] However, most of these non-genetic studies on histopathological BCC subtypes did not adjust for potential confounders.[6-9,13] Therefore, it is possible that the associations found, were spurious. We included 11 potential confounders in our non-genetic multivariable model and found that patients with a superficial first BCC were significantly younger, almost twice as likely to be female and 12-18 times more likely to have their BCC on the trunk or extremities than patients with a non-superficial first BCC. These differences in age, sex and localization could suggest that a different pattern of UVR exposure, namely intense intermittent, plays a role in the etiology of sBCC as compared to nsBCC. A British and Australian cohort study showed that excessive recreational UVR exposure significantly increased the risk of truncal (superficial) BCCs,[17,28] whereas Dutch and Italian case-control studies showed no relation between cumulative lifetime UVR exposure and sBCC.[10,16]

Another potential explanation for the significantly higher risk of sBCC in younger women could be behavior. Women tend to use tanning beds more often than men[29,30] and pay closer attention the their health and physical appearance than men, which may lead to more medical visits.[31]

It is also possible that tumor biology differs at various anatomical sites. A Dutch renal transplant study showed that transplant recipients more often developed sBCCs and that their BCCs were located more frequently on the trunk and extremities than in the non-immunosuppressed, which may point at role for the immune system.[8]

Superficial first BCC patients were significantly younger (approximately 5 years) than non-superficial first BCC patients and developed their BCCs more often on relatively sun-unexposed sites, which could mean that they have a different genetic predisposition which makes them more vulnerable to develop (superficial) BCC. It is possible that they, for example, have a reduced DNA repair capacity or other risk increasing DNA differences.[32] Hence, we compared carefully selected BCC candidate SNPs between these two patient groups. Of the 19 included candidate SNPs in the multivariable regression analysis, rs12203592 looked most promising (OR: 1.83, 95% CI: 1.13-2.97, p-value 0.014), but lost its significance after adjusting the FDR. This SNP is an intron variant mapped to the interferon regulatory factor 4 (IRF4) gene, which belongs to a well-known family of transcription factors that are important in the regulation of the immune system. It is possible that certain SNPs downregulate the immune system which could lead to the formation of sBCC in relatively sun-unexposed areas earlier in life. A recent genetic analysis of melanoma patients showed a significant association with the bimodal (early- and late-onset) age distribution of melanoma for different rs12203592 genotypes.[33] In addition, IRF4 also plays a key role in the pigmentation pathway and in the formation of (pre)malignancies of the skin.[34-36] These pre-malignancies (i.e., actinic keratosis) have a superficial growth pattern which is comparable to that of sBCCs.

Limitations

Misclassification of BCC subtypes by pathologists most likely occurred throughout the study period, but it is unlikely that this misclassification was differential. However, we could not check the tissue samples as we only received excerpts from PALGA.

The total number of BCCs could have been underestimated, since we only included histopathologically confirmed BCCs. This underestimation will be most pronounced for superficial BCCs, because physicians could diagnose these BCCs visually and treat them non-invasively. However, a recent Dutch observational study showed that only a small percentage (ca. 7%) of patients with metachronous BCCs had subsequent non-histologically confirmed BCCs.[37] In addition, the evidence-based BCC guideline from the Dutch Society for Dermatology and Venereology states that histopathological verification is needed for all for BCC suspicious lesions.[38] Finally, the distribution pattern of the histopathological subtypes in our study population is in line with other studies, with the nodular type being the most common, followed by the superficial type and infiltrative type, while mixed types were frequently found as well.[6-11] Our candidate SNP approach likely lacked sufficient power (26 degrees of freedom used and 159 patients with a superficial first BCC) despite the FDR approach taken. Detailed information about other limitations of the Rotterdam Study, the phenotype collection and the non-genetic and genetic predictors can be found in two earlier publications.[19,39]

Conclusion

Patients with a superficial first BCC differ from non-superficial first BCC patients with respect to environmental factors (tumor localization as a proxy for UVR exposure) and phenotypic characteristics (age and sex), but (as far as we could find) not in genotype. As sBCC patients develop their first BCCs at a younger age, they could be at higher risk for subsequent skin cancers. Further study of the interplay between environmental, phenotypic and genotypic predictors and BCC subtypes may provide useful knowledge for BCC pathogenesis and the design of programs for prevention and early detection of BCC.

ACKNOWLEDGMENTS

REFERENCES

1   Lomas A, Leonardi-Bee J, Bath-Hextall F. A systematic review of worldwide incidence of nonmelanoma skin cancer. *Br J Dermatol* 2012; **166**: 1069-80.

2   Christenson LJ, Borrowman TA, Vachon CM *et al.* Incidence of basal cell and squamous cell carcinomas in a population younger than 40 years. *JAMA* 2005; **294**: 681-90.

3   Flohil SC, van der Leest RJ, Arends LR *et al.* Risk of subsequent cutaneous malignancy in patients with prior keratinocyte carcinoma: a systematic review and meta-analysis. *Eur J Cancer* 2013; **49**: 2365-75.

4   Hollestein LM, de Vries E, Aarts MJ *et al.* Burden of disease caused by keratinocyte cancer has increased in The Netherlands since 1989. *J Am Acad Dermatol* 2014; **71**: 896-903.

5   Gordon LG, Rowell D. Health system costs of skin cancer and cost-effectiveness of skin cancer prevention and screening: a systematic review. *Eur J Cancer Prev* 2015; **24**: 141-9.

6   Betti R, Inselvini E, Carducci M *et al.* Age and site prevalence of histologic subtypes of basal cell carcinomas. *Int J Dermatol* 1995; **34**: 174-6.

7   McCormack CJ, Kelly JW, Dorevitch AP. Differences in age and body site distribution of the histological subtypes of basal cell carcinoma. A possible indicator of differing causes. *Arch Dermatol* 1997; **133**: 593-6.

8   Bastiaens MT, Hoefnagel JJ, Bruijn JA *et al.* Differences in age, site distribution, and sex between nodular and superficial basal cell carcinoma indicate different types of tumors. *J Invest Dermatol* 1998; **110**: 880-4.

9   Scrivener Y, Grosshans E, Cribier B. Variations of basal cell carcinomas according to gender, age, location and histopathological subtype. *Br J Dermatol* 2002; **147**: 41-7.

10   Pelucchi C, Di Landro A, Naldi L *et al.* Risk factors for histological types and anatomic sites of cutaneous basal-cell carcinoma: an italian case-control study. *J Invest Dermatol* 2007; **127**: 935-44.

11   Betti R, Radaelli G, Crosti C *et al.* Margin involvement and clinical pattern of basal cell carcinoma with mixed histology. *J Eur Acad Dermatol Venereol* 2012; **26**: 483-7.

12   Rippey JJ. Why classify basal cell carcinomas? *Histopathology* 1998; **32**: 393-8.

13   Arits AH, Schlangen MH, Nelemans PJ *et al.* Trends in the incidence of basal cell carcinoma by histopathological subtype. *J Eur Acad Dermatol Venereol* 2011; **25**: 565-9.

14   Kopf AW. Computer analysis of 3531 basal-cell carcinomas of the skin. *J Dermatol* 1979; **6**: 267-81.

15   Lovatt TJ, Lear JT, Bastrilles J *et al.* Associations between ultraviolet radiation, basal cell carcinoma site and histology, host characteristics, and rate of development of further tumors. *J Am Acad Dermatol* 2005; **52**: 468-73.

16   Kennedy C, Bajdik CD, Willemze R *et al.* The influence of painful sunburns and lifetime sun exposure on the risk of actinic keratoses, seborrheic warts, melanocytic nevi, atypical nevi, and skin cancer. *J Invest Dermatol* 2003; **120**: 1087-93.

17   Lovatt TJ, Lear JT, Bastrilles J *et al.* Associations between UVR exposure and basal cell carcinoma site and histology. *Cancer Lett* 2004; **216**: 191-7.

18   Hofman A, Brusselle GG, Darwish Murad S *et al.* The Rotterdam Study: 2016 objectives and design update. *Eur J Epidemiol* 2015; **30**: 661-708.

19   Verkouteren JA, Smedinga H, Steyerberg EW *et al.* Predicting the Risk of a Second Basal Cell Carcinoma. *J Invest Dermatol* 2015; **135**: 2649-56.

20 Casparie M, Tiebosch AT, Burger G *et al*. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* 2007; **29**: 19-24.

21 LeBoit PE, Burg G, Weedon D *et al*. *World Health Organization Classification of Tumours. Pathology and Genetics of Skin Tumours.* Lyon: IARC Press. 2006.

22 De Hertog SA, Wensveen CA, Bastiaens MT *et al*. Relation between smoking and skin cancer. *J Clin Oncol* 2001; **19**: 231-8.

23 Husein-Elahmed H, Aneiros-Fernandez J, Gutierrez-Salmeron MT *et al*. Alcohol intake and risk of aggressive histological basal cell carcinoma: a case-control study. *Eur J Dermatol* 2012; **22**: 525-30.

24 http://www.ebi.ac.uk/gwas/. In, Vol. 2015: National Human Genome Research Institute (NHGRI) and the European Bioinformatics Institute (EMBL-EBI). 16th December 2015.

25 Yang J, Lee SH, Goddard ME *et al*. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**: 76-82.

26 Hoaglin DC, Iglewicz B. Fine-Tuning Some Resistant Rules for Outlier Labeling. *J Am Stat Assoc* 1987; **82**: 1147-9.

27 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* 1995; **57**: 289-300.

28 Neale RE, Davis M, Pandeya N *et al*. Basal cell carcinoma on the trunk is associated with excessive sun exposure. *J Am Acad Dermatol* 2007; **56**: 380-6.

29 Robinson JK, Rigel DS, Amonette RA. Trends in sun exposure knowledge, attitudes, and behaviors: 1986 to 1996. *J Am Acad Dermatol* 1997; **37**: 179-86.

30 Koster B, Thorgaard C, Clemmensen IH *et al*. Sunbed use in the Danish population in 2007: a cross-sectional study. *Prev Med* 2009; **48**: 288-90.

31 Swetter SM, Layton CJ, Johnson TM *et al*. Gender differences in melanoma awareness and detection practices between middle-aged and older men with melanoma and their female spouses. *Arch Dermatol* 2009; **145**: 488-90.

32 Wei Q, Matanoski GM, Farmer ER *et al*. DNA repair and aging in basal cell carcinoma: a molecular epidemiology study. *Proc Natl Acad Sci U S A* 1993; **90**: 1614-8.

33 Gibbs DC, Orlow I, Bramson JI *et al*. Association of Interferon Regulatory Factor-4 Polymorphism rs12203592 With Divergent Melanoma Pathways. *J Natl Cancer Inst* 2016; **108**.

34 Han J, Kraft P, Nan H *et al*. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 2008; **4**: e1000074.

35 Gerstenblith MR, Shi J, Landi MT. Genome-wide association studies of pigmentation and skin cancer: a review and meta-analysis. *Pigment Cell Melanoma Res* 2010; **23**: 587-606.

36 Jacobs LC, Liu F, Pardo LM *et al*. IRF4, MC1R and TYR genes are risk factors for actinic keratosis independent of skin color. *Hum Mol Genet* 2015; **24**: 3296-303.

37 Flohil SC, van Tiel S, Koljenovic S *et al*. Frequency of non-histologically diagnosed basal cell carcinomas in daily Dutch practice. *J Eur Acad Dermatol Venereol* 2013; **27**: 907-11.

38 (NVDV) NVvD. Evidence-based Guideline Basal Cell Carcinoma. In. 2015.

39 Verkouteren JA, Pardo LM, Uitterlinden AG *et al*. Common Variants Affecting Susceptibility to Develop Multiple Basal Cell Carcinomas. *J Invest Dermatol* 2015; **135**: 2135-8.

SUPPLEMENTARY MATERIAL

eTable 1. Candidate single nucleotide polymorphisms for basal cell carcinoma or non-melanoma skin cancer

First Author	Date	Disease/Trait	Region	Reported Gene(s)	Mapped gene	Strongest SNP-Risk Allele	Context	Risk Allele Frequency	p-Value	OR or beta	95% CI
Stacey SN	11-10-2008	Basal cell carcinoma	1p36.13	RCC2, ARHGEF10L, PADI4, PADI6	PADI6	rs7538876-A	intron	0,35	4,00E-12	1,28	[1.19-1.37]
Stacey SN	11-10-2008	Basal cell carcinoma	1q42.13	RHOU	FTH1P2 - ISCA1P2	rs801114-G		0,33	6,00E-12	1,28	[1.19-1.37]
Rafnar T	18-01-2009	Basal cell carcinoma	5p15.33	TERT, CLPTM1L	CLPTM1L	rs401681-C	intron	0,56	4,00E-12	1,25	[1.18-1.34]
Nan H	22-06-2011	Basal cell carcinoma	13q32.3	UBAC2	HMGB3P4 - CCR12P	rs7335046-G		0,12	3,00E-08	1,26	[1.18-1.34]
Nan H	22-06-2011	Basal cell carcinoma	16q24.3	MC1R	MC1R	rs1805007-T	missense	0,07	4,00E-17	1,55	[1.45-1.66]
Nan H	22-06-2011	Basal cell carcinoma	6p25.3	EXOC2	IRF4 - EXOC2	rs12210050-T		0,17	1,00E-09	1,24	[1.17-1.31]
Stacey SN	24-09-2011	Basal cell carcinoma	17p13.1	TP53	TP53	rs78378222-C	UTR-3	NR	2,00E-20	2,16	[1.83-2.54]
Zhang M	31-03-2013	Non-melanoma skin cancer	6p25.3	IRF4	IRF4	rs12203592-T	intron	NR	7,00E-14	0,21	[0.15-0.27]
Zhang M	31-03-2013	Non-melanoma skin cancer	6p25.3	EXOC2, IRF4	IRF4 - EXOC2	rs12202284-A		NR	5,00E-08	0,15	[0.091-0.209]
Zhang M	31-03-2013	Non-melanoma skin cancer	14q22.1	GNG2	OR7E159P - GNG2	rs8015138-C		NR	7,00E-08	0,11	[0.071-0.149]
Stacey SN	08-01-2014	Basal cell carcinoma	20p13	TGM3	TGM3	rs214782-G	intron	0,17	6,00E-17	1,29	[1.22-1.37]
Stacey SN	08-01-2014	Basal cell carcinoma	8q22.2	RGS22	RGS22	rs7006527-?	intron	0,86	9,00E-13	1,3	[1.22-1.41]
Stacey SN	08-01-2014	Basal cell carcinoma	7q32.3	KLF14	LINC-PINT	rs157935-T	intron		9,00E-11	1,23	[1.15-1.31]

eTable 1. Candidate single nucleotide polymorphisms for basal cell carcinoma or non-melanoma skin cancer (*continued*)

First Author	Date	Disease/Trait	Region	Reported Gene(s)	Mapped gene	Strongest SNP-Risk Allele	Context	Risk Allele Frequency	p-Value	OR or beta	95% CI
Stacey SN	08-01-2014	Basal cell carcinoma	20p13	TGM3	STK35 - TGM3	rs59586681-T		0,61	3,00E-09	1,16	[1.11-1.22]
Stacey SN	08-01-2014	Basal cell carcinoma	9p21.3	CDKN2A, CDKN2B	CDKN2B-AS1	rs2151280-G	intron		3,00E-10	1,2	[1.14-1.27]
Stacey SN	08-04-2015	Basal cell carcinoma	2q33.1	ALS2CR12, CASP8, CASP10	ALS2CR12	rs13014235-C	missense	0,456	2,00E-09	1,15	[1.10-1.20]
Stacey SN	08-04-2015	Basal cell carcinoma	8q21.13	ZFHX4, ZFHX4-AS1	LINC01111 - MRPL9P1	rs28727938-C		0,938	4,00E-12	1,43	[1.30-1.59]
Stacey SN	08-04-2015	Basal cell carcinoma	12q13.13 / 5q11.2	KRT5	KRT5	rs11170164-T	missense	0,087	9,00E-09	1,29	[NR]
Stacey SN	08-04-2015	Basal cell carcinoma	2p24.3	MYCN, FAM49A	MYCN - FAM49A	rs57244888-T		0,898	5,00E-12	1,32	[1.22-1.43]
Stacey SN	08-04-2015	Basal cell carcinoma	10p14	GATA3, RP11-428L9.1, RP11-428L9.2	RNA5SP299 - LINC00709	rs73635312-G		0,874	2,00E-16	1,35	[1.25-1.45]

SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval.

eTable 2. Non-genetic characteristics of 633 Rotterdam Study patients with a primary BCC

Patient and tumor characteristics	Coding	Overall[1]	Superficial BCC	Nodular BCC
Number of patients		633 (100%)	137 (100%)	496 (100%)
Age at first BCC (years)	Median (IQR)	74.0 (67.3-80.4)	70.2 (64.3-76.0)	75.4 (68.2-81.5)
Sex	Female	339 (54%)	87 (64%)	252 (51%)
Pigment status	Dark	106 (17%)	24 (18%)	82 (17%)
	Intermediate	296 (47%)	70 (51%)	226 (46%)
	Light	151 (24%)	30 (22%)	121 (24%)
	Missing	80 (13%)	13 (9%)	67 (14%)
Easily sunburned	Yes	219 (35%)	53 (39%)	166 (33%)
	Missing	38 (6%)	5 (4%)	33 (7%)
Outdoor work	Yes	83 (13%)	14 (10%)	69 (14%)
	Missing	173 (27%)	42 (31%)	131 (26%)
Sun protection	No, never or hardly ever	234 (37%)	44 (32%)	190 (38%)
	Missing	34 (5%)	4 (3%)	30 (6%)
Smoking	Current or former	422 (67%)	92 (67%)	330 (67%)
	Missing	9 (1%)	1 (1%)	8 (2%)
Alcohol consumption (glasses/day)	Median (IQR)	0.5 (0.04-1.8)	0.6 (0.1-1.4)	0.5 (0.03-2.1)
	Missing	126 (20%)	18 (13%)	108 (22%)
Coffee consumption (cups/day)	Median (IQR)	3.3 (2.0-4.0)	3.3 (1.5-4.0)	4.0 (2.0-4.5)
	Missing	126 (20%)	18 (13%)	108 (22%)
>1 BCC at initial diagnosis	Yes	86 (14%)	24 (18%)	62 (13%)
Localization of first BCC	Head and neck	393 (62%)	24 (18%)	369 (74%)
	Extremities	101 (16%)	54 (39%)	47 (9%)
	Trunk	135 (21%)	58 (42%)	77 (16%)
	Missing	4 (1%)	1 (1%)	3 (1%)

[1] Participants with a mixed-type BCC with a superficial component were excluded.

BCC, basal cell carcinoma; IQR, interquartile range.

eTable 3. Genetic characteristics of 681 Rotterdam Study patients with a primary BCC

Patient and tumor characteristics	Coding	Overall[1]	Superficial BCC	Nodular BCC
Number of patients		681 (100%)	159 (100%)	522 (100%)
Age at first BCC (years)	Median (IQR)	72.2 (63.1-79.0)	68.0 (60.8-75.6)	73.3 (64.8-80.3)
Sex	Female	363 (53%)	103 (65%)	260 (50%)

[1] Participants with a mixed-type BCC with a superficial component were excluded.
BCC, basal cell carcinoma; IQR, interquartile range.

eTable 4. Associations between predictors and occurrence of superficial primary BCC (n = 633)[1]

Patient and tumor characteristics	Coding	Univariable models[2]	Multivariable model[2,3]
Age at first BCC (years)	Continuous	0.95 (0.93-0.97)***	0.96 (0.94-0.99)**
Sex	Female	1.69 (1.14-2.49)**	2.29 (1.36-3.83)**
Pigment status	Dark	Reference	Reference
	Intermediate	1.07 (0.63-1.80)	0.99 (0.52-1.86)
	Light	0.89 (0.49-1.63)	0.73 (0.35-1.52)
Easily sunburned	Yes	1.19 (0.80-1.77)	1.17 (0.71-1.95)
Outdoor work	Yes	0.77 (0.41-1.42)	0.99 (0.47-2.10)
Sun protection	No or hardly ever	0.72 (0.48-1.08)	0.89 (0.54-1.45)
Smoking	Current or former	1.00 (0.66-1.49)	1.33 (0.77-2.29)
Alcohol consumption (glasses/day)	Continuous	0.88 (0.75-1.04)	0.85 (0.70-1.04)
Coffee consumption (cups/day)	Continuous	0.92 (0.81-1.03)	0.92 (0.80-1.05)
>1 BCC at initial diagnosis	Yes	1.49 (0.89-2.49)	1.34 (0.72-2.52)
Localization of first BCC	Head and neck	Reference	Reference
	Extremities	17.34 (9.83-30.59)***	15.92 (8.80-28.80)***
	Trunk	11.38 (6.67-19.43)***	13.28 (7.40-23.82)***

[1] Compared to nodular BCCs only; all mixed-type BCCs with a superficial component were excluded.
[2] Pooled ORs with 95% CIs between brackets.
[3] Full model, no selection procedures used.
* P-value < 0.05; ** P-value < 0.01 ; *** P-value < 0.001.
BCC, basal cell carcinoma.

Model fit

Cox and Snell R^2:	25.0%
Nagelkerke R^2:	38.6%
Hosmer and Lemeshow Test:	0/30 had a p-value < 0.05

eTable 5. Associations between predictors and occurrence of first superficial BCC (n = 681)[1]

Patient and tumor characteristics	Coding	Univariable models[2]	Multivariable models[2,3]	Multivariable model[2,4]
Age at first BCC (years)	Continuous	0.97 (0.95-0.98)***		0.97 (0.95-0.98)***
Sex	Female	1.85 (1.28-2.68)**		1.95 (1.32-2.87)***
rs73635312	Yes	1.02 (0.67-1.56)	0.99 (0.64-1.53)	0.97 (0.63-1.51)
rs11170164	Yes	0.94 (0.59-1.48)	0.95 (0.59-1.53)	0.99 (0.60-1.61)
rs7335046	Yes	0.83 (0.57-1.22)	0.79 (0.53-1.18)	0.78 (0.51-1.19)
rs8015138	Yes	0.78 (0.61-1.01)	0.81 (0.63-1.06)	0.83 (0.64-1.08)
rs1805007	Yes	0.97 (0.59-1.59)	0.93 (0.56-1.55)	0.92 (0.54-1.55)
rs78378222	Yes	0.77 (0.31-1.92)	0.82 (0.32-2.08)	0.84 (0.32-2.19)
rs7538876	Yes	1.06 (0.82-1.38)	1.03 (0.78-1.34)	1.03 (0.78-1.36)
rs801114	Yes	0.94 (0.72-1.22)	1.01 (0.77-1.33)	0.99 (0.75-1.32)
rs214782	Yes	0.91 (0.67-1.23)	0.93 (0.68-1.26)	0.95 (0.69-1.31)
rs13014235	Yes	1.05 (0.81-1.35)	1.04 (0.80-1.35)	1.05 (0.80-1.38)
rs57244888	Yes	0.96 (0.59-1.57)	1.02 (0.62-1.69)	1.05 (0.63-1.76)
rs401681	Yes	1.00 (0.78-1.29)	0.93 (0.72-1.21)	0.92 (0.70-1.21)
rs12203592	Yes	1.50 (0.97-2.33)	1.50 (0.94-2.32)	2.11 (1.25-3.58)**
rs12202284	Yes	0.79 (0.53-1.17)	0.73 (0.49-1.10)	0.55 (0.35-0.88)*
rs12210050	Yes	0.88 (0.61-1.26)	0.86 (0.60-1.24)	
rs157935	Yes	0.97 (0.73-1.28)	0.95 (0.71-1.27)	0.95 (0.71-1.29)
rs28727938	Yes	0.87 (0.49-1.53)	0.74 (0.41-1.32)	0.70 (0.38-1.27)
rs7006527	Yes	0.76 (0.53-1.09)	0.72 (0.50-1.05)	0.72 (0.49-1.06)
rs2151280	Yes	1.15 (0.90-1.47)	1.15 (0.89-1.49)	1.18 (0.91-1.53)
rs59586681	Yes	1.07 (0.82-1.38)	1.10 (0.84-1.44)	1.07 (0.81-1.42)

[1] Compared to nodular BCCs only; all mixed-type BCCs with a superficial component were excluded.

[2] Odds ratios with 95% confidence intervals between parentheses.

[3] Included one SNP at a time, adjusted for age at first BCC, sex and first 4 principal components.

[4] Full model, adjusted for age at first BCC, sex and first 4 principal components. No selection procedures used.

* P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001.

BCC, basal cell carcinoma.

Model fit full multivariable model

Cox and Snell R^2:	6.9%
Nagelkerke R^2:	10.4%
Hosmer and Lemeshow Test:	p-value 0.802